

Quick guide to using Python preprocessing in SIMCA®-online

This document describes how to practically add a script to SIMCA® and use it in SIMCA®-online.

General information

Since version 16 it is possible for SIMCA® to use and embed Python scripts for preprocessing of data. This lets you morph the dataset for both modelling and predictions (including SIMCA®-Q and SIMCA®-online) row by row.

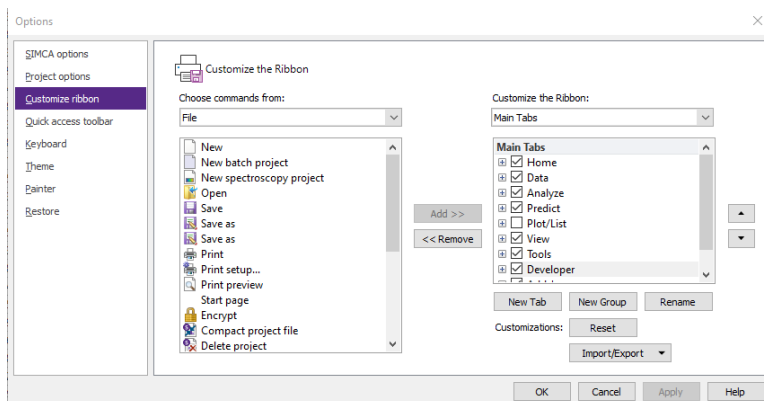
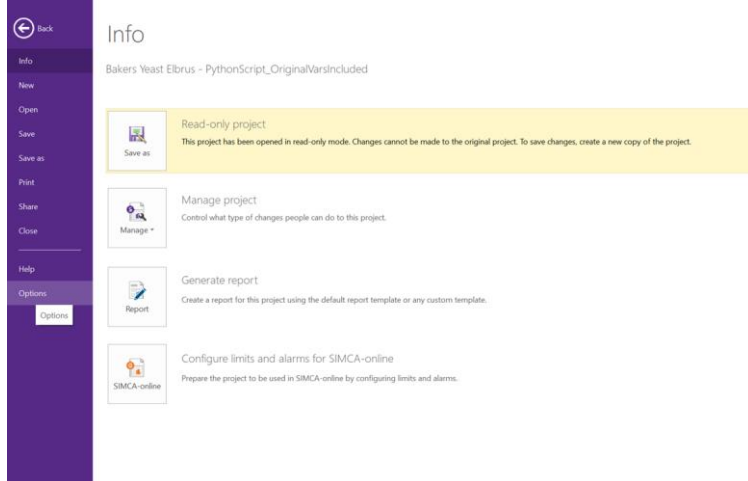
Note: Applying a Python script might slow down performance of the SIMCA®-Q and SIMCA®-online systems.

Creating a Python preprocessing script

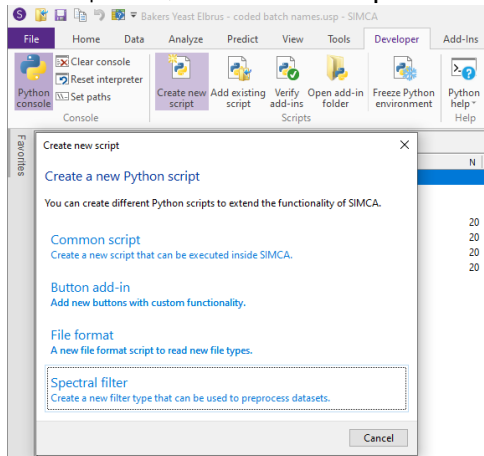
This section describes how to create a Python preprocessing script. The actual Python code is left for you to implement.

1. Add the Development tab to SIMCA® (if it is not there already):

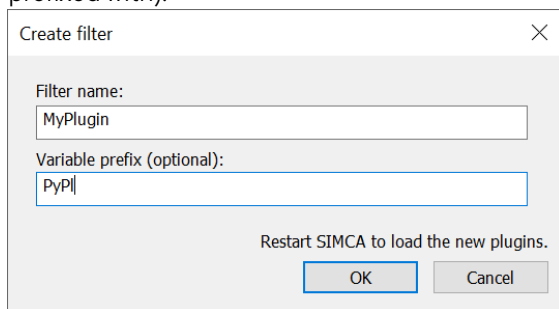
In **File | Options**, in the Customize ribbon section, select to add the Development tab and click **OK**



2. Create a new Spectral filter with Python by clicking **Create new script** in the Scripts group on the Developer tab, and then select **Spectral filter**



3. Give the script a name, and optionally also a prefix (that all new variables in the new dataset will be prefixed with).



A template is created and automatically opened. This is where you add your Python code



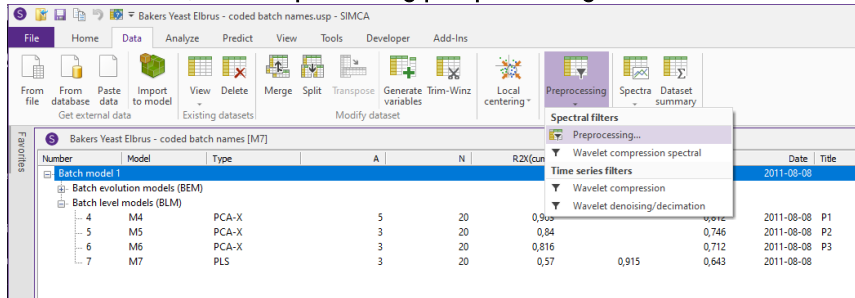
MyPlugin.py

The example function (transform) just takes the square of all values, but this is where you need to enter your own transformation/preprocessing function. The matrix returned must always have the same number of rows as the input data collection but other than that anything is possible.

Applying a preprocessing script

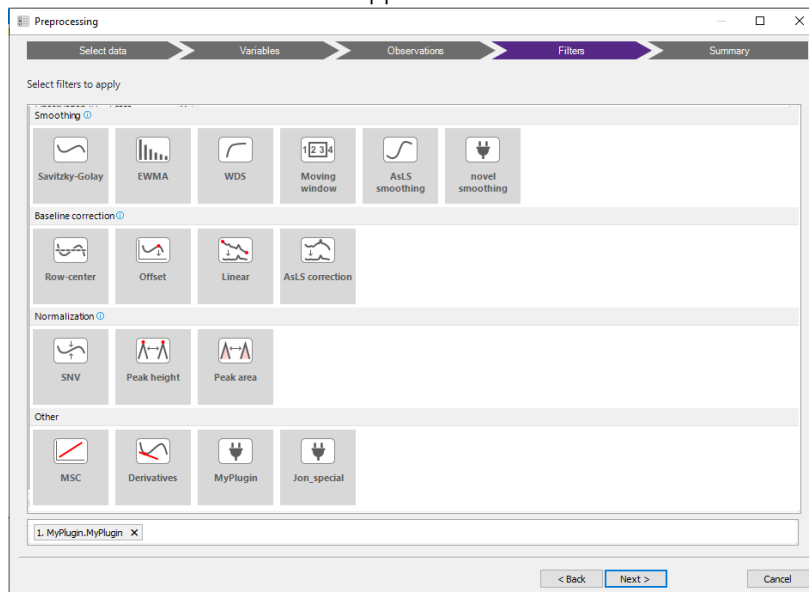
This section describes how a spectral filter and Python preprocessing script can be used in SIMCA®. Note that SIMCA® needs to be restarted to load any new Python add-ins from the add-in folder, and that all preprocessing scripts that you want to use need to be in the add-in folder of SIMCA®.

1. On the Data tab, click **Preprocessing | Preprocessing...**

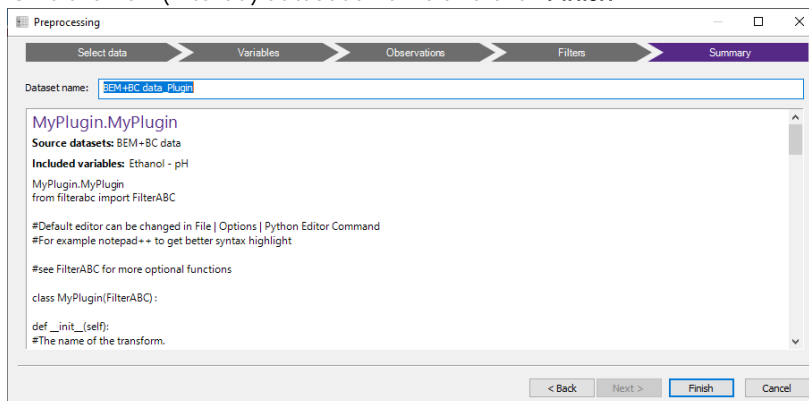


2. Select the Source dataset to apply the filtering to and then go through the wizard and select the Python script (or any of the built-in filters) that you want to use and click **Next** to continue to the Summary page.

Note that several filters can be applied in a chain



3. Give the new (filtered) dataset a name and click **Finish**

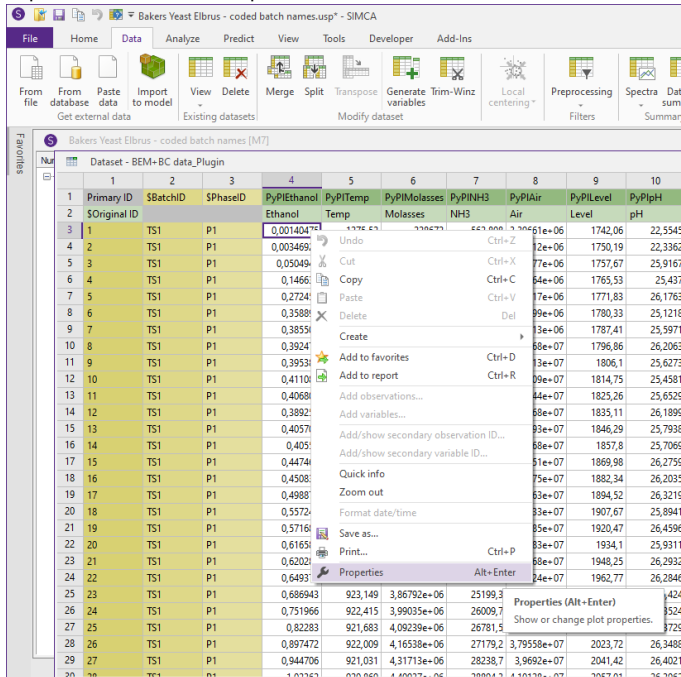


Viewing an existing embedded Python preprocessing script

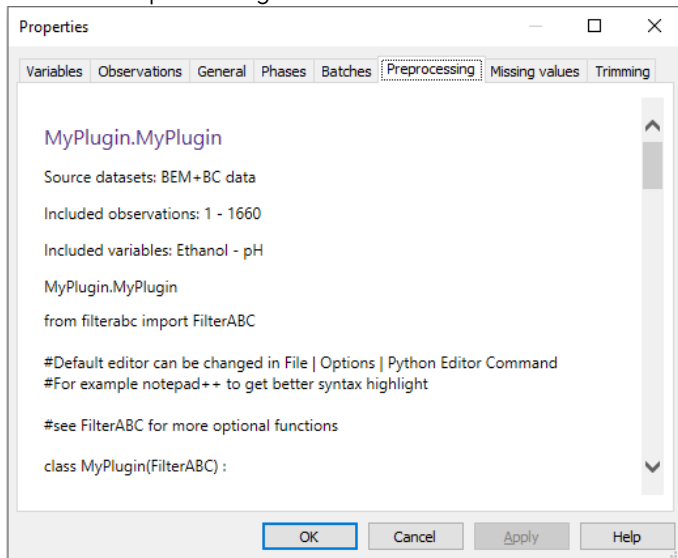
This section describes how to view the embedded Python script for a specific dataset in SIMCA®.

To view the existing Python preprocessing script that is already embedded in a dataset:

1. Open the dataset Properties:



2. View the Preprocessing tab:

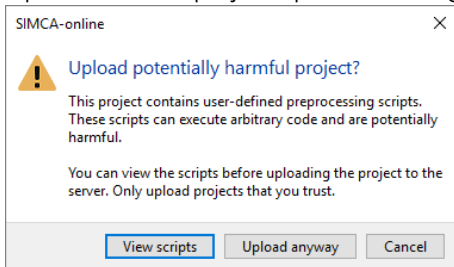


Using a SIMCA project with an embedded Python preprocessing script in SIMCA®-online

This section describes using a SIMCA® project with an embedded Python script in SIMCA®-online.

Note that since Python scripts can potentially contain dangerous code, any project with existing Python scripts needs to be approved before it can be used in the products.

1. Upload SIMCA® project opens a warning-message:



To see the Python script content, click **View scripts**. After reviewing the script, provided you are content, click **Upload anyway** to move forward

2. Step through the project configuration as with any other project.
3. Display your new/filtered variables in the charts