Tutorial

Machine Learning

Table of contents

1. General Information

This tutorial describes some application examples of Machine Learning tool of KOMVOS. The Machine Learning functionality is implemented in KOMVOS and powered by the ML Toolkit.
By using the Machine Learning (ML) functionality it is possible to predict the impact of Design Changes in the early design stages of product development. This can be done by using two types of Machine learning algorithms that are available.

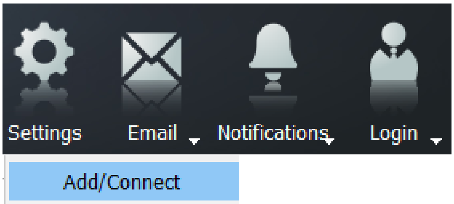
• DATA Driven Design Variable based algorithm: Trained based on a model's design variable values and selected responses. The Design Variable based training uses machine learning algorithms to create a predictive model that is able to predict single scalar values, 2d curves and 3d field results.
• Simulation Driven Feature based algorithm: Trained based on the Finite Element model (Body in white) and first torsional and first bending modes frequency values. The Feature based ML training performs a feature extraction process on the finite element models that are the input during the training process. The extracted features contain information related to mass, materials, shape, etc. of the finite element model. The trained machine learning model is able to predict single scalar values, related to the first torsional and bending mode frequency.
Machine learning requires Training of a given "dataset" that consists of existing models (experiments) and their responses. Such datasets are created by ANSA's Optimization tool that can create a Design of Experiments (DOE) and automatically save them as Simulation Runs in the selected DM, or imported using CSV files. Utilizing the DM structure of Simulation Models and Simulation runs, ML training can be performed on selected DOE Studies, which will create a machine learning predictive model, known as Predictor.
Such predictors are able to predict a theoretical model's behavior (keyvalue, 2d plots, 3d results) by changing design variable values (DV Based) or using a new FE model design (Feature based).
To use the Machine Learning functionality, the ML Toolkit has to be downloaded and installed, and the ML_SERVER license feature needs to be enabled upon request. The ML Toolkit installs an environment that contains all the required latest libraries to perform Machine learning (Tensorflow, Pandas, Numpy, etc.

The files required for this tutorial reside inside the following directory:

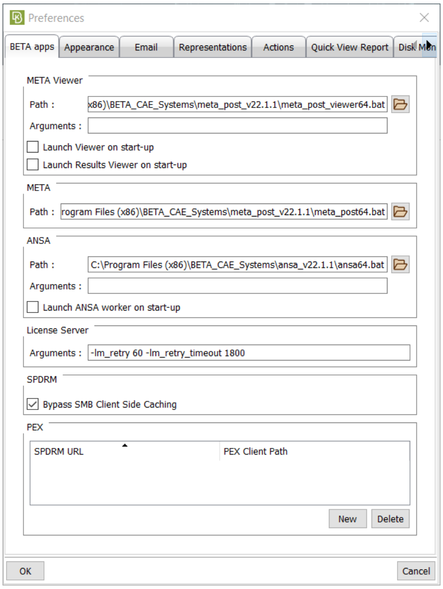
```
/komvos_vxx.x.x_tutorials/Machine_Learning/tutorial_files/
```

1.1. Starting with KOMVOS

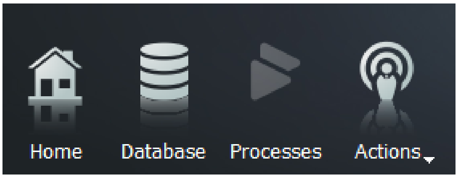
KOMVOS is a solution for interactive browsing, visualization and handling of Simulation Runs, Key Results and Reports. KOMVOS provides access to data that exists in corporate solutions such as small scale file based ANSA-DM or even SPDRM or any other 3rd party SDM system. In this case KOMVOS will connect to the DM created earlier, to access the Simulation Model and Simulation runs created by the Optimization tool.



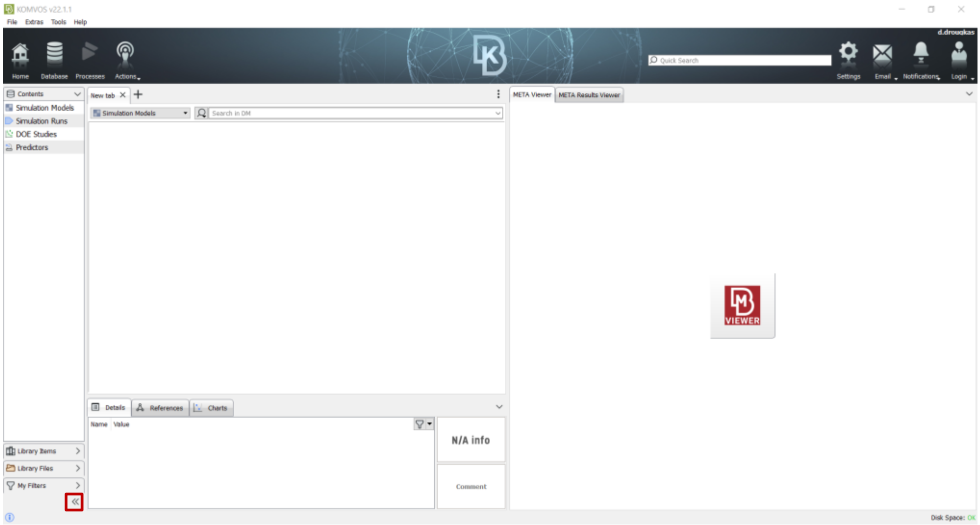
To connect with the DM, press the Login icon and select **Add/Connect.**
In the File Browser that appears, select a DM folder that contains the data, Simulation models and Simulation runs, created by ANSA.



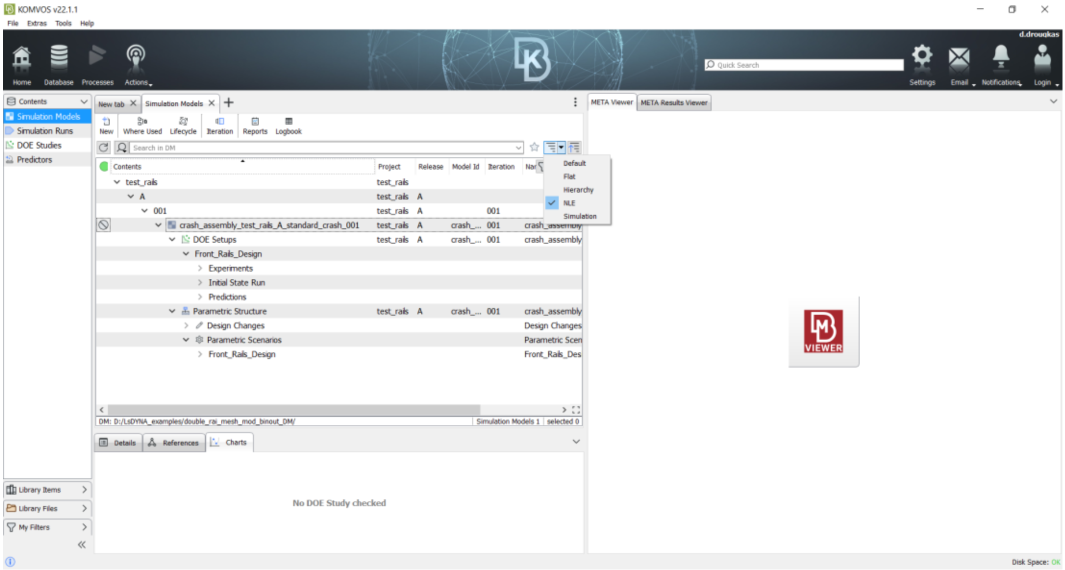
In order for KOMVOS to be able to perform several actions, including Machine Learning Training, and prediction, the BETA apps paths need to be completed in the respective fields. Open the settings by pressing the respective icon and select the paths for the BETA apps.



The Database icon will open the database page of KOMVOS where the handling and overview of the DM takes place.
Pressing the button will open the database page.



On the left side, the contents and other items that exist in the DM can be selected. This toolbar can be reduced and expanded by the relative button at the lower part of the screen.
In the Contents it is possible to see and navigate to the lists of any object that exists in the DM.

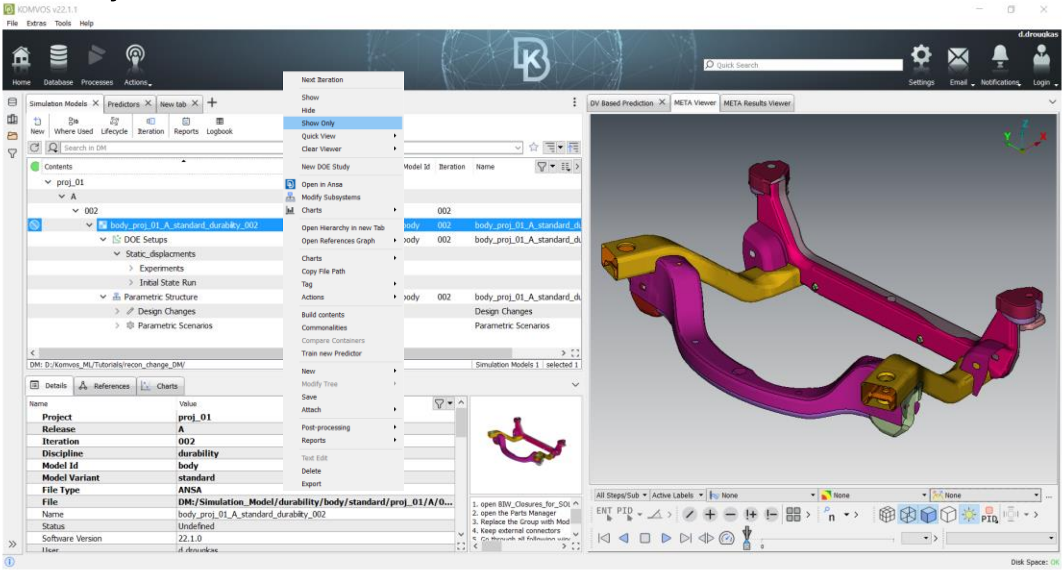


In order to access the full functionality of the Machine Learning, the NLE view of the Simulation models should be used when inspecting, and handling Simulation models and Simulation runs. On the right side of the KOVMOS user interface, META Viewer windows can open in order to view models, parts, reports, images, videos, etc. and explore the available data.

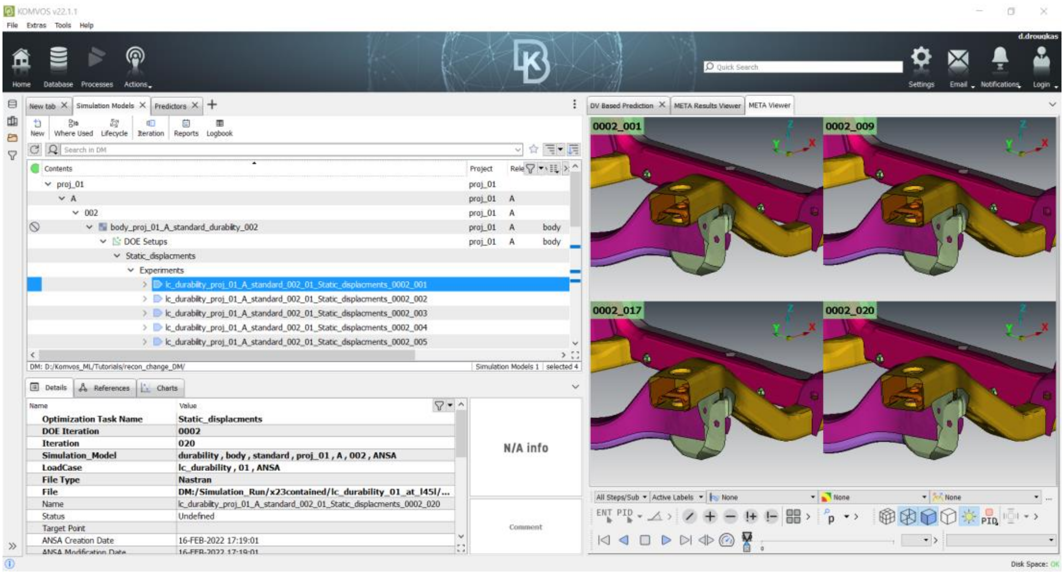
1.2. Parametric model viewing and handling

The Simulation Models that are created from ANSA's Optimization tool contain the DOE Setups and Parametric Structure items. These contain the created experiments and the information about the Design Variables of the model, respectively. These items are visible only in the NLE view of the Simulation models.

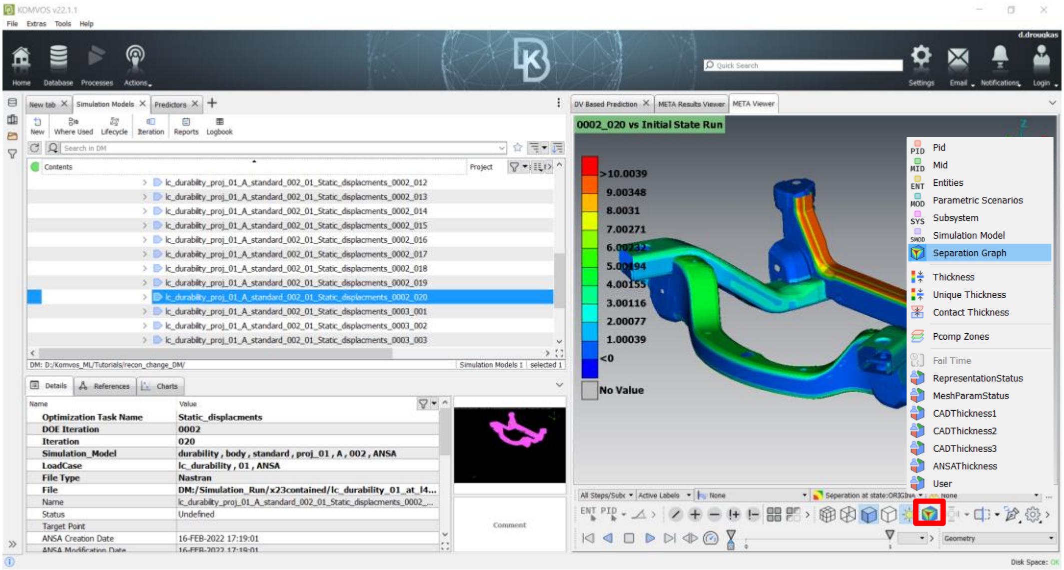
In order to see the model, select the simulation model and with right click select **Show in Viewer** or **Show only**



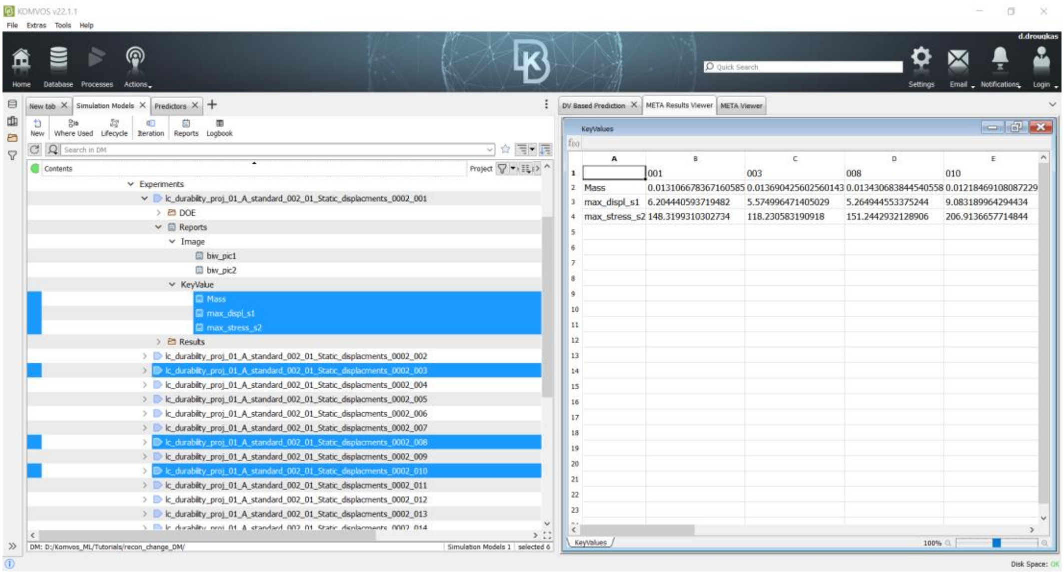
In order to show the geometry of each experiment, expand the Experiments item and click on each experiment. The selected experiment geometry will appear in the META Viewer window.
To compare mulitiple experiments in the Viewer, hold Ctrl button and pick up to four experiments.



In case there are geometry modifications between the models, in order to compare the geometry of each experiment with the initial model, there is separation graph option that creates a separation graph between the selected experiment and the initial state.

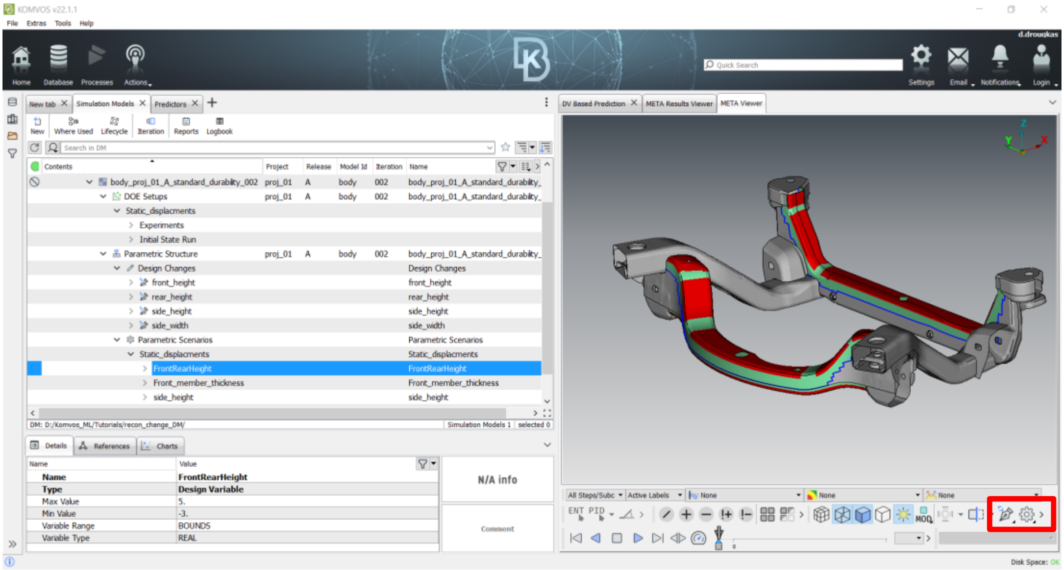


Each experiment can contain Report items from its post processing. Such items can be Images, KeyValues, Tables, Videos, etc. These can be viewed in the Results Viewer when selected, and it is possible to compare values from multiple experiments by selecting the report item, KeyValue for example, and while holding the Ctrl button, selecting the other experiments.



The parametric structure item under the Simulation model contains the Design Changes the model may have and the Parametric Scenarios. The Design Changes are the direct morphing parameters that are being used by design variables in the workflow of the Optimization Tool in of the ANSA model. The Parametric scenario is a description of the Optimization workflow in the ANSA model that created the Simulation model.

Selecting the Design changes or Parametric scenario, the relative parameters and their details appear in the model and respectively in the details tab.



2. Design Variable Based Machine Learning

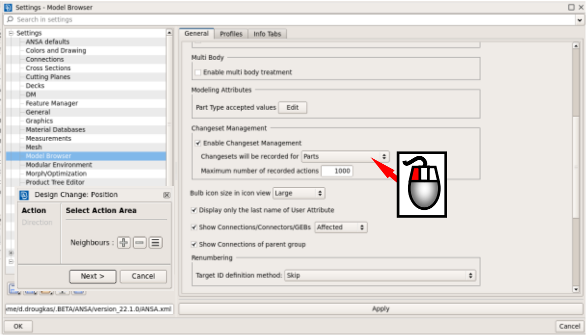
2.1. Introduction

Design variable based predictors are advanced data driven machine learning models that can be used for any parametric study or exploration. DV based predictors can be trained using as input the design variable values of a parametric model and its responses. These can be single scalar values (KeyValue) 2d curves and 3d field results. Prediction of the responses is done through the DV Based prediction window, where values can be applied on the relative Design Variables, in order to define a “what if” scenario and predict the responses.

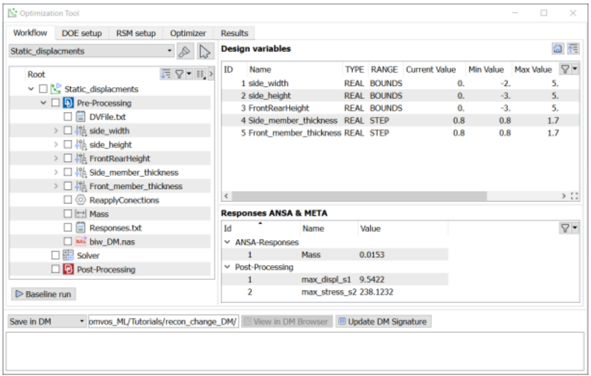
2.2. Creating the dataset

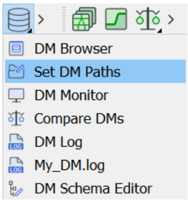
2.2.1. Model preparation in ANSA

In order to create the required training dataset, several Simulation runs need to be created and added in a DM. This is done through the ANSA Optimization tool.

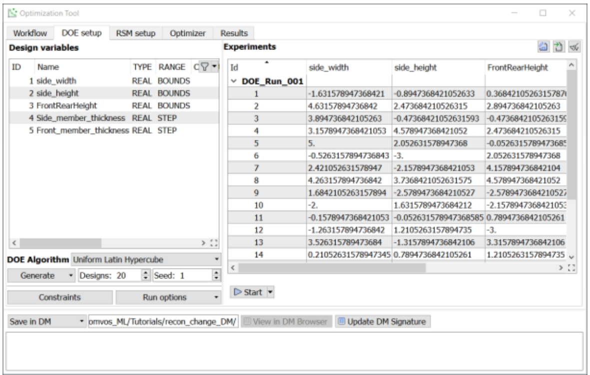


In order to allow for training of 3D results on models with mesh modifications, **enable Changeset Management for Parts**, from the:
Tools> Settings> Settings> Model Browser > General menu.



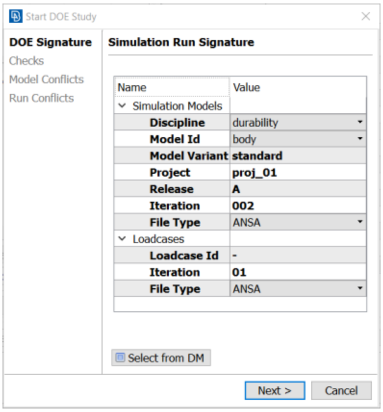


Once a DM path is defined, it is possible to define a complete Optimization task with Solver and post processing item, and a DOE Setup. Run a baseline run to make sure the workflow is running without an error and any desired responses are created.

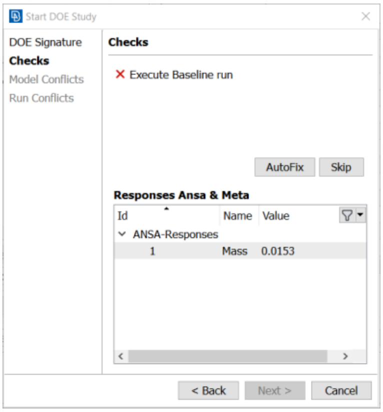


In the DOE setup tab, the Option Save in DM will automatically save a new Simulation model and as many Simulation runs as are the defined experiments. Select a desired **DOE Algorithm** and Define a number of Designs and Seeds.
Press **Generate** to create the Experiments.

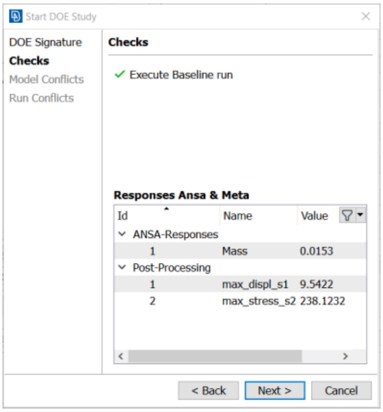
NOTE! Instructions on how to create an optimization task with Solver and Post Processing item with the session in DM, are described in the Optimization with ANSA & META tutorial.



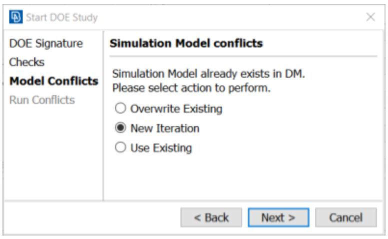
Press **Start** and type in the “Signature” of the Simulation model and Simulation runs.
Select the attributes of the model and the load-case.



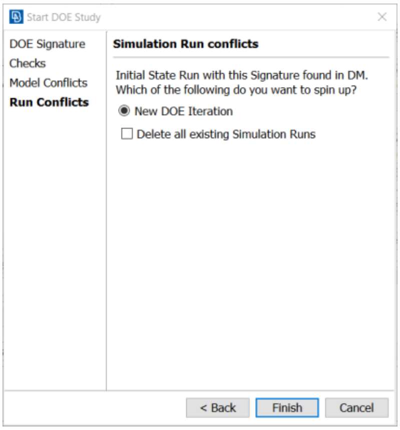
The next step, **Checks** if the Baseline Run was not executed, or it failed, the process will not be allowed to continue. Press the **Autofix** button that will start a Baseline Run. If the **Next** button is not activated after the run is completed, some correction actions need to be performed in the workflow. If the **Next** button is activated, press it to continue.
In case the workflow is correct and the Baseline check does not pass, **Skip** button can be used, to omit this step, stating that the workflow is correct.



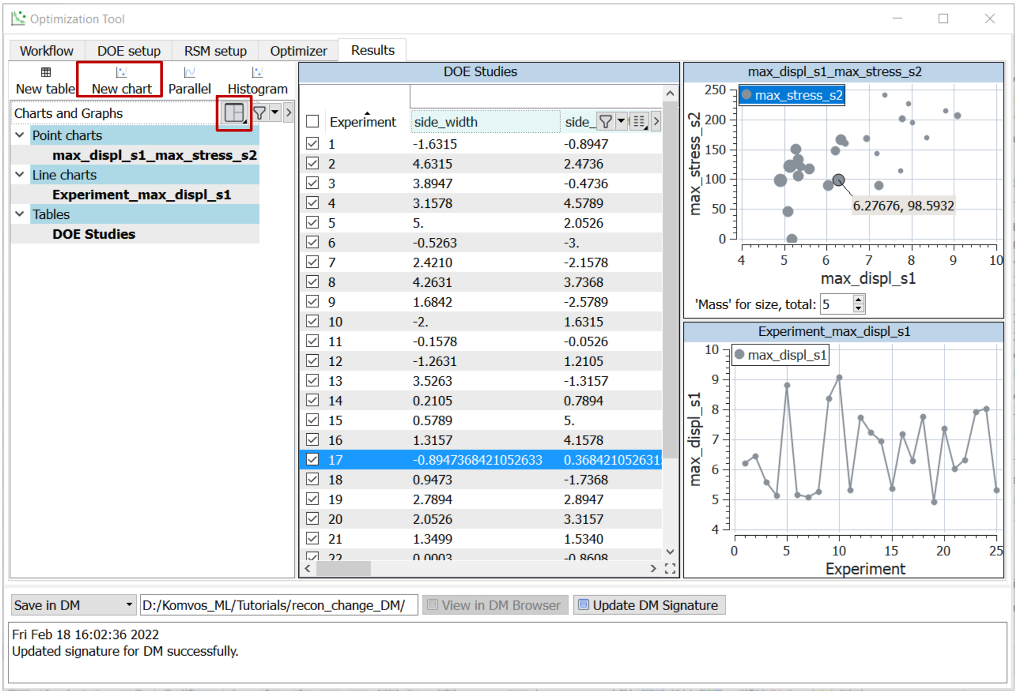
If the Checks step identifies that the Baseline run was finished successfully, all responses will appear in the Responses table. Press **Next** to continue.



In the next, the tool checks if there are existing Simulation Models in the DM with the same name and other attributes.
If there are, the options of **Overwrite**, **New Iteration** and **Use existing** are available for selection.
Overwrite, will delete the existing Simulation model in the DM and create a new one.
New Iteration will create a new iteration of the Simulation model, keeping the old one.
Use existing, will not update the Simulation model. It will keep the existing and not save a new one.

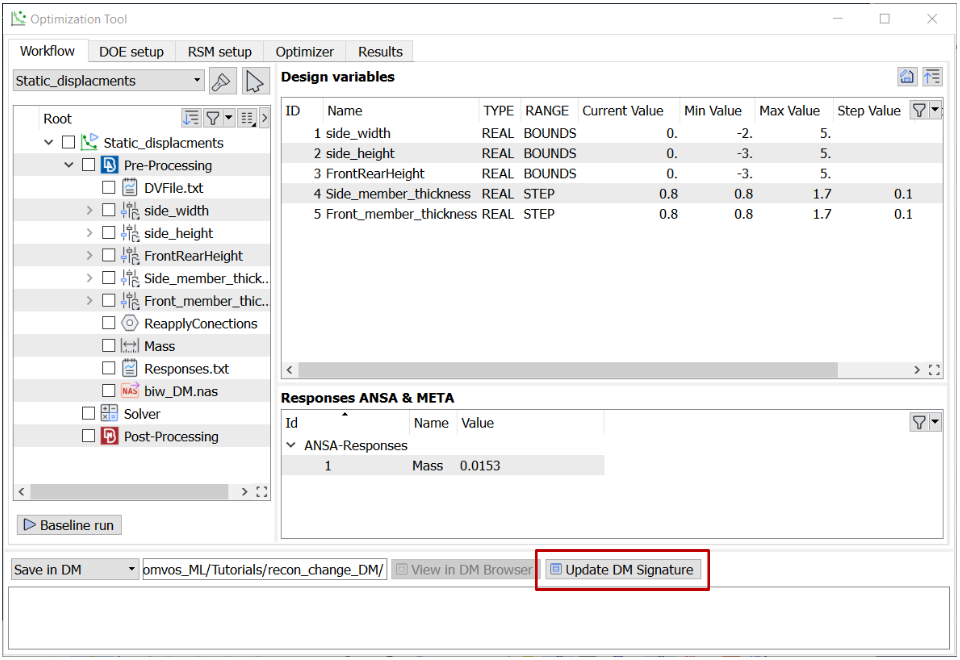


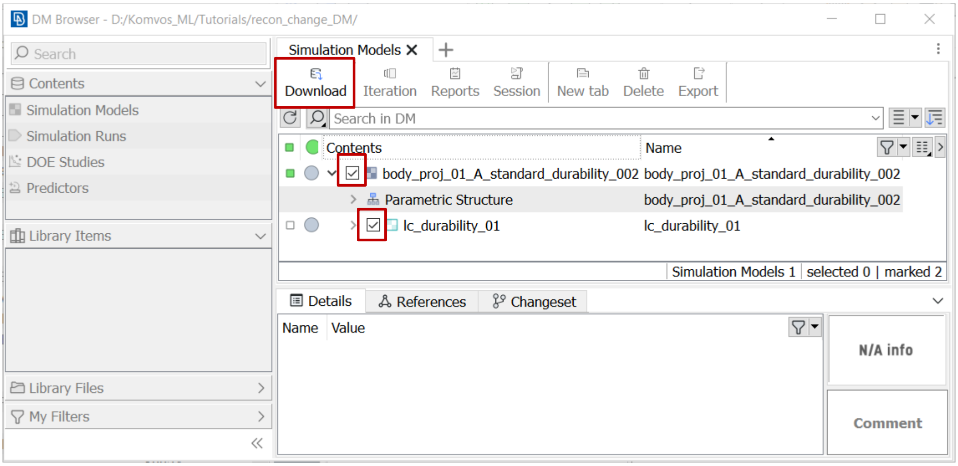
The next page, checks if there are existing Simulation Runs in the DM with the same attributes. If there are, dependting on the selection of the Simulation model, it will ask to save the new Simulation Runs under a new DOE iteration, with the option to Delete existing Simulation runs, to avoid the conflict, or aggregate the new experiments to an existing DOE iteration.
Pressing Finish, will start running the DOE.



Once the DOE run is finished, the results can be viewed in the results tab. The charts and tables are in the Results tab are saved, when the ansa file is saved. Charts can be created through the **New Chart** button and once created can be draged and droped from the **Charts and Graphs** list to the Layout on the right of the screen. The layout can also be modified by selecting one from the relative icon.

In case a saved ANSA file is opened, and a DM that contains existing Simulation run is defined, in order to view the responses and the simulation runs in the Results tab, use the **Update DM Signature** button and select the desired simulation model from the DM browser window that opens up.

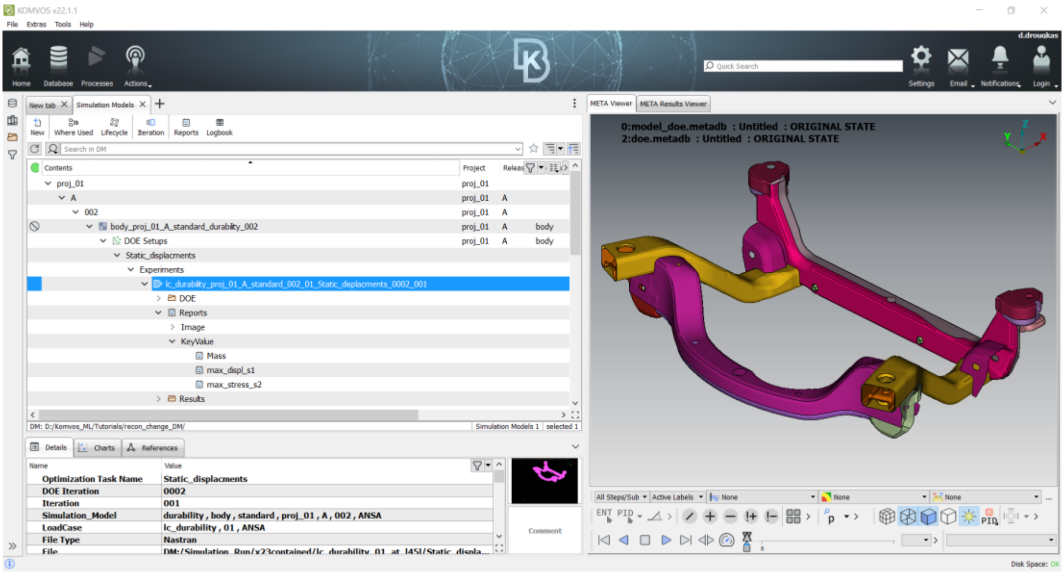


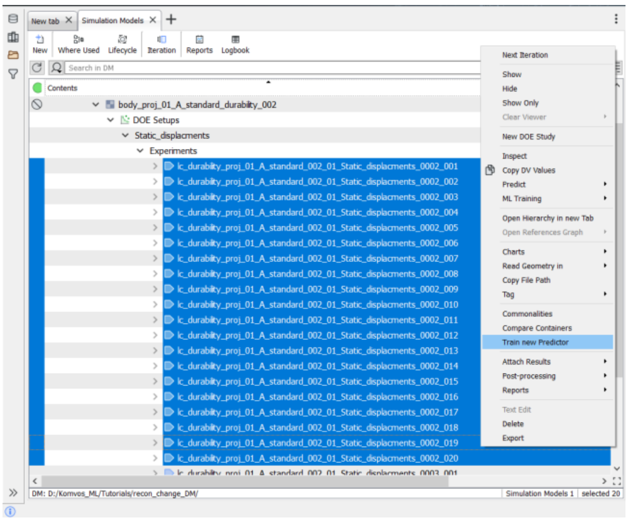


Then press **Download** to update the information in the Optimization tool.

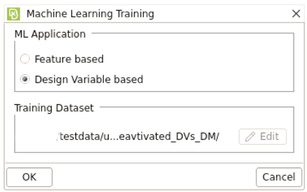
2.3. Machine learning Training

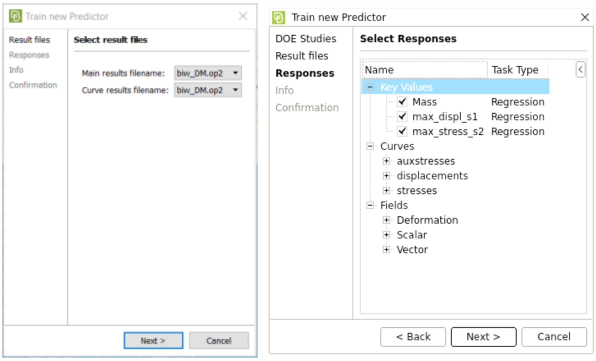
After connecting KOMVOS to the previously created DM, open the Simulation models list in the Database. Each experiment under the DOE Studies contains the DOE information, Reports and Results. The KeyValues that appear under the Report items can be used for ML Training. Each of these KeyValues can be used, in order to create a machine learning model that will predict this response for any values of the Design Variables.
In a similar manner, the result files that are available in the Results folder of each experiment will be able to be used in order to train machine learning models that can predict 2d (curves) and 3d (field scalar, and vector) results.
Once a Machine Learning model is trained, a new entity called Predictor will be created and saved in the DM.



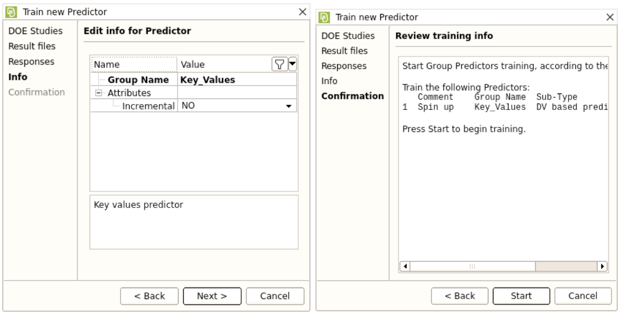


If there are no Predictors available in the DM, in order to start the ML training, select the Experiments that will be used for the training and with right click select **Train new Predictor**. At least 10 experiments need to be selected in order to start the ML training. For this tutorial, leave out of the training 3 simulation runs, to be used for prediction validation in the Prediction chapter.
In the Machine Learning Algorithm window that appears, select **Design Variable based**.

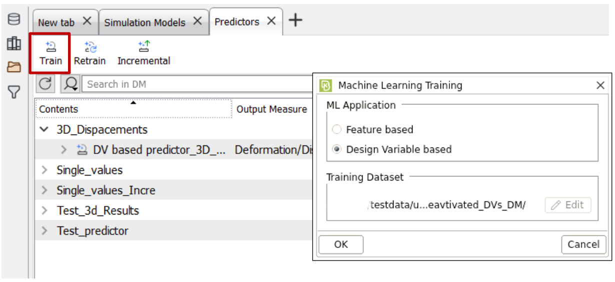




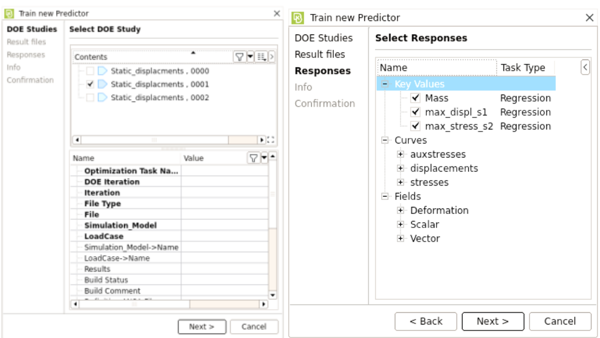
In the wizard that appears, in the Result files page, if there are result files in each experiment, a selection of main and curve results files appears. This selection will provide the 3D and 2D results that can be used for training.
In the Responses page, it is possible to select Key Values, Curve, Field, Scalar and Vector results to train for. In case the values are non-numerical, the predictor will be classification type. Otherwise it will be a regression type.



In the Info page, provide a name (Key_Values) and optionally some comments for the Predictor that is about to be created.
Finally confirm and begin the training by pressing the **Start** button.

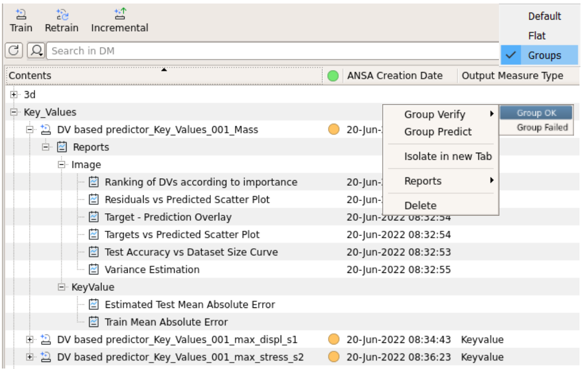


If predictors already exist in the DM, a new training can start from the Predictors Tab. Once the Train button is pressed, the Machine learning Algorithm window appears.
Select **Design Variable** based.

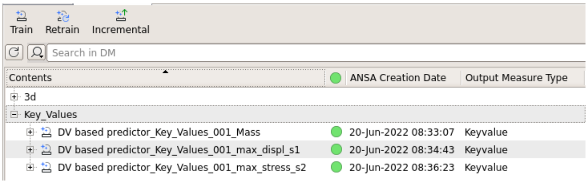


The new Predictor wizard starts, this time requiring a DOE Study in the first page. Select a DOE Study that contains more than 10 experiments and continue with the next pages as before, this time selecting the three Key Values as Responses.
Name this Predictor Key_Values and start the training.

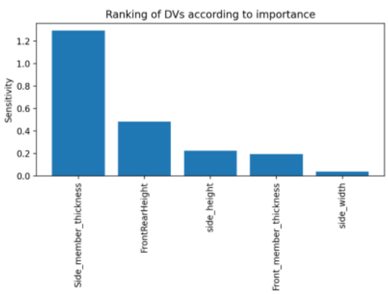
During the training, the training status can be seen in the Home tab.

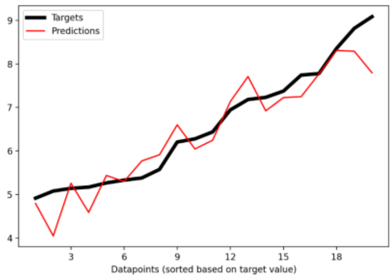


Once the training is completed, the Predictor entities will appear in the Predictors list.
Predictors appear in Groups (Groups view). Each group contains as many predictors, as the number of responses that were selected during training definition.
Predictor Status in Orange means the predictor needs to be verified. Failed predictors cannot make predictions.

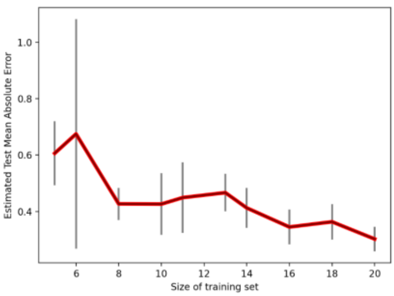


Each predictor contains Report items that provide indicators about the predictor's performance, and the Estimated Mean Absolute Error of the Prediction.
More information about these reports can be found in KOMVOS User's Guide in the Machine Learning chapter, **Predictor Reports**.





* Based on test predictions obtained using nested cross-validation
(due to small dataset size)

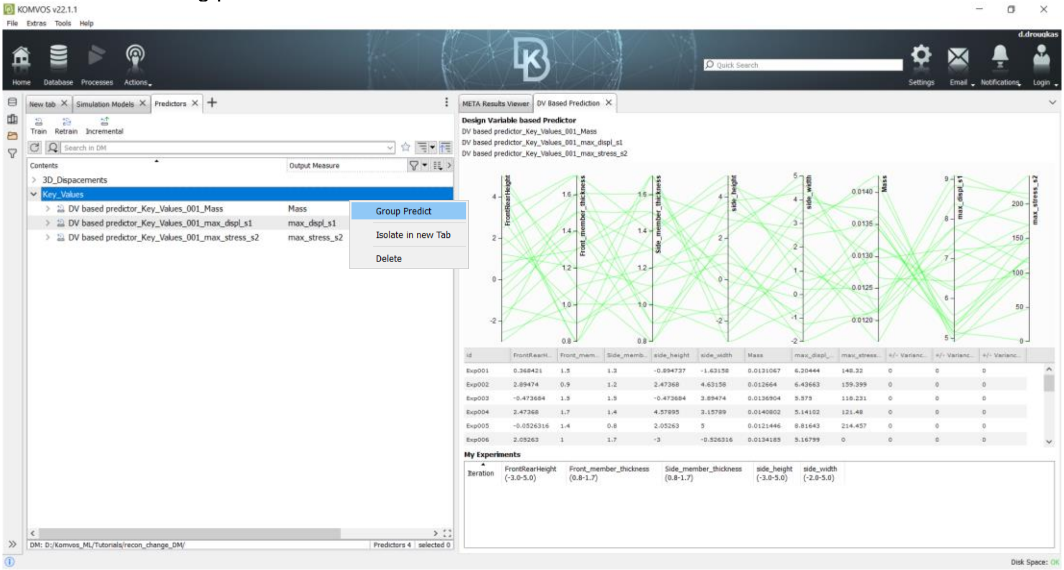


* Test Mean Absolute Error was estimated using nested cross-validation
(due to small dataset size)

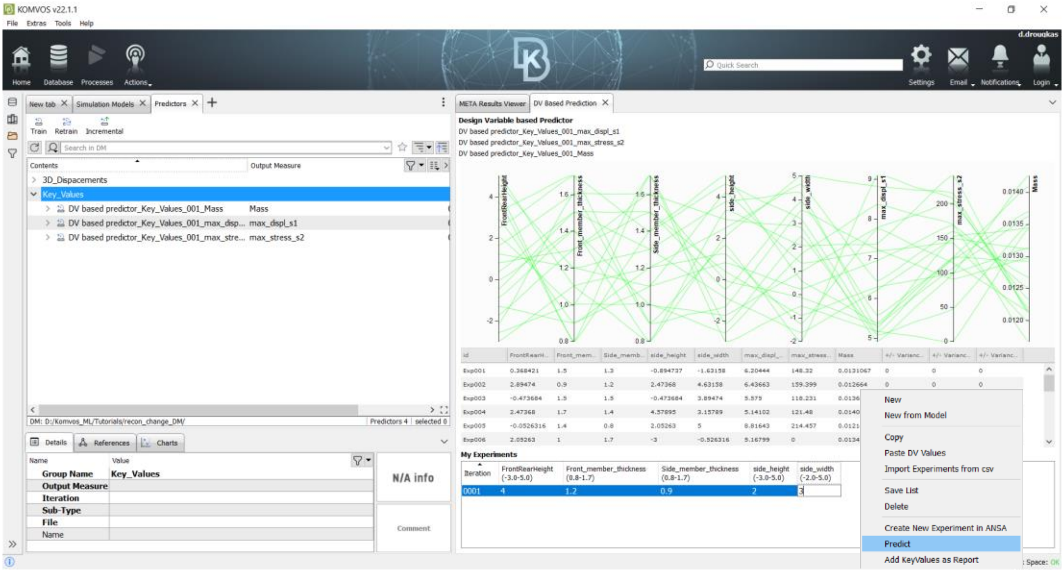
2.4. Machine Learning Prediction

2.4.1. Key Values prediction

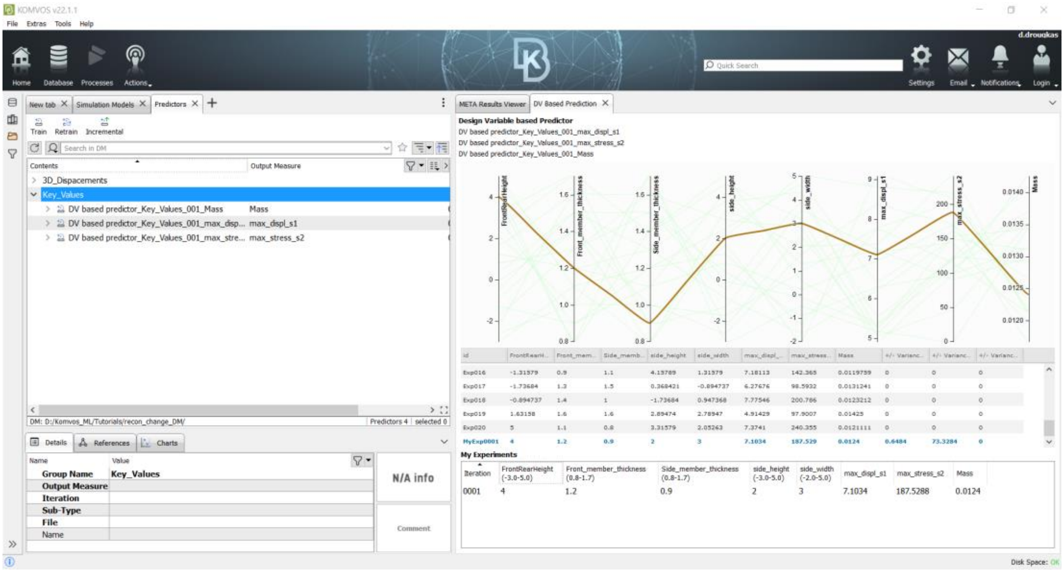
While viewing the Predictors tab in Groups view, right click on the Group name and select **Group Predict**. This will open the DV Based Prediction window on the right side of the window. The Prediction window contains a parallel coordinates chart, showing the Design Variables and the trained Responses as columns. Bellow there is a list of experiments that were used in order to train the predictors. These experiments are already shown in the parallel coordinates chart. The **My Experiments** list below is used in order to create new experiments by typing design variable values, and making predictions.



Right click in the **My Experiments** list and select **New**, to create a new theoretical experiment. The initial values that appear for each design variable are the values of the Initial state run.

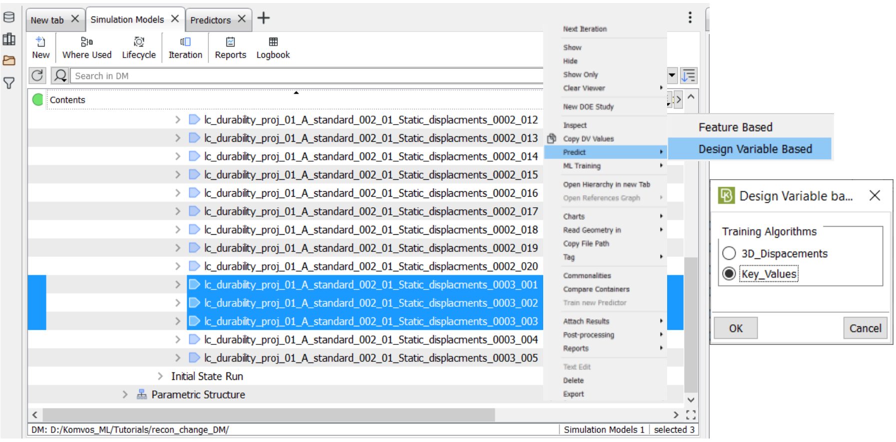


Type in some values, select the experiment line and with right click, select **Predict**

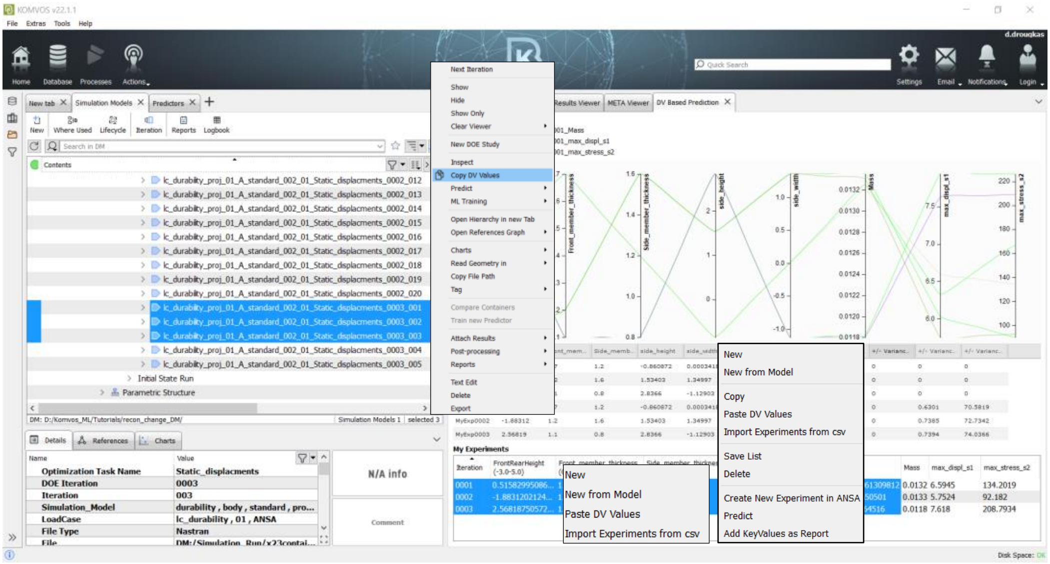


A new experiment is added in the list and in the chart. The predicted values and their variances are added in the respective columnts and there is a new line in the chart that represents this experiment.

Such predictions can be done also for experiments that exist in the DM. From the Simulation Models tab, select three experiments that were not used for the training of the Predictors and with right click select **Predict>Design Variable Based**. Pick the **Key_Values** predictor and press OK.

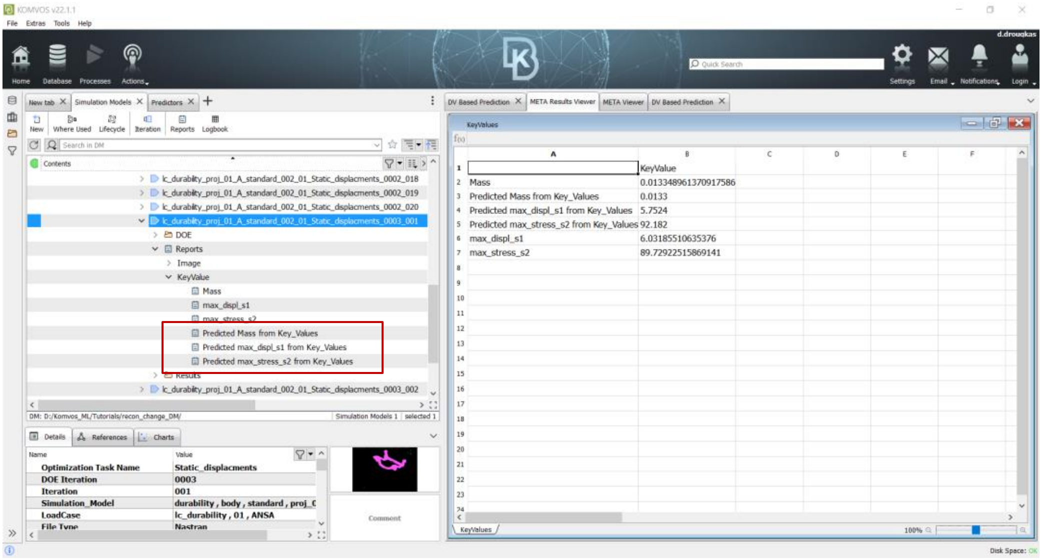


After the Prediction is done, the results and experiment lines will be added in the Parallel Coordinates chart.



In a similar way select the experiments and with right click select the option **Copy DV value**s. The Design Variable values of each selected experiment are copied to clipboard. Then in the **My Experiments** list of the Prediction Window, right click on the empty space and select the Paste **DV values** option. The experiment values are added in the list and selecting the **Predict** option will start the prediction.

Once the prediction is done and the results appear in the list, it is possible to create Report items for these KeyValues and add them as Reports under the existing Simulation Run. Select the lines in the **My Experiments** list and with right click select **Add KeyValues as Reports**. New KeyValue report items are created under each relative experiment.

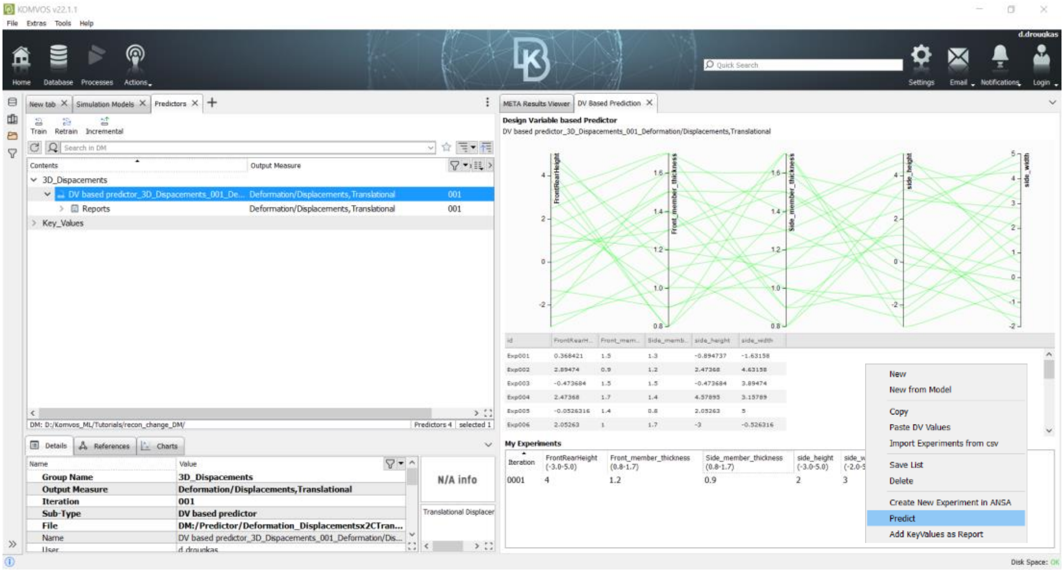


The experiments in the **My Experiments** list can be saved as text (*.txt) or CSV (*.csv) files with the option **Save list**. Similarly such a list can be imported with the option **Import Experiments from csv**.

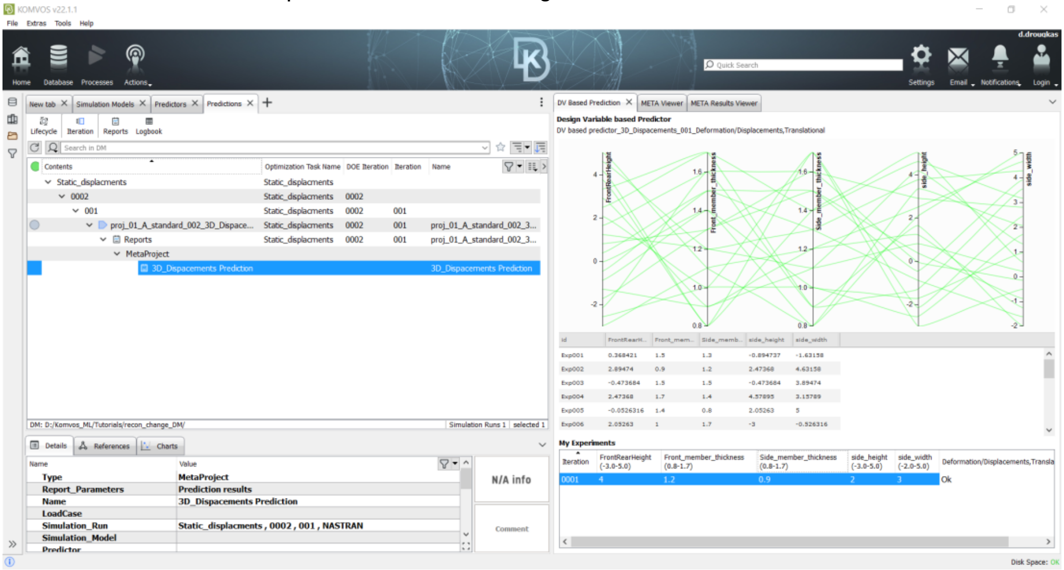
From a selected experiment in the list it is also possible to **Create new Experiment in ANSA**. This action will open the baseline simulation model in ANSA and apply the selected DV values in the optimization tool's DOE tab, ready to create a new experiment.

2.4.2. Curves and Field results prediction

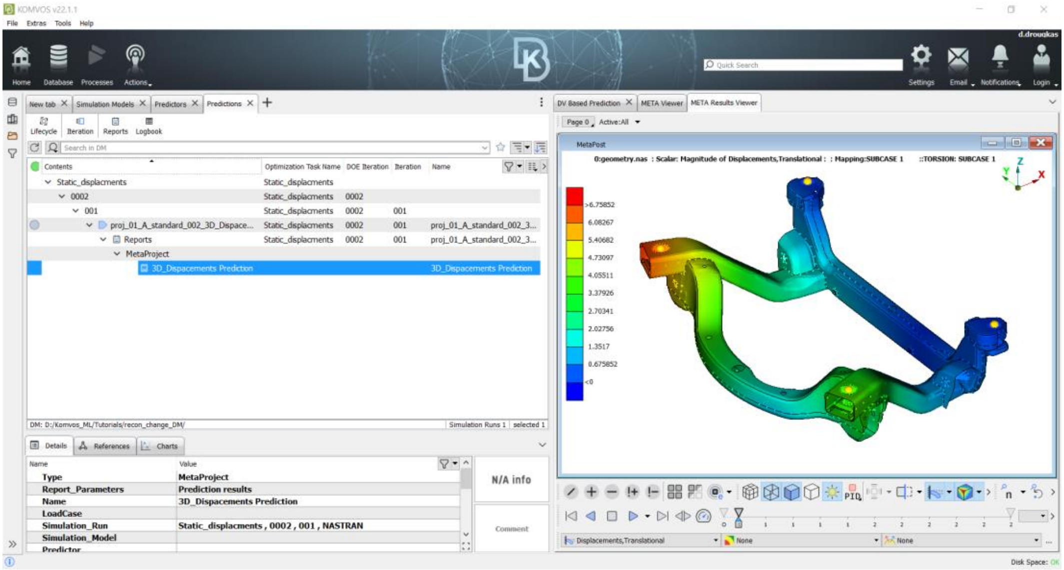
Similar to the prediction of the Key Values, while viewing the Predictors tab in Groups view, right click on the Group name and select **Group Predict**. This will open the DV Based Prediction window on the right side of the window. The Prediction window contains a parallel coordinates chart, showing the Design Variables and the trained Responses as columns. Bellow there is a list of experiments that were used in order to train the predictors. These experiments are already shown in the parallel coordinates chart. The **My Experiments** list below is used in order to create new experiments by typing design variable values, and making predictions.



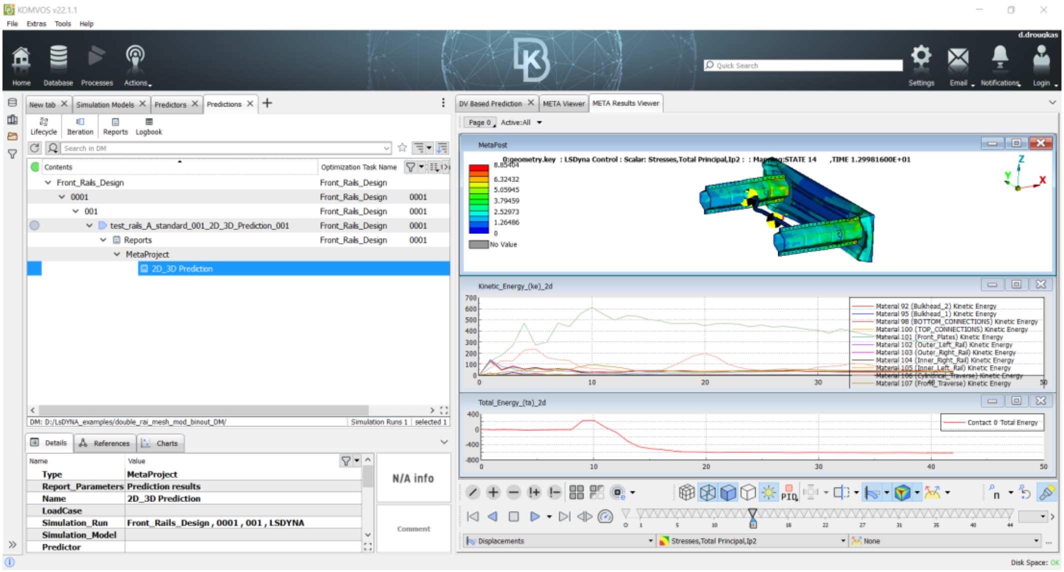
Right click in the **My Experiments** list and select **New**, to create a new theoretical experiment. The initial values that appear for each design variable are the values of the Initial state run. Type in some values, select the experiment line and with right click, select **Predict.**



Once the prediction is done, an **Ok** will appear in the list as results. This is normal as there is no single value that can be calculated. At the same time there is no result column in the parallel coordinates chart. However, in the left side of the window a new tab called Predictions appears a new Prediction simulation run is created. This simulation run contains a new Report item that contains the prediction as a META Project that can be viewed in the META Viewer



In case the Predictor is trained for more 3d field results, like stresses, etc., the labels in the META Viewer are active and can switch between results.
In a similar manner, if a Predictor was trained for 2D results (curves), the prediction will create a META Project Report item that can also be viewed in the META Viewer as it appears below.



Such META Project files can open in META, in order to compare with a validation model.

3. Feature Based Machine Learning

3.1. Introduction

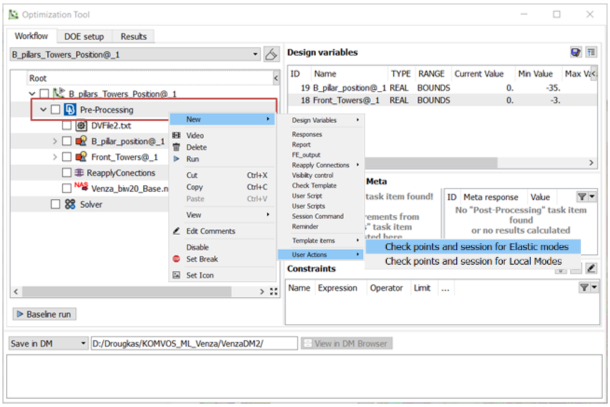
Feature based Machine Learning uses as input the Finite element model and a selected response and performs feature extraction. Through this process specific features of the FE model are captured and train the Machine learning model. Feature based predictors are tuned to be trained and predict First Torsional, Vertical and Lateral Bending modes frequency values of a finite element body in white. No design variables are needed in order to train this type of Predictors.

3.2. Creating the dataset

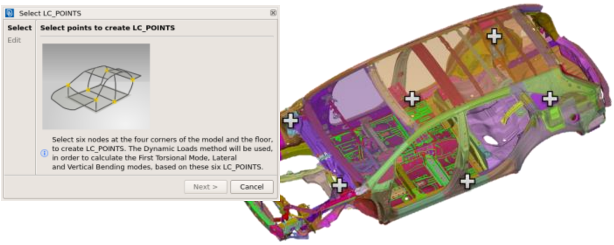
As described, Feature based Machine Learning does not require a parametric model in order to learn. Finite Element Body in White models and their responses should be used as a dataset in order to train the Predictor. This dataset needs to be added in the DM in order for KOMVOS to be able to collect the data and perform the training. The process of adding the models in the DM can be done with two different ways. One way is to prepare the model in ANSA, using the Optimization tool and design variables, in order to create the variations of the model. In this case, even though Design Variables are used in order to create the “different” design of the model, the purpose is just to have the FE Body in White models and their responses in the DM.
The other way is to have data that are not created by ANSA documented in an informative CSV file. The Import DOE Study tool will be able to import the data in the DM so that Machine Learning training can be performed.

3.2.1. Model preparation for elastic modes in ANSA

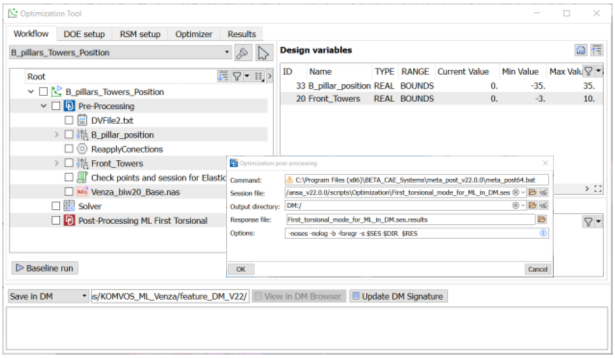
In order to perform Feature based Machine Learning, specific actions need to be taken in the preparation of the model in ANSA. This is required in order to create the required responses of First Torsional mode, vertical and lateral bending model frequency values.
The model is subjected to a 103 Normal modes analysis and a specifically defined post processing session is applied in order to identify the required responses. The session is using named Interface points which will pin point critical positions of the Body In white in order to identify the first Torsional and Bending mode frequency values.
A wizard is available in order to facilitate the creation of these interface points.



Once a DM path is defined, it is possible to define a complete Optimization.
The option Save in DM must be defined in order to use the DM to save a new Simulation model and as many Simulation runs as are the defined experiments.
A dedicated User action is available and when it runs, it will check if the appropriate interface points are available. Once the item is added in the task, move it before the FE Output item. Select it and with right click **Run** it will check if the model is prepared, or will start the wizard if it is not. The six interface points need to be on the strut tower points and on stiff areas closer to the middle of the body in white. Interface points on panel areas are not recommended.



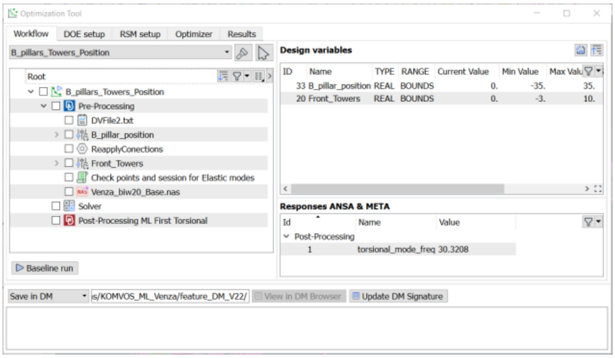
The created LC_Points will be named according to their position. F_LHS: Front left hand side, M_RHS: Middle right hand side, etc. If the given names do not match the position, they should be manually renamed accordingly.



Once the LC_Points are created, a new post processing item will be automatically created, using a specific post processing session that will produce the needed frequency value responses.

Select right click **Reset** on the User item, to reset its status.

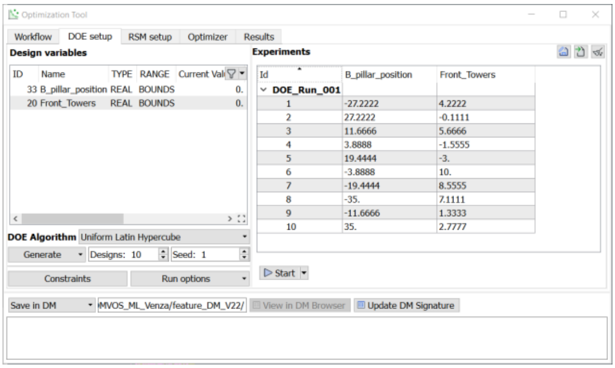
If four LC_Points are used at the corners of the BiW, they are enough to get the First Torsional mode value response. With the additional two middle points, first Lateral and Vertical Bending modes frequency values can be identified.



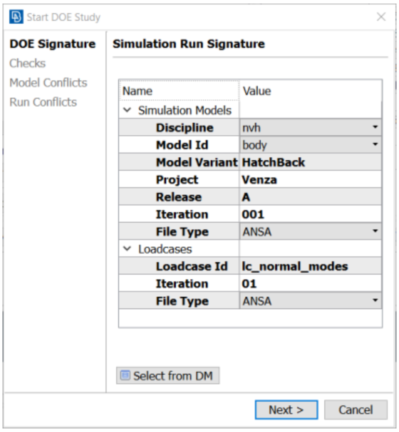
Once the workflow is ready, press the **Baseline Run** button to start the baseline analysis.

If the workflow is correct and the analysis finishes successfully, the responses will appear in the Responses area of the Optimization Tool.

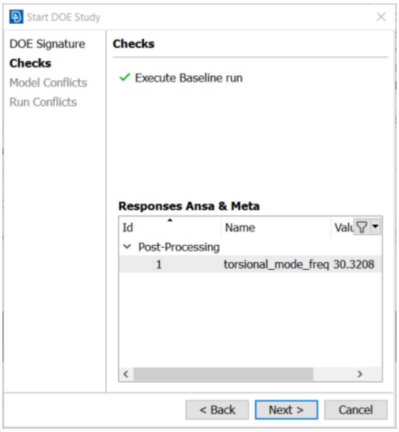
In case the Interface points are not defined, or the specific post processing item is not available, the Baseline Run or DOE will stop with an error.



To continue to the creation of the dataset, switch to the DOE Setup tab and generate Designs using one of the available DOE Algorithms.

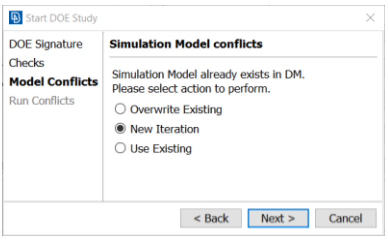


Press **Start** and type in the "Signature" of the Simulation model and Simulation runs.
Select the attributes of the model and the load-case.

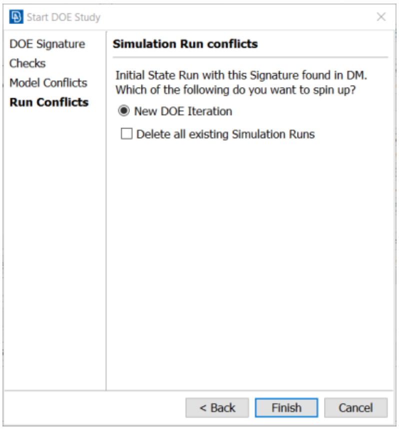


The next step, **Checks** if the Baseline Run was not executed, or it failed, the process will not be allowed to continue. Press the **Autofix** button that will start a Baseline Run. If the **Next** button is not activated after the run is completed, some correction actions need to be performed in the workflow. If the **Next** button is activated, press it to continue.
In case the workflow is correct and the Baseline check does not pass, **Skip** button can be used, to omit this step, stating that the workflow is correct.

If the Checks step identifies that the Baseline run was finished successfully, all responses will appear in the Responses table. Press **Next** to continue.

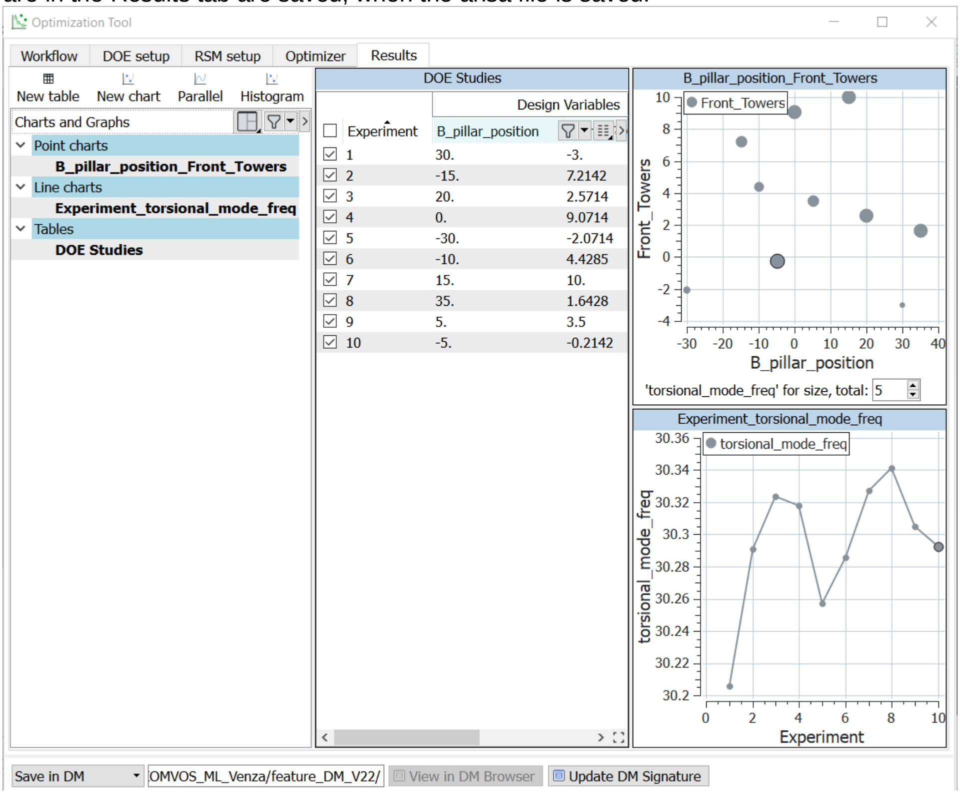


In the next, the tool checks if there are existing Simulation Models in the DM with the same name and other attributes.
If there are, the options of **Overwrite**, **New Iteration** and **Use existing** are available for selection.
Overwrite, will delete the existing Simulation model in the DM and create a new one.
New Iteration will create a new iteration of the Simulation model, keeping the old one.
Use existing, will not update the Simulation model. It will keep the existing and not save a new one.



The next page, checks if there are existing Simulation Runs in the DM with the same attributes. If there are, dependting on the selection of the Simulation model, it will ask to save the new Simulation Runs under a new DOE iteration, with the option to Delete existing Simulation runs, to avoid the conflict, or aggregate the new experiments to an existing DOE iteration.
Pressing Finish, will start running the DOE.

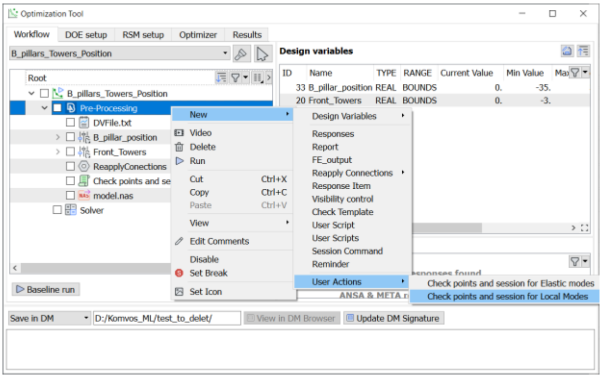
Once the DOE run is finished, the results can be viewed in the results tab. The charts and tables are in the Results tab are saved, when the ansa file is saved.



Charts can be created through the **New Chart** button and once created can be draged and droped from the **Charts and Graphs** list to the Layout on the right of the screen. The layout can also be modified by selecting one from the relative icon.
In case a saved ANSA file is opened, and a DM that contains existing Simulation run is defined, in order to view the responses and the simulation runs in the Results tab, use the **Update DM Signature** button and select the desired simulation model from the DM browser window that opens.

3.2.2. Model preparation for local modes in ANSA

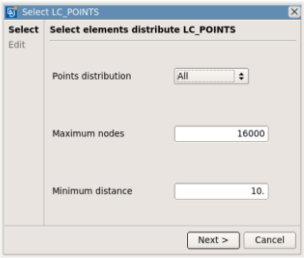
Local Modes of a Body in White can be identified through a specific process, in order to be used Feature based Machine Learning. A dedicated User Action in the Optimization task prepares the model and a specific post processing session is able to identify the local modes and create responses. These responses can be used for training machine learning models.

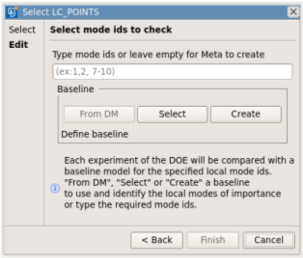


Once a DM path is defined, it is possible to define a complete Optimization.

The option <u>Save in DM</u> must be defined in order to use the DM to save a new Simulation model and as many Simulation runs as are the defined experiments.

A dedicated User action is available that will check if the appropriate interface points are available. Once the item is added in the task, move it before the FE Output item. Select it and with right click **Run** it will check if the model is prepared, or will start the wizard if it is not.



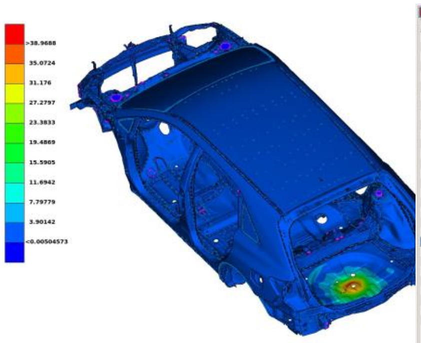


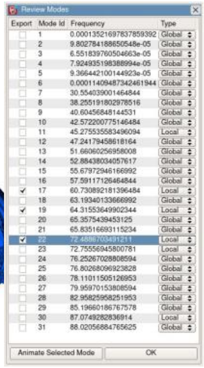
The maximum number of nodes will automatically be defined. The Minimum distance refers to the distance between interface points.

In the next step, the wizard requires a directory where baseline results of the model exist. This will be used in order to open META for the baseline local modes selection.

If baseline results exist in the current DM select **From DM**. If results exist in a different directory, use **Select**, to pick the directory where the baseline run results exist.

If results are not available, the **Create** button will start a new analysis when **Finish** is pressed. (FE output and Solver items need to be defined in the Optimization task for the Baseline run to start.)

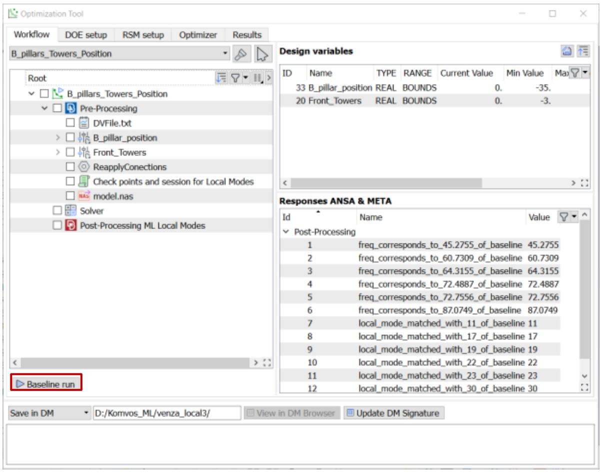




Once the baseline run is finished, META will open and the Review modes window will appear. This window shows the modes and frequencies and has automatically labeled them as Local and Global.

Once the desired local modes are selected, when META is closed, the interface points will be created, and also the dedicated post processing item will be defined, with a session specific for identifying the selected local modes. Finally the local modes task needs to be reset. Right click on the Check points and session for Local Modes task item and select **Reset.**

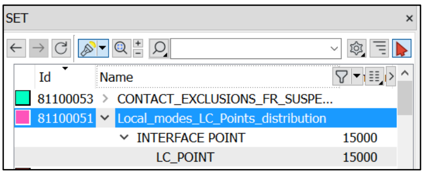
The task is ready.

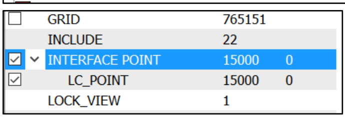


The Post Processing item is automatically defined and contains a pre-defined session and options in order to identify the selected local modes and their frequencies for each of the experiments.

The responses will appear in the Responses list when a new Baseline Run is started.

During a Baseline Run or DOE, in case the Interface points or the post processing item are not defined, the process will stop with an error.





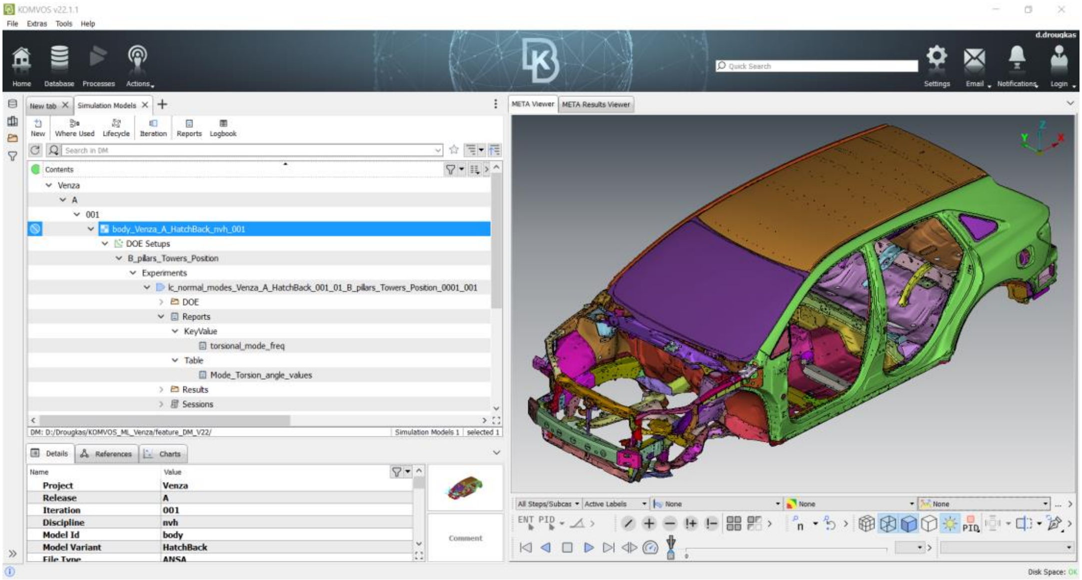
At the same time when the post processing item is created and the workflow is ready, the interface points have already been defined in the model and a specific set has been created that contains the interface points.

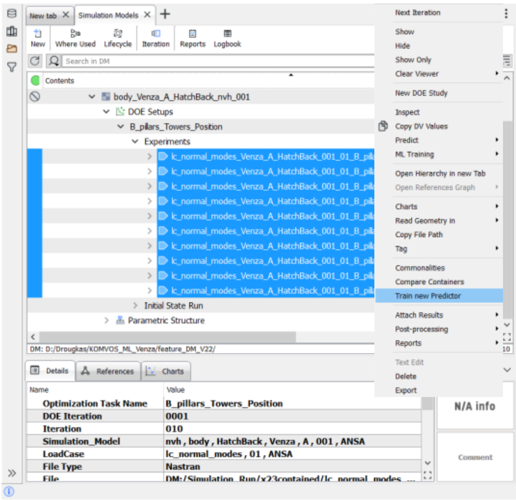
As long as this set exists in the model, the local modes wizard will not open during the run of the workflow.

Similar to the process defined in 3.2.1, switch to the DOE Setup tab and generate experiments using a DOE Algorithm. When **Start** is pressed, define the Simulation Model’s attributes and the load-case and continue to the creation of the dataset.

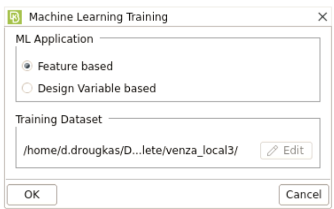
3.3. Machine Learning Training

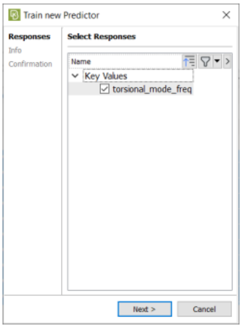
After connecting KOMVOS to the previously created DM, open the Simulation models list in the Database. Each experiment under the DOE Studies contains the mode file, DOE information, Reports and Results. The KeyValues that appear under the Report items can be used for ML Training. Each of these KeyValues can be used, in order to create a machine learning model that will predict this response Finite Element Body in White model.
Once a Machine Learning model is trained, a new entity called Predictor will be created and saved in the DM.

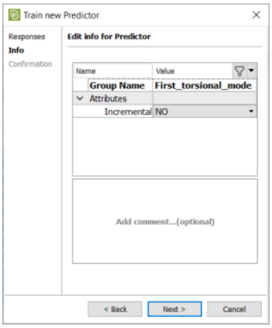




If there are no Predictors available in the DM, in order to start the ML training, select the Experiments that will be used for the training and with right click select **Train new Predictor**. At least 10 experiments need to be selected in order to start the ML training.
In the Machine Learning Algorithm window that appears, select **Feature based**.

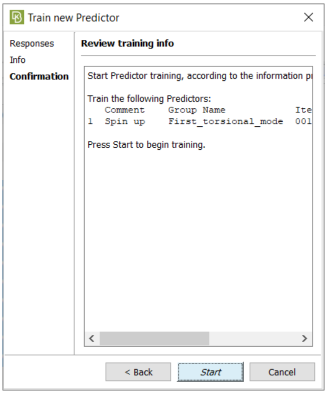




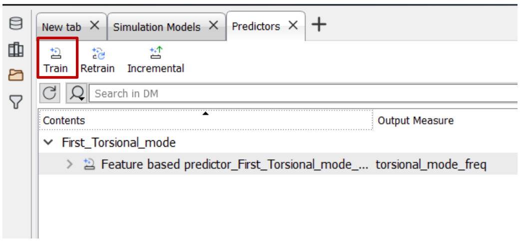


In the wizard that appears, in the Responses page, it is possible to select Key Values results to train for. Select the *torsional_mode_freq* response. The predictor task type will be regression.
In case the values are non-numerical, the predictor type will be classification.

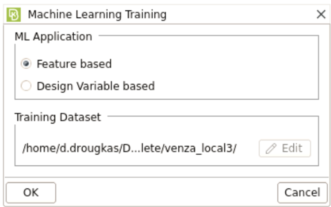
In the Info page, provide a name and optionally some comments for the Predictor that is about to be created.

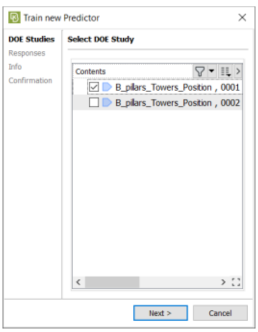


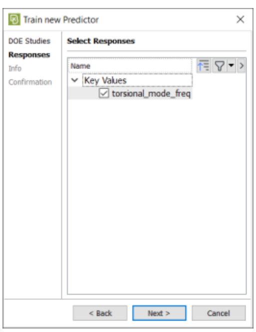
Finally Confirm and begin the training by pressing the **Start** button.



If predictors already exist in the DM, a new training can start from the Predictors Tab. Once the Train button is pressed, the Machine learning Algorithm window appears.
Select **Feature based**.

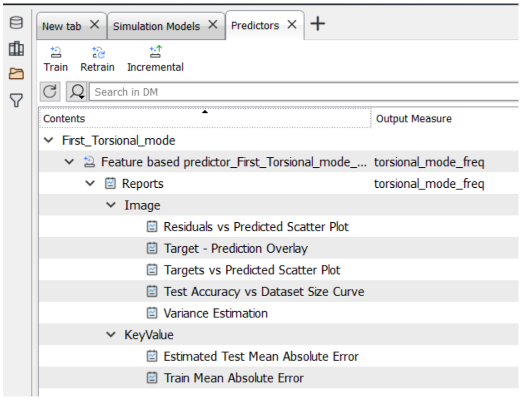






The new Predictor wizard starts, this time requiring a DOE Study in the first page. Select a DOE Study that contains more than 10 experiments and continue with the next pages as before, selecting the *torsional_mode_freq* response and Start the training.

During the training, the training status can be seen in the Home tab.

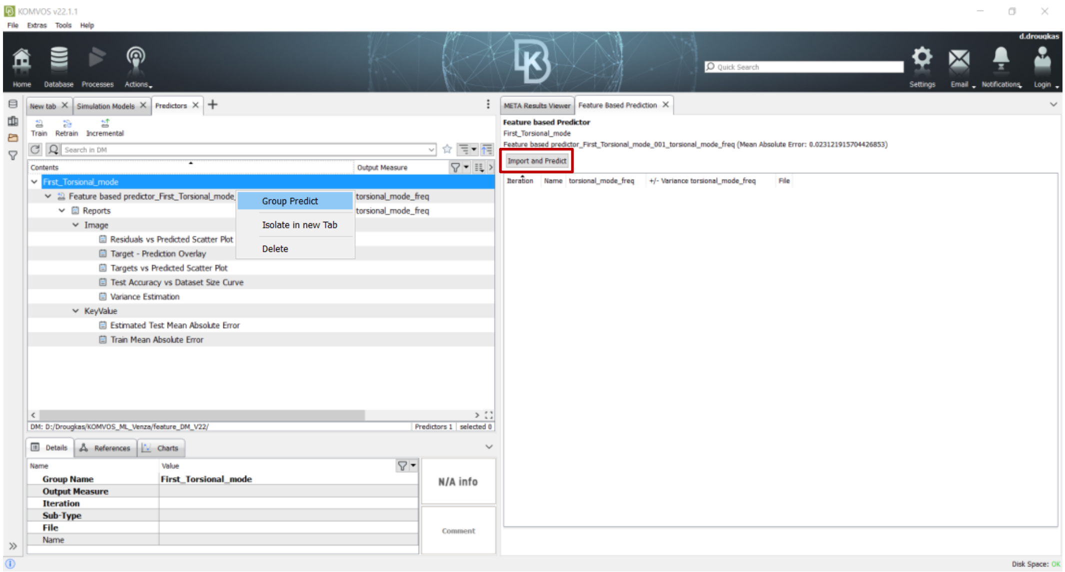


Once the training is completed, the Predictor entities will appear in the Predictors list.
Predictors appear in Groups (Groups view). Each group contains as many predictors, as the number of responses that were selected during training definition.

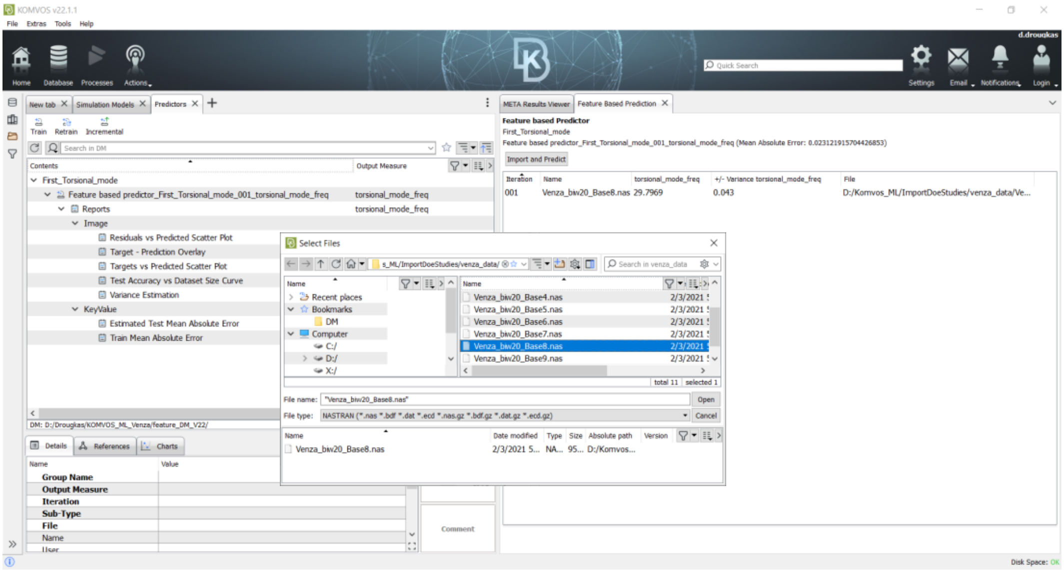
Each predictor contains Report items that provide indicators about the predictor's performance, and the Estimated Mean Absolute Error of the Prediction.
More information about these reports can be found in KOMVOS User's Guide in the Machine Learning chapter, **Predictor Reports**.

3.4. Machine Learning Prediction

While viewing the Predictors tab in Groups view, right click on the Group name and select **Group Predict**. This will open the Feature Based Prediction window on the right side of the window. The Prediction window contains a list where the name and predicted value will appear.

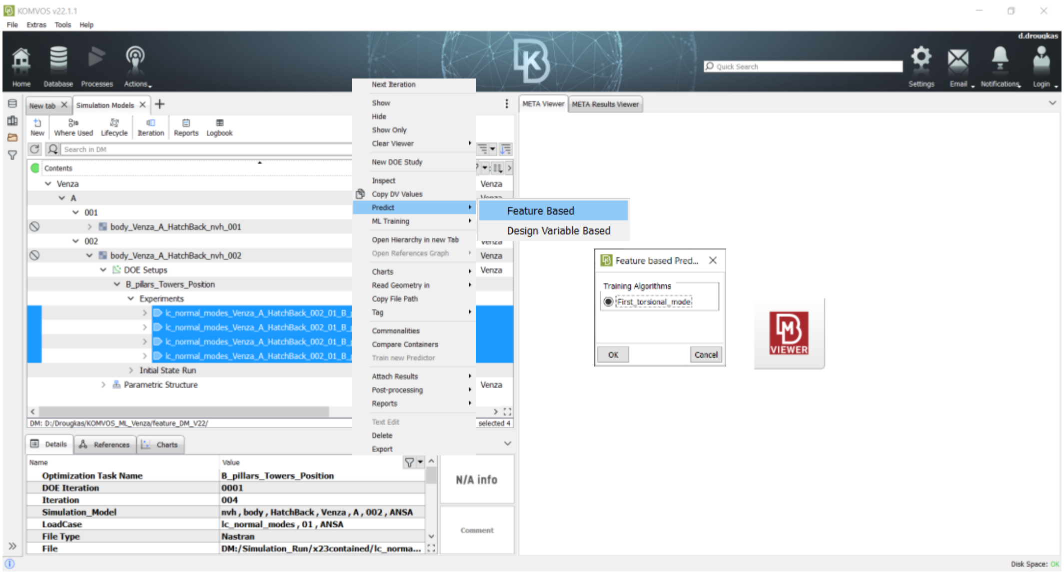


To make a prediction for one or more Body in White models that are not added in the DM, press the **Import and Predict** button and select a model from a directory.



A few seconds later, the predicted value will appear in the prediction window, along with the variance predicted value.

To make a prediction for models that are already in the DM, switch to the Simulation models tab and select the desired experiments. With right click, select Predict and select the existing Feature based predictor, to start the prediction.



4. DOE Studies

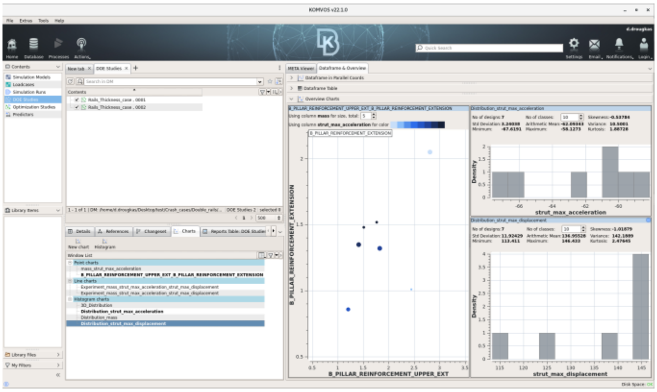
4.1. Introduction

The DOE Study entity contains DOE data existing in the DM. Such entities are automatically created when a DOE Study is created from ANSA and saved in the DM. Also DOE Studies entities are automatically created when a DOE Study is imported through the Import DOE action.
The DOE Studies list hosts the DOE Study entities existing in the DM.
DOE Study entities can be created for previous versions through the Simulation runs list (DOE View), using the *Create DOE_Study* action on a DOE Iteration.

4.2. Handling DOE Studies

From the Contents list double click on the DOE Studies to open the DOE Studies tab. The list in the left of the screen shows the DOE Study items per DOE iteration. On the right side of the screen the Dataframe & Overview tab holds a parallel coordinates plot and the Dataframe table. The charts and table are linked, so that filtering can be applied from the chart to the table.

Through the Dataframe & Overview tab it is possible to explore the experiments' Design Variable values and Responses. Filtering can take place in the Parallel Coordinates plot in order to distinguish desired experiments. Click on the axis and drag, to mark the filter bounds. To remove the filter, left click on the side of the axis where the cross appears.
In the Charts tab, press the **New Chart**, and select **Point 2D** to create a 2D Points chart. In the window that appears select what DV or response will be reoriented from each axis. If **Charts** tab is missing, open it from the *Show/Hide tabs* drop menu.



Similarly **Line 2D** chart creates a line chart. X axis on the line chart is always the experiment id.

Filtered and checked experiments can be added in a new Chart (line, point, histogram, etc.) from the respective button in the Charts tab for post processing.

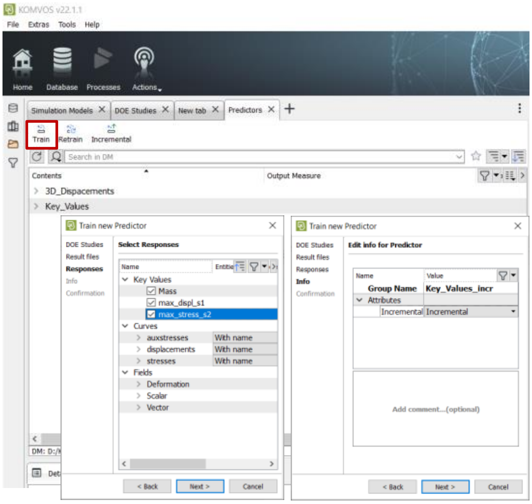
A reports table can be generated for checked experiments, to show the reports for each one in the respective Reports Table tab that is created.
Right click on the Dataframe table and select **Reports Table**, to create the Reports table for the checked experiments.

5. Incremental Training

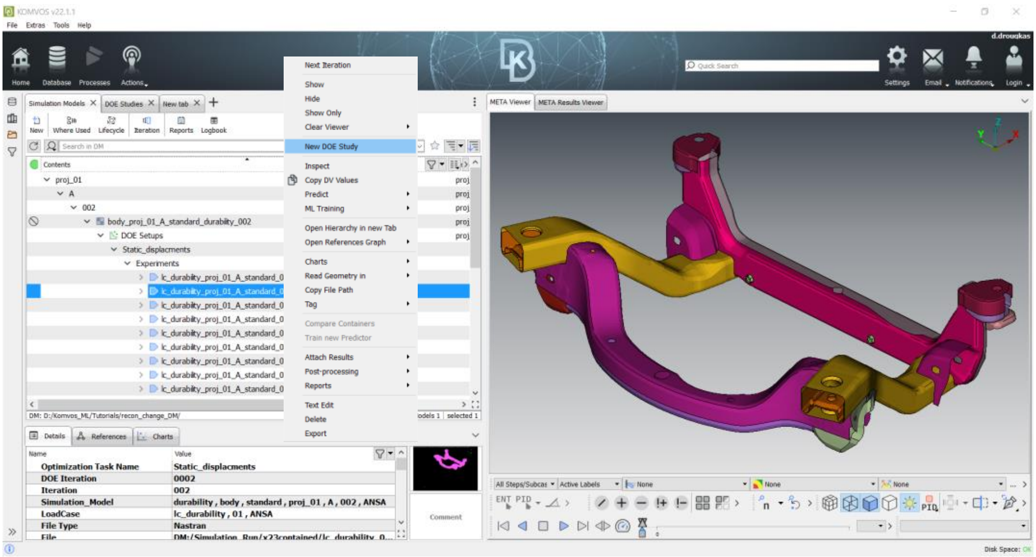
5.1. Introduction

Incremental Learning allows for addition of new data in existing predictors, without the time expense of training again for all previous data.
In order to be able to “update” an existing predictor with new data, the predictor needs to support Incremental learning. This is an option that has to be activated when the predictor is originally created.
Incremental learning can be used for both DV and Feature based predictors. On DV based Machine Learning, the additional data need to have the same Design Variables and responses as the original training data. On Feature base Machine Learning, the additional data need to have the same responses as the original training data.

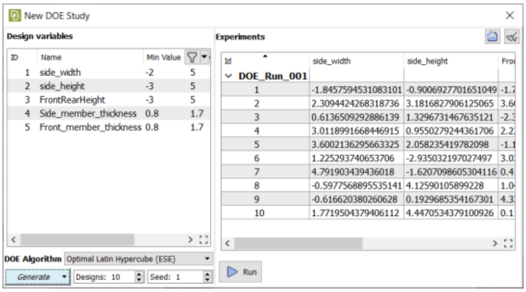
5.2. Incremental Training



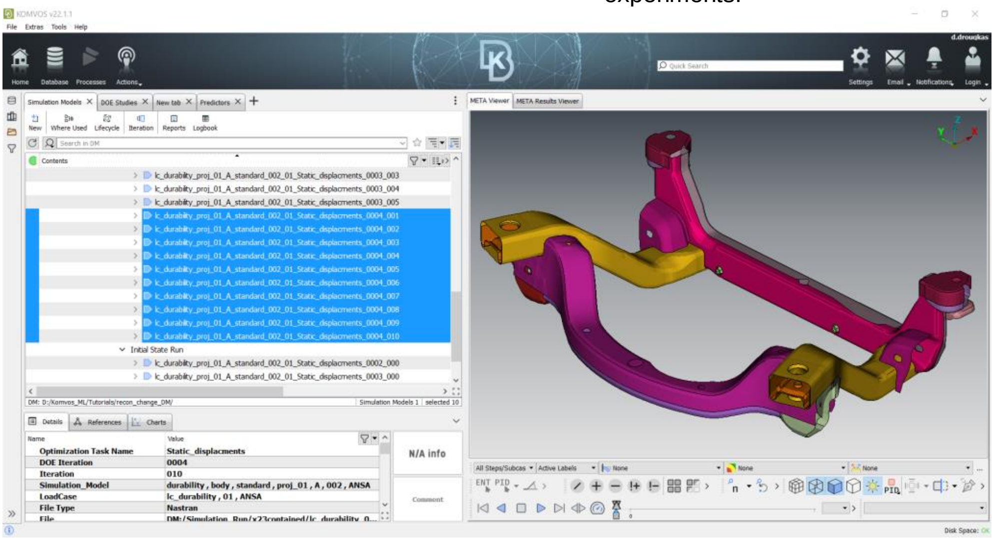
Connect to the DM used in the DV based Machine Learning (chapter 2).
Create a new DV based group of predictors using the initial DOE Study, select the three KeyValues as Responses and in the Info page, select **Incremental**.
This option will allow the predictors that will be created, to be updated at a later stage with additional experiments.



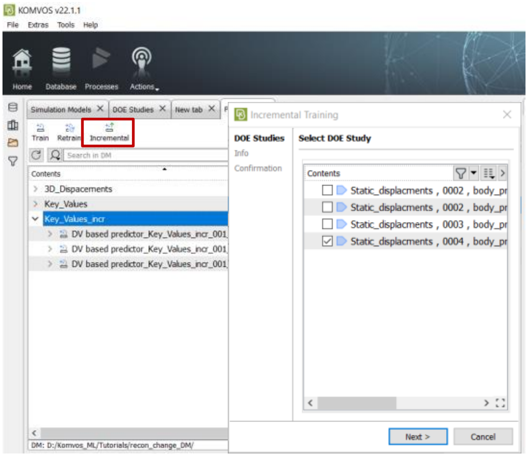
Additional experiments can be added in the DM through ANSA, or using the **New DOE Study Option** from the right click menu of the Simulation models (NLE View).



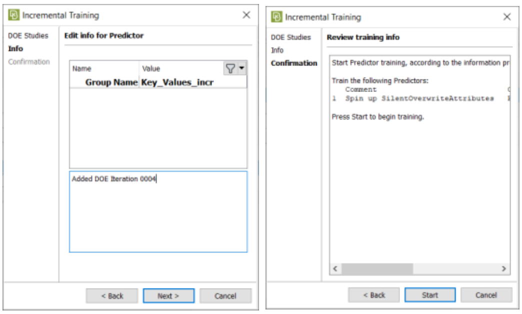
The New DOE Study window opens where the existing design variables and their boundary values appear. Select a DOE Algorithm, a number of Designs and a Seed number.
Press **Generate** to see the experiments in the respective table.
Press Run to start the creation of the experiments.
This process will open ANSA in the background and the original parametric model will be used in order to create the new DOE. Any items that were included in the initial workflow (Design Variables, scripts, Solver, post processing, etc.) will be used in batch, to create the experiments.



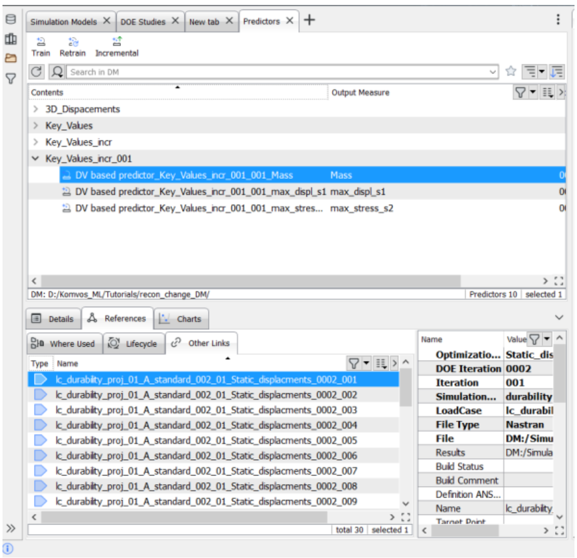
The new experiments are added under the same Simulation model, with a new DOE Iteration id.



Select the recently created predictor group and press the **Incremental** button and in the first page of the wizard, select the new DOE Studies iteration to be used for the training.



In the next page optionally type in some comments, and press **Next**.
Press **Start** to begin the training.



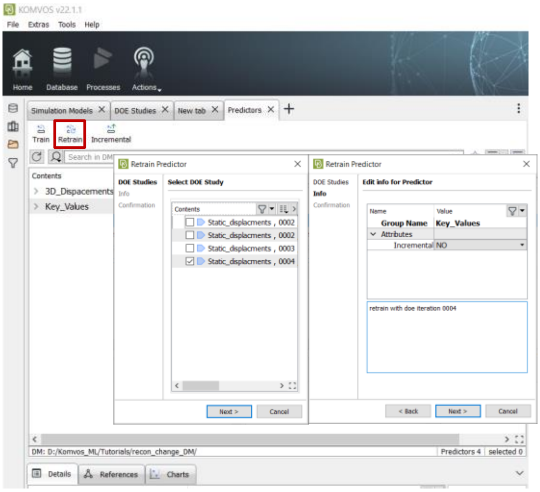
The new Group of predictors is created and it was trained with both 0002 and 0004 DOE iteration experiments.

6. Retrain

6.1. Introduction

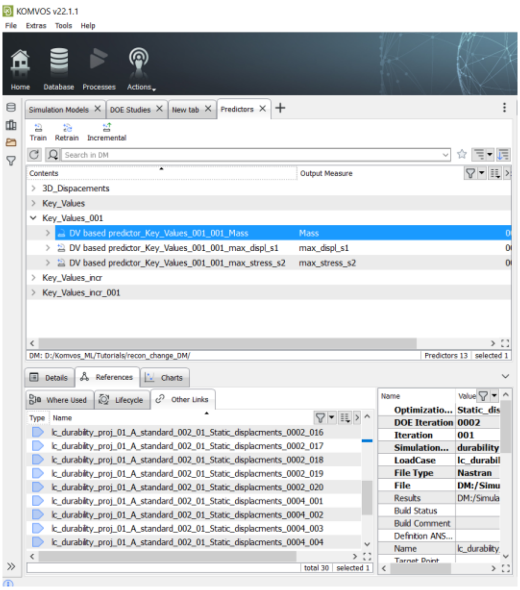
The Retrain option for a selected predictor will create a new predictor that will be trained with both the original data of the selected predictor and also new selected data.
The new predictor that will be created will be able to predict for the responses of the original predictor.
Retrain option can be used for both DV and Feature based predictors, basic or incremental. On DV based Machine Learning, the additional data need to have the same Design Variables and responses as the original training data. On Feature base Machine Learning, the additional data need to have the same responses as the original training data.

6.2. Retrain



Connect to the DM used in the DV based Machine Learning (chapter 2).
Select the Key_Values group and press the **Retrain** button.

In the first page of the wizard, select the new DOE Iteration. This way the new Predictor group that will be create that will be trained for both the DOE Iterations, old and new (0002 and 0004).
Press **Next** and in the next page optionally add comments and Continue to **Start** the training.



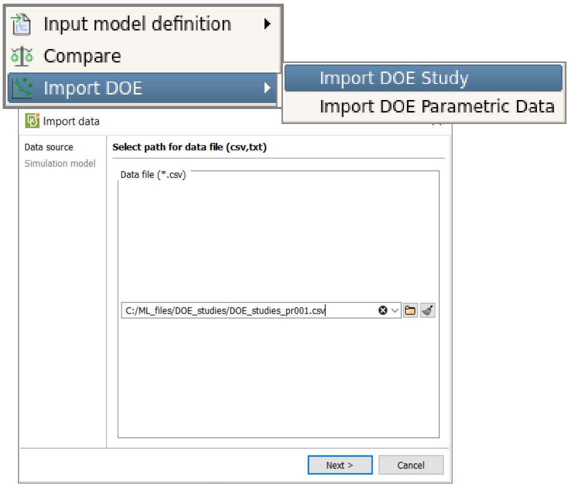
The new Group of predictors is created and it was trained with both 0002 and 0004 DOE iteration experiments.
If we compare the time needed for Incremental And Retrain options to create a group trained with both DOE iterations, the Incremental option is much faster.
The accuracy between the two may have differences and Incremental and Retrain options should be used according to the nature of the problem.

7. Import DOE Studies

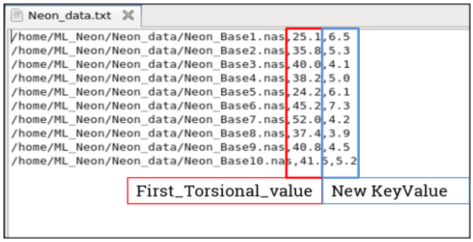
Machine Learning in KOMVOS can be performed even in cases where the “data set” was not created using ANSA’s Optimization tool. Using Import DOE action, it is possible to import experiments as Simulation Runs in a new DM and perform Machine Learning actions.

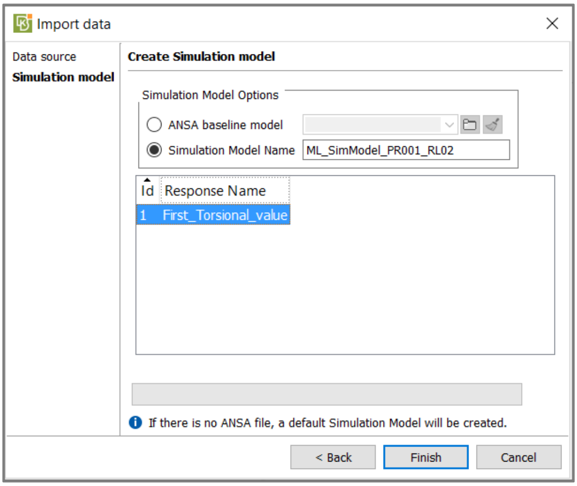
7.1. Import DOE Study

In case past DOE are available and the Feature Based Machine learning will be used, it is possible to use the **Import DOE Study**, to store the data in the DM. This way Feature based Machine learning will be possible.

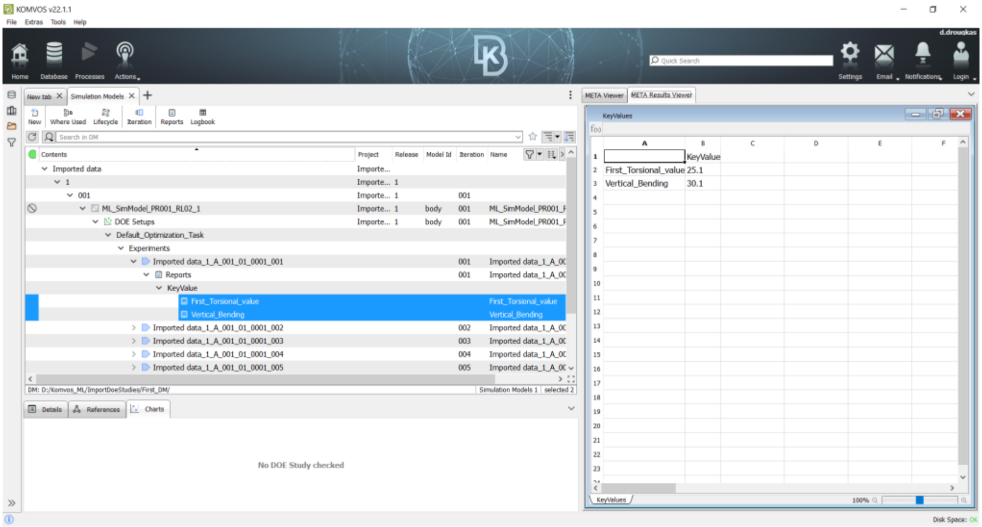


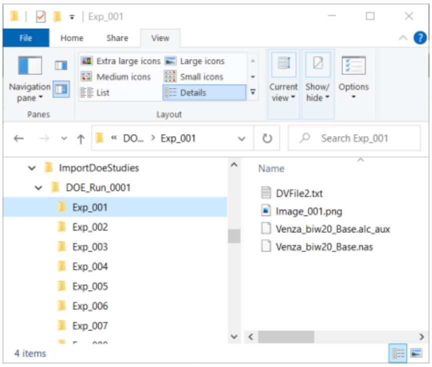
Connect to a new empty DM directory and using the **Import DOE >Study** option from the Actions Menu, select a CSV file that contains information about the DOE Study. The CSV file should contain the paths of each model and comma separated the response values.





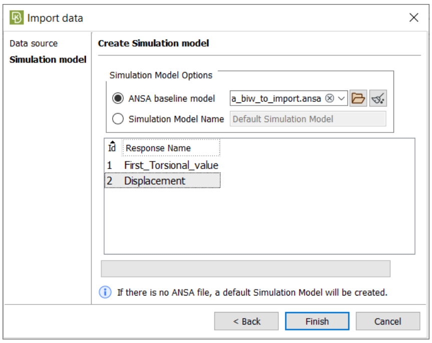
In the Responses window, a response name needs to be added, that corresponds to the value in the respective column after the file path.
If an ANSA file is available and the DOE was performed in ANSA, the design variable information can also be acquired. The information will be stored automatically in the appropriated format under a Simulation Model, so that it can be used from the Machine Learning functionality



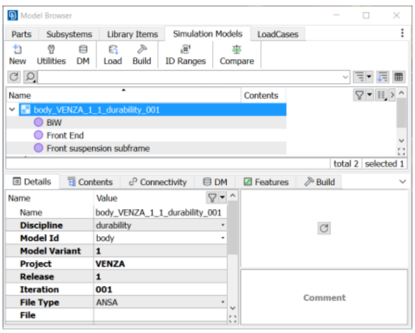


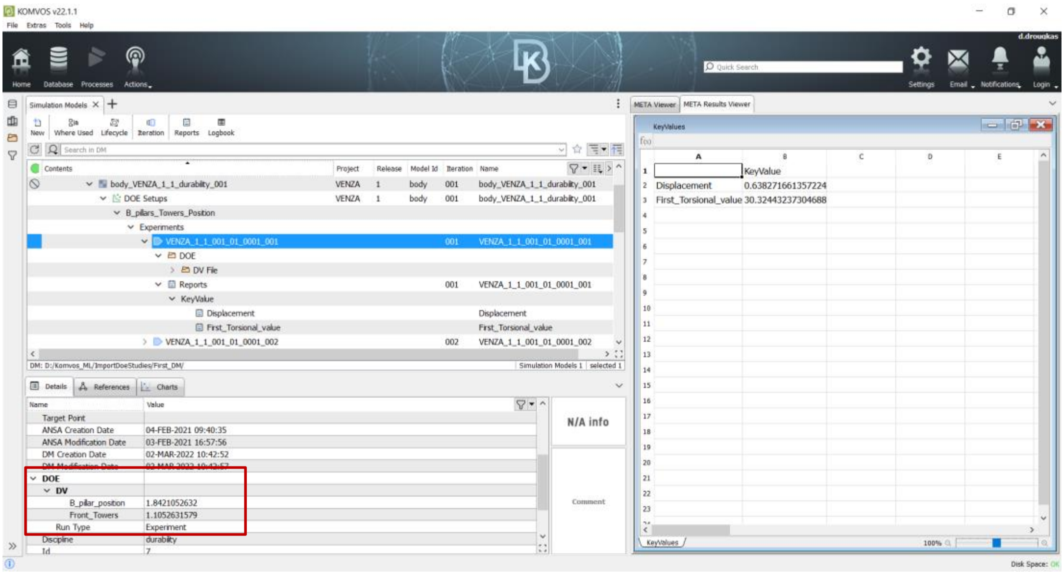
In case the DOE Study was created in ANSA through the optimization tool (ANSA Tutorials,"Optimization with ANSA and META"), the CSV file should contain the relative paths to each experiment's model. In this case, the DV file in each experiment will be used in order to read the parameters and provide them in the DOE Study in DM.

	A	B	C
1	../DOE_Run_0001/Exp_001/Venza_biw20_Base.nas	30.3244	0.6383
2	../DOE_Run_0001/Exp_002/Venza_biw20_Base.nas	30.2601	0.6363
3	../DOE_Run_0001/Exp_003/Venza_biw20_Base.nas	30.2913	0.6371
4	../DOE_Run_0001/Exp_004/Venza_biw20_Base.nas	30.3328	0.6381
5	../DOE_Run_0001/Exp_005/Venza_biw20_Base.nas	30.2764	0.6368
6	../DOE_Run_0001/Exp_006/Venza_biw20_Base.nas	30.2564	0.6361
7	../DOE_Run_0001/Exp_007/Venza_biw20_Base.nas	30.3062	0.6387
8	../DOE_Run_0001/Exp_008/Venza_biw20_Base.nas	30.2458	0.6352
9	../DOE_Run_0001/Exp_009/Venza_biw20_Base.nas	30.2886	0.6363
10	../DOE_Run_0001/Exp_010/Venza_biw20_Base.nas	30.2567	0.6372
11	../DOE_Run_0001/Exp_011/Venza_biw20_Base.nas	30.3140	0.6365
12	../DOE_Run_0001/Exp_012/Venza_biw20_Base.nas	30.3310	0.6364
13	../DOE_Run_0001/Exp_013/Venza_biw20_Base.nas	30.3177	0.6356
14	../DOE_Run_0001/Exp_014/Venza_biw20_Base.nas	30.2913	0.6353
15	../DOE_Run_0001/Exp_015/Venza_biw20_Base.nas	30.2854	0.6367
16	../DOE_Run_0001/Exp_016/Venza_biw20_Base.nas	30.2928	0.6349
17	../DOE_Run_0001/Exp_017/Venza_biw20_Base.nas	30.3105	0.6384
18	../DOE_Run_0001/Exp_018/Venza_biw20_Base.nas	30.2928	0.6369
19	../DOE_Run_0001/Exp_019/Venza_biw20_Base.nas	30.3088	0.6376
20	../DOE_Run_0001/Exp_020/Venza_biw20_Base.nas	30.3313	0.6372



In case the ANSA file is available and contains a Simulation model, it can be selected as baseline model. This way the information of the existing simulation model will be used in the creation of the simulation model in the DM.

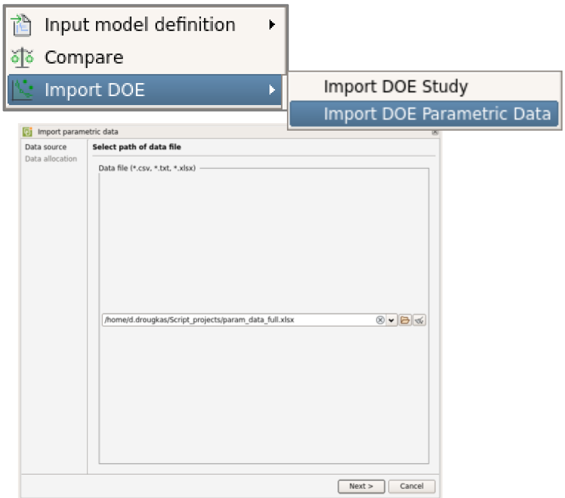




The result of this Import action is a DM with a Simulation model that has attributes collected from the given ANSA file. Key Values have been created for each of the responses given in the CSV file and also Design Variable information are available for each experiment. This will allow the use of these experiments fot training DV based or Feature based Predictors.

7.2. Import DOE Parametric DATA

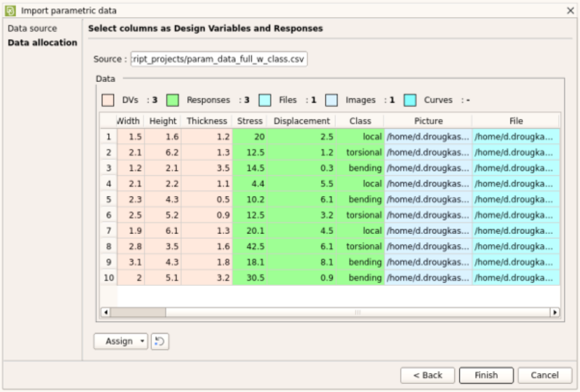
In case past DOE are available and the parameterization was not done in ANSA, it is possible to use the **Import DOE Parametric Data** option, to store the data in the DM. This way Design Variable based Machine learning will be possible.



Using the **Import DOE > Parametric Data** option from the Actions Menu, it is possible to import already existing parametric data from an Excel spreadsheet (*.xls) or CSV file (*.csv).

	A	B	C	D	E	F	G	H	I
1	Exp id	Width	Height	Thickness	Stress	Displacement	Class	Picture	File
2	1	1.5	1.6	1.2	20	2.5	local	.../Image_001.png	.../biw1.nas
3	2	2.1	6.2	1.3	12.5	1.2	torsional	.../Image_002.png	.../biw2.nas
4	3	1.2	2.1	3.5	14.5	0.3	bending	.../Image_003.png	.../biw3.nas
5	4	2.1	2.2	1.1	4.4	5.5	local	.../Image_004.png	.../biw4.nas
6	5	2.3	4.3	0.5	10.2	6.1	bending	.../Image_005.png	.../biw5.nas
7	6	2.5	5.2	0.9	12.5	3.2	torsional	.../Image_006.png	../biw6.nas
8	7	1.9	6.1	1.3	20.1	4.5	local	.../Image_007.png	.../biw7.nas
9	8	2.8	3.5	1.6	42.5	6.1	torsional	.../Image_008.png	.../biw8.nas
10	9	3.1	4.3	1.8	18.1	8.1	bending	.../Image_009.png	.../biw9.nas
11	10	2	5.1	3.2	30.5	0.9	bending	.../Image_010.png	.../biw10.nas

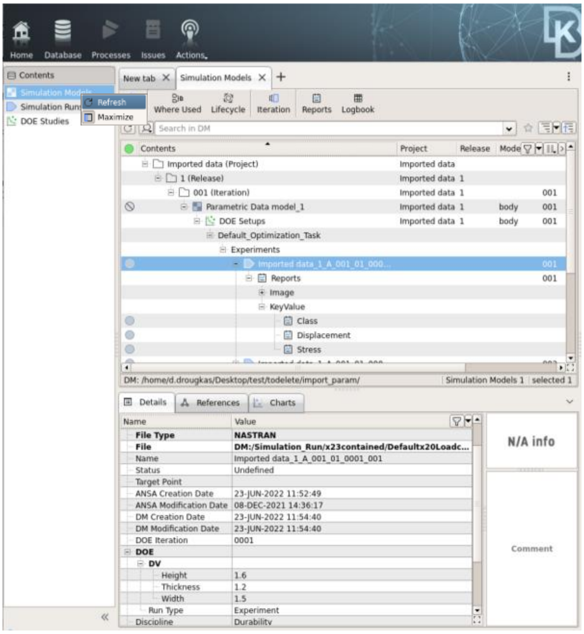
The spreadsheet should have the experiment id in the first column. The other columns should contain Design Variables, Responses, and optionally pictures and the files of the models.



The rows and columns of the column formatted file will be read and displayed in the data window. Each column can be assigned as Design variable, Response, Image, File or Response curve.
The image and file paths should point to existing images and files.

The data will be added in the DM structure and will be available to be used for Design Variable based Machine Learning.

NOTE! The first column of the file needs to contain the ids of the experiments. The first row of the file is the header and needs to contain the type of value it contains(names of DVs or responses, pictures, files)



The result of this Import action is a DM with a default empty Simulation model that has default attributes. Key Values have been created for each of the responses given in the spreadsheet file and also Design Variable information are available for each experiment. This will allow the use of these experiments fot training DV based Predictors.