Monitoring Toolbox V1.2
Quick Start Guide

Note: The current version will expire on December 31, 2021. New versions will be released.

Predictive Monitoring Toolbox

Objective of this toolbox

- Monitor the operation of industrial processes, and detect the occurrence of events (e.g., faults, near faults, shutdown) based on process data

Key features of this toolbox

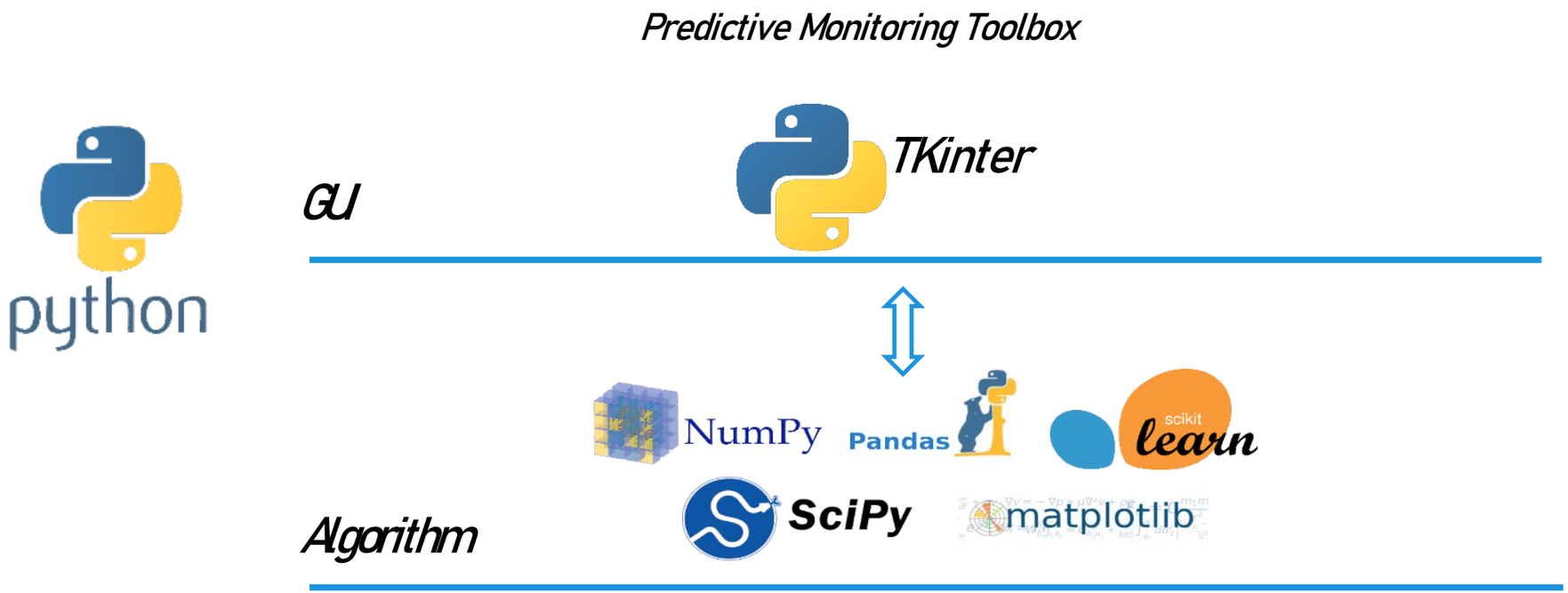
- Five causal analysis algorithms are provided to investigate the cause-effect relationships among process variables
- Five unsupervised learning algorithms are provided to conduct predictive monitoring of the process and to detect events taking place in processes
- Online monitoring feature is provided to simulate the online implementation of monitoring schemes for industrial processes
- Hierarchical distributed monitoring is provided to account for larger processes of more process variables

Overview of Predictive Monitoring Toolbox

Predictive Monitoring Toolbox

An overview of the toolbox

The toolbox was developed using Python



- *Python has open source license (free to use)*
- *Python possesses stable libraries*
- *Python is developer friendly software*

Predictive Monitoring Toolbox

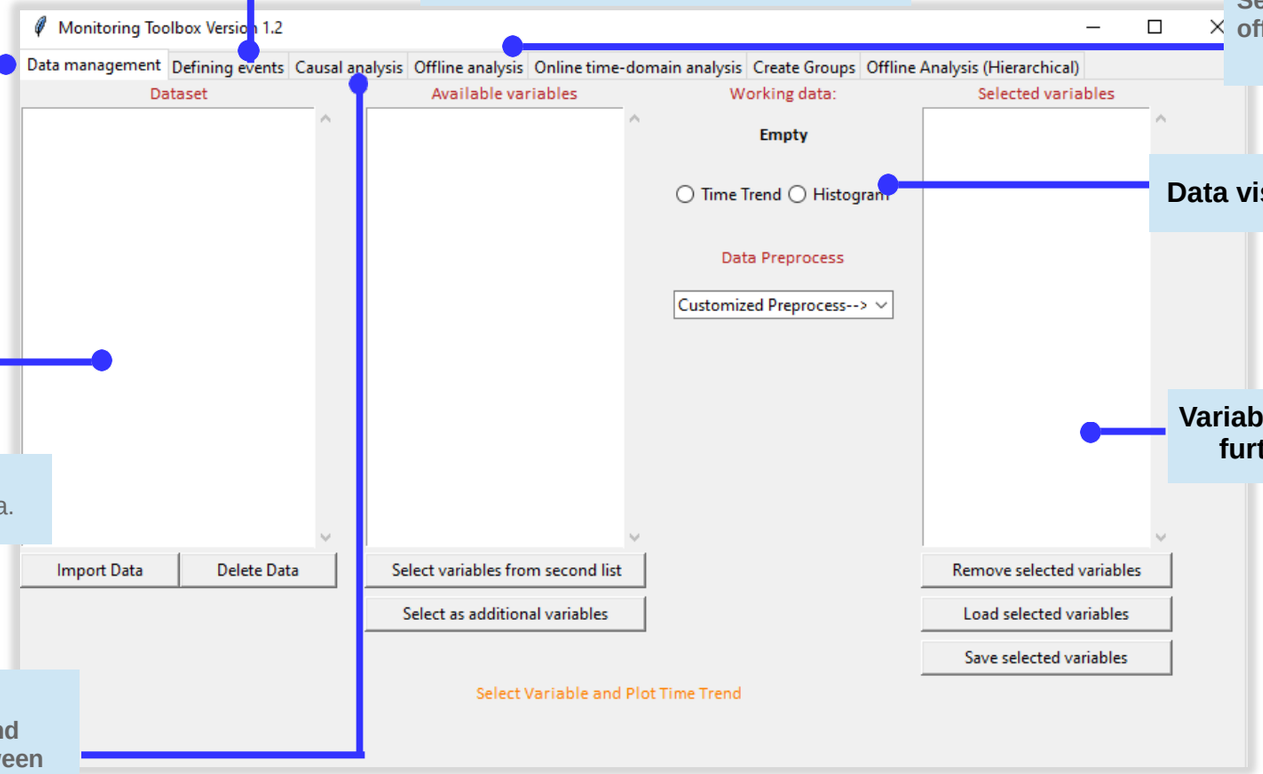
An overview of the toolbox

We create this guide to help you minimize the learning curve.

Defining events
Define events manually based on the visualization of trajectories

Offline and Online monitoring
Several algorithms are available for offline/online monitoring of a process

Data management
Import datasets and process data necessary for analysis



Data visualization

Data management
Import and **delete** data.

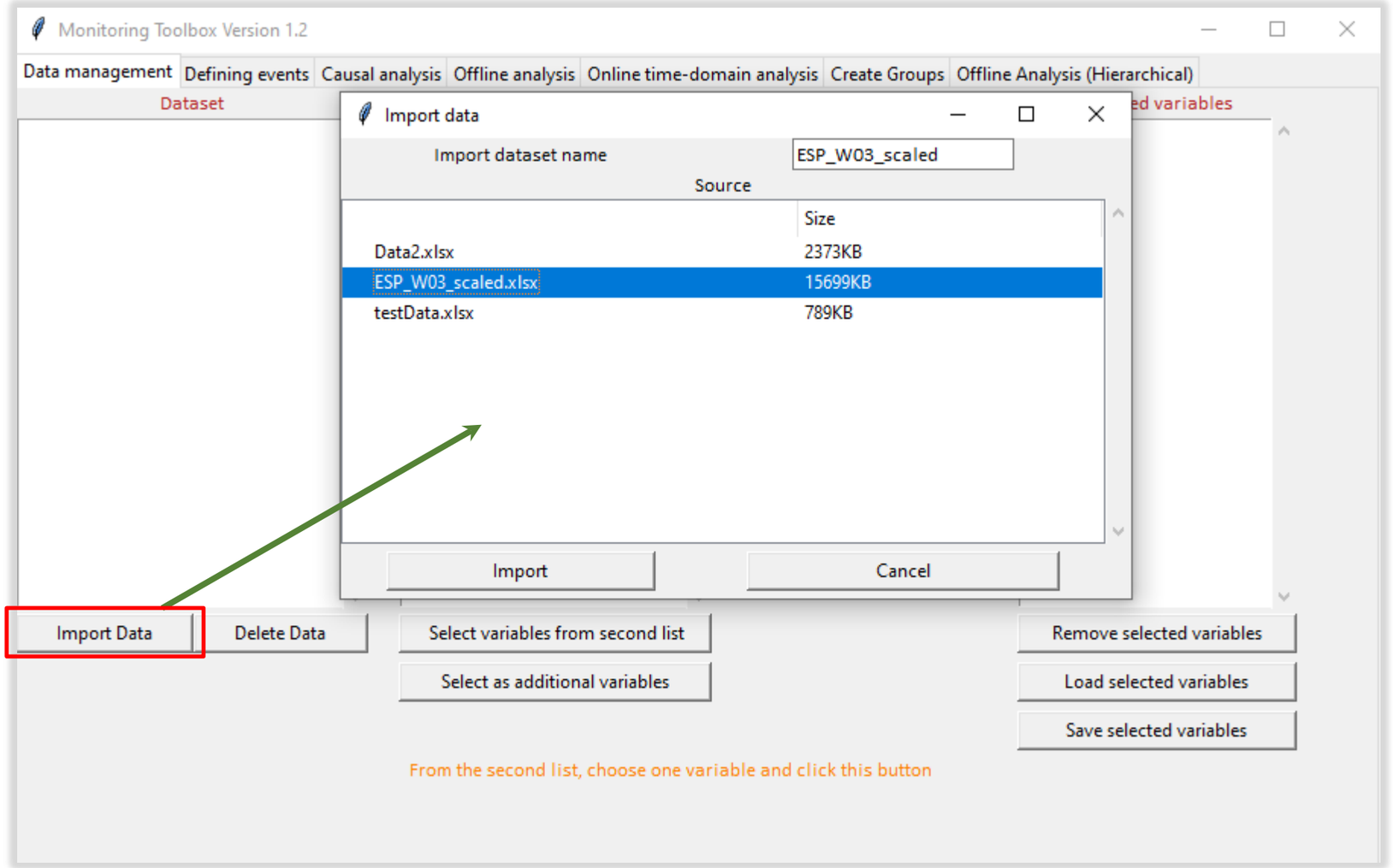
Variables selected for further analysis

Causal analysis
To investigate cause and effect relationship between pair of variables

Predictive Monitoring Toolbox

An overview of the toolbox – Data management

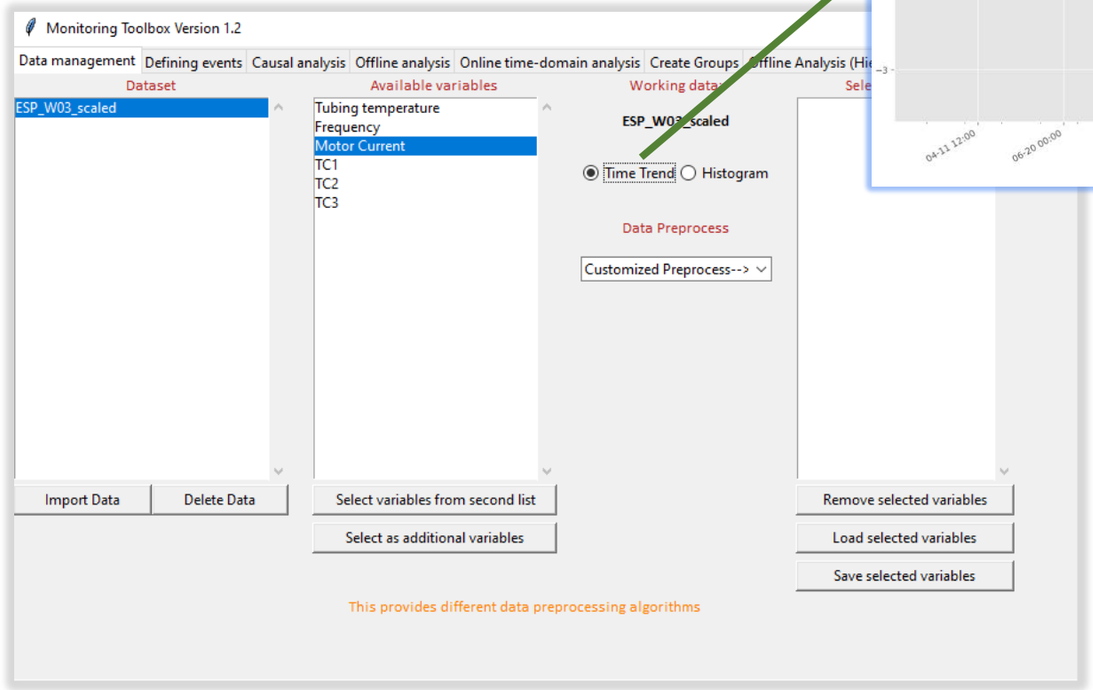
- Data Management
- Data Visualization
- Data Preprocessing
- Defining Events
- Causal Analysis
- Offline/Online Analysis

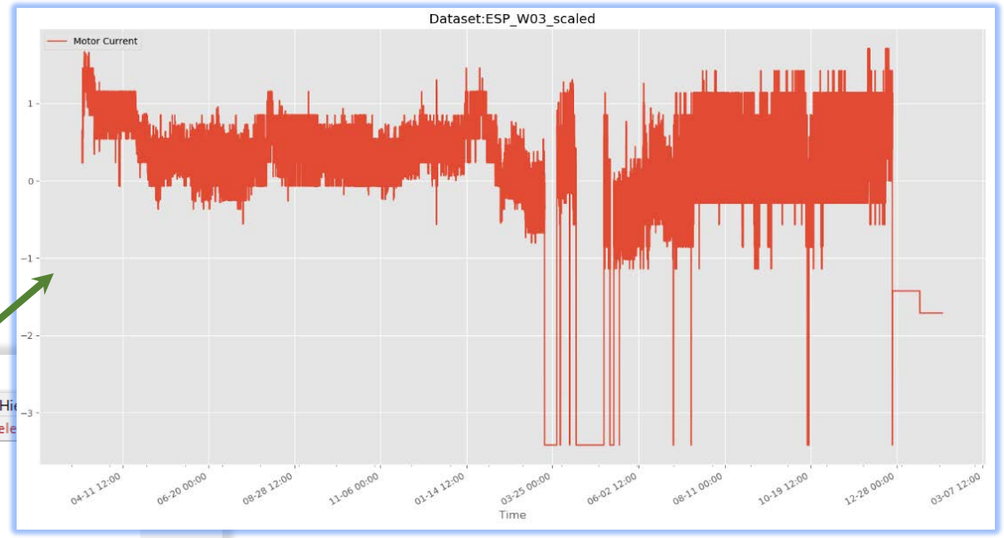


Predictive Monitoring Toolbox

An overview of the toolbox – Data visualization

- Data Management
- Data Visualization
- Data Preprocessing
- Defining Events
- Causal Analysis
- Offline/Online Analysis

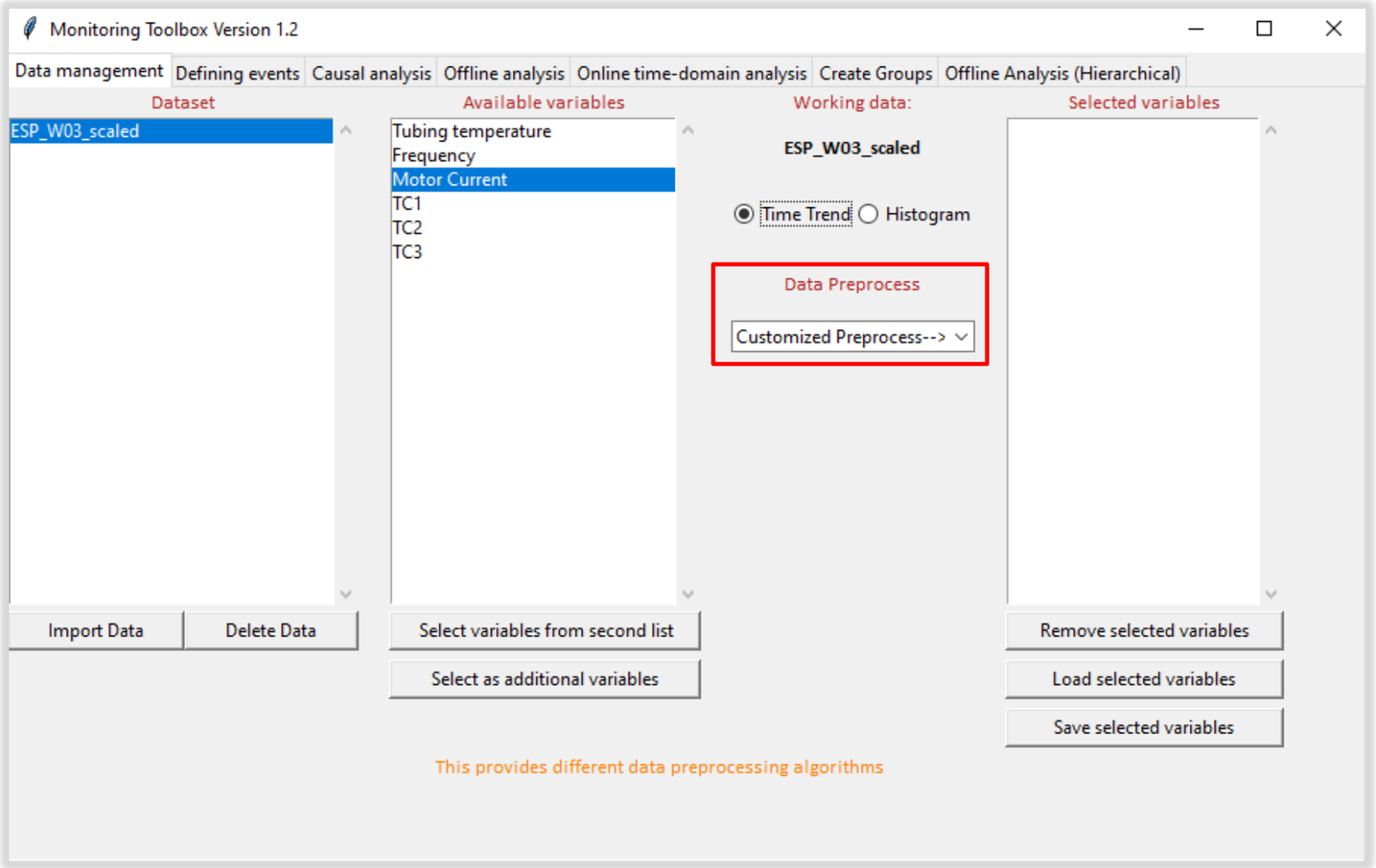




Predictive Monitoring Toolbox

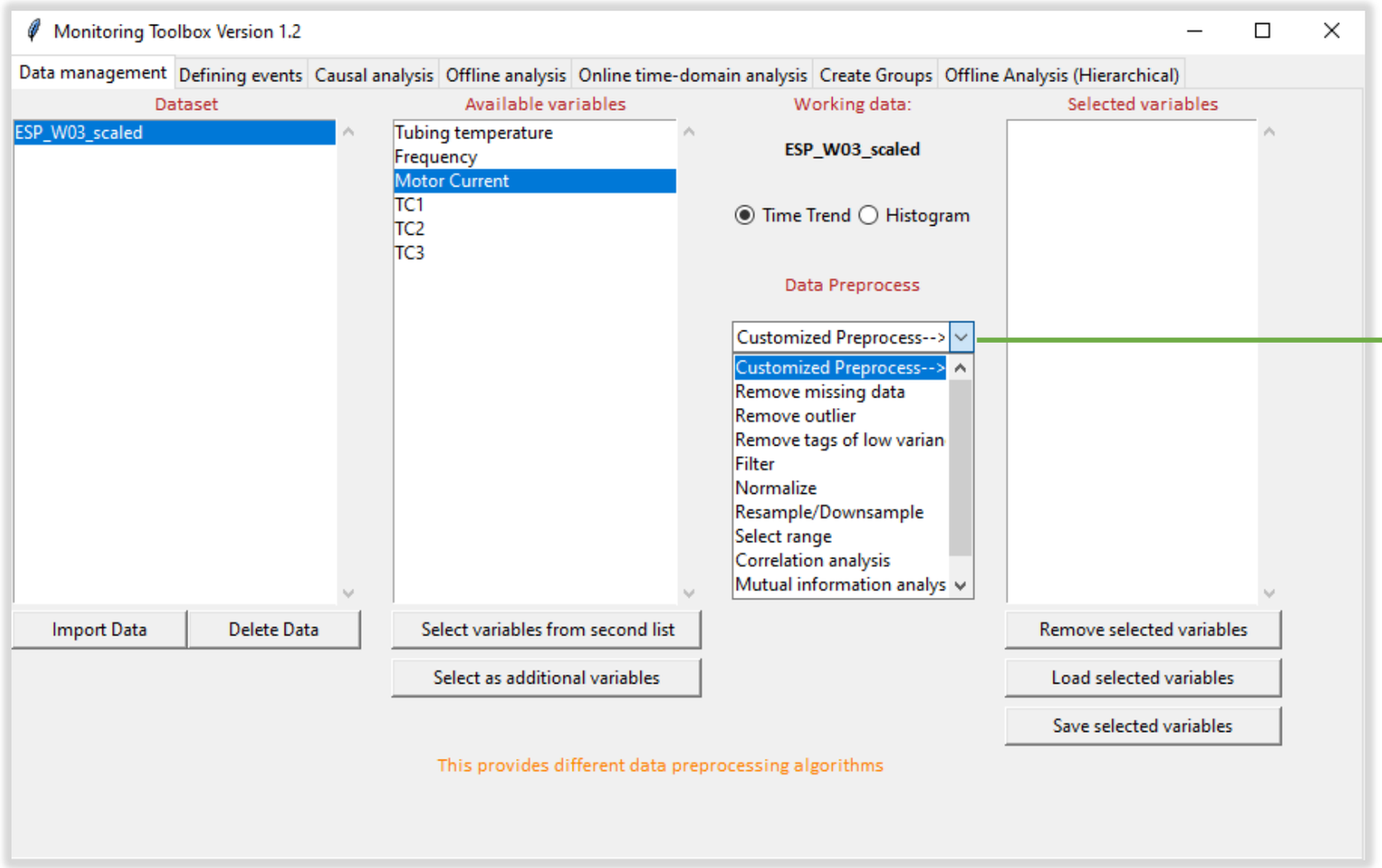
An overview of the toolbox – Data preprocessing

- Data Management
- Data Visualization
- **Data Preprocessing**
- Defining Events
- Causal Analysis
- Offline/Online Analysis



Predictive Monitoring Toolbox

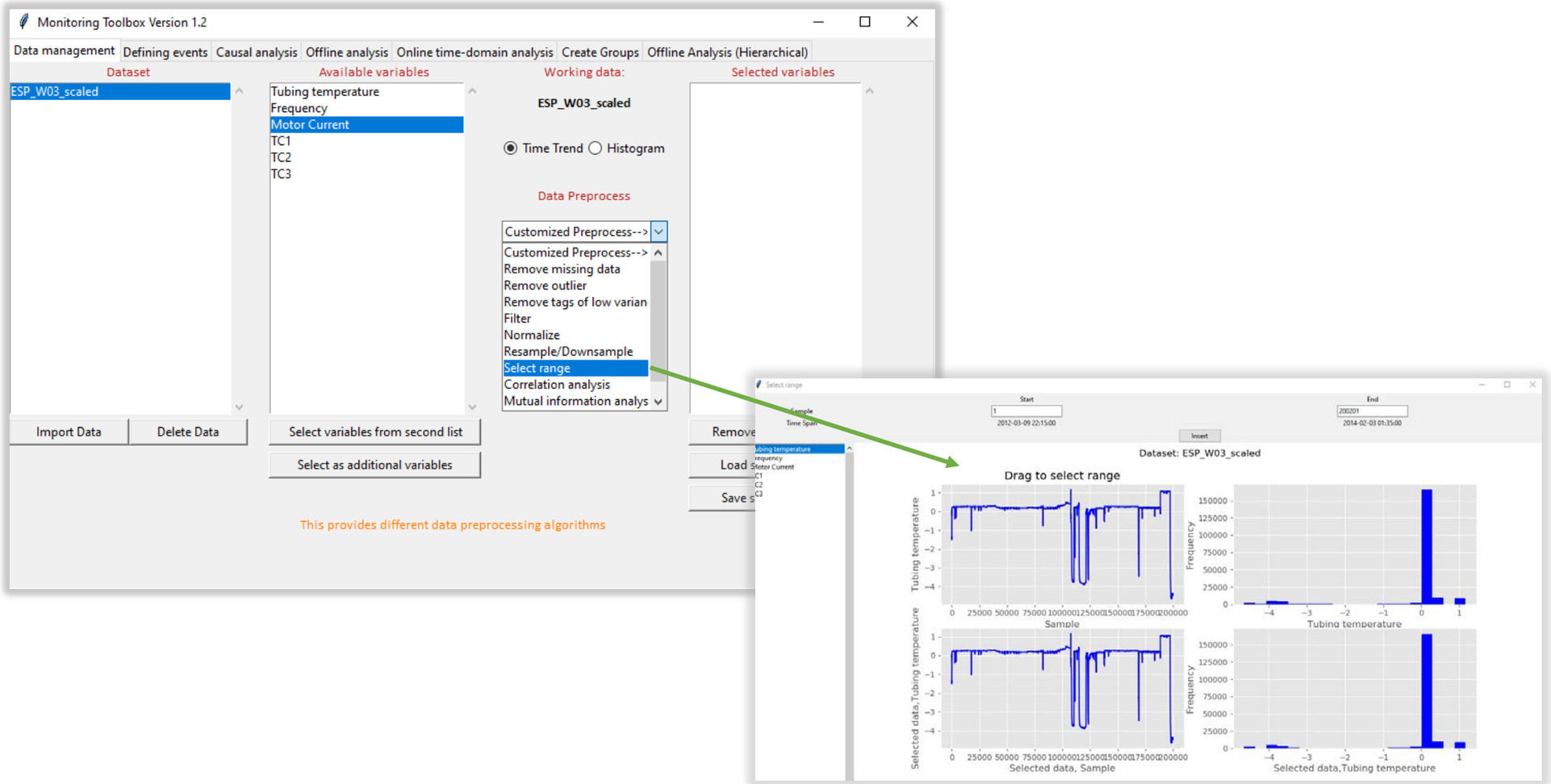
An overview of the toolbox – Data preprocessing



- Missing data
- Remove outliers
- Low variance
- Filter
- Normalize
- Resample
- Range selection
- Correlation
- Mutual information
- Auxiliary tags

Predictive Monitoring Toolbox

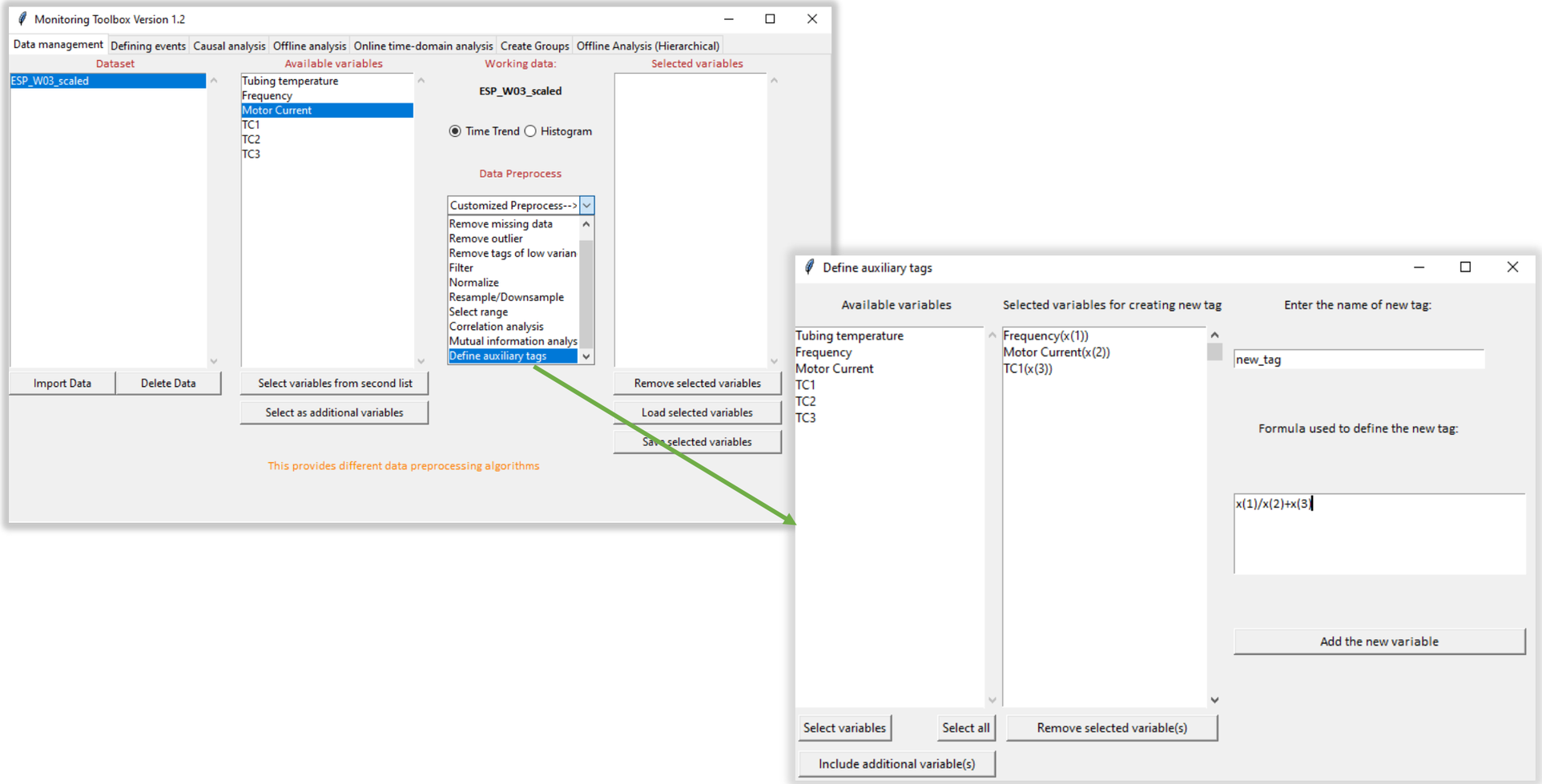
An overview of the toolbox – Data preprocessing



- Missing data
- Remove outliers
- Low variance
- Filter
- Normalize
- Resample
- **Range selection**
- Correlation
- Mutual information
- Auxiliary tags

Predictive Monitoring Toolbox

An overview of the toolbox – Data preprocessing

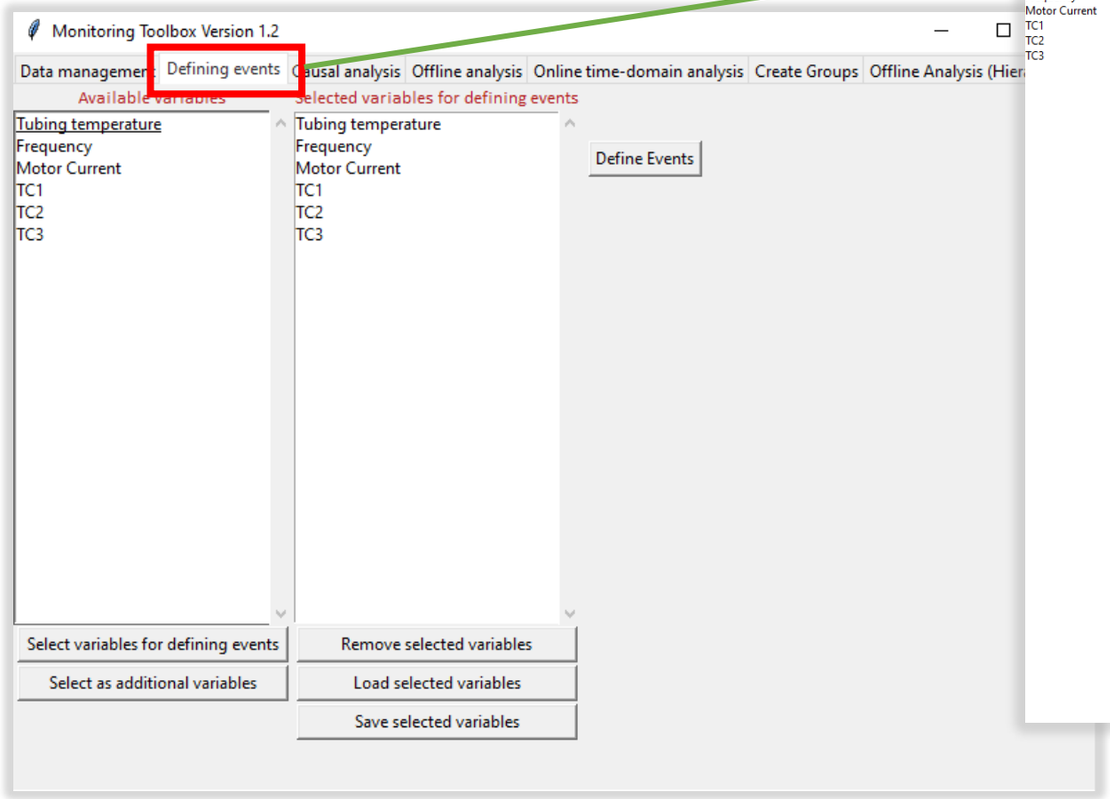


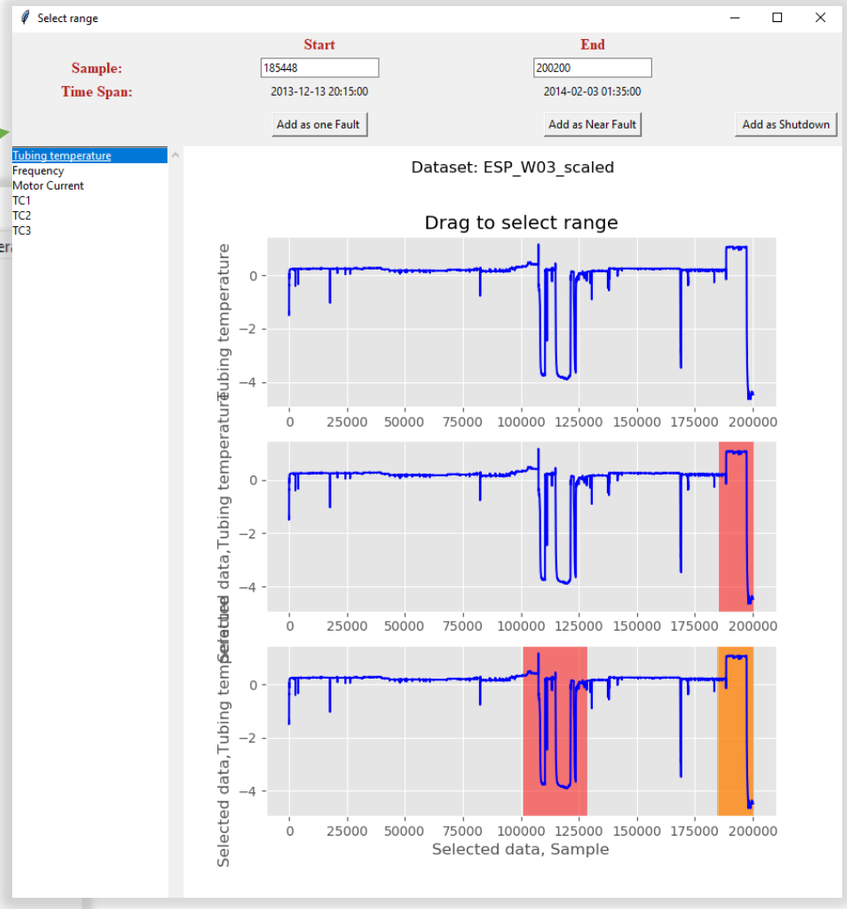
- Missing data
- Remove outliers
- Low variance
- Filter
- Normalize
- Resample
- Range selection
- Correlation
- Mutual information
- Auxiliary tags

Predictive Monitoring Toolbox

An overview of the toolbox – Defining events

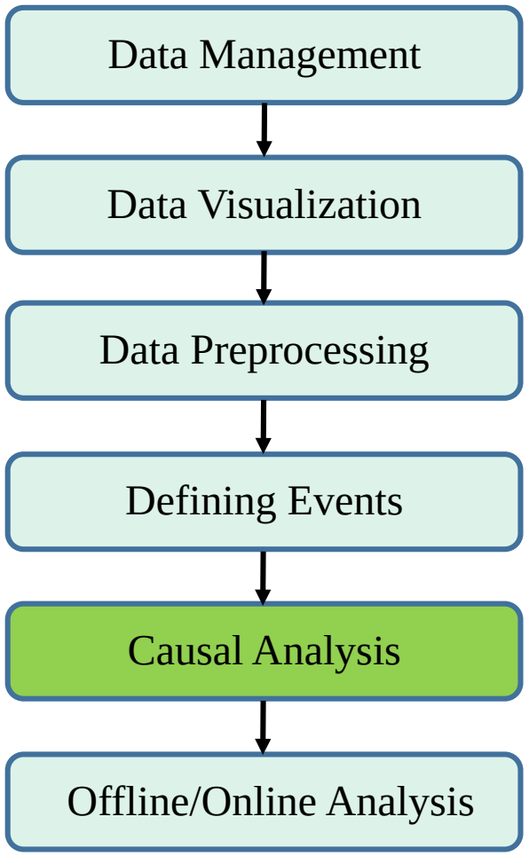
- Data Management
- Data Visualization
- Data Preprocessing
- Defining Events
- Causal Analysis
- Offline/Online Analysis

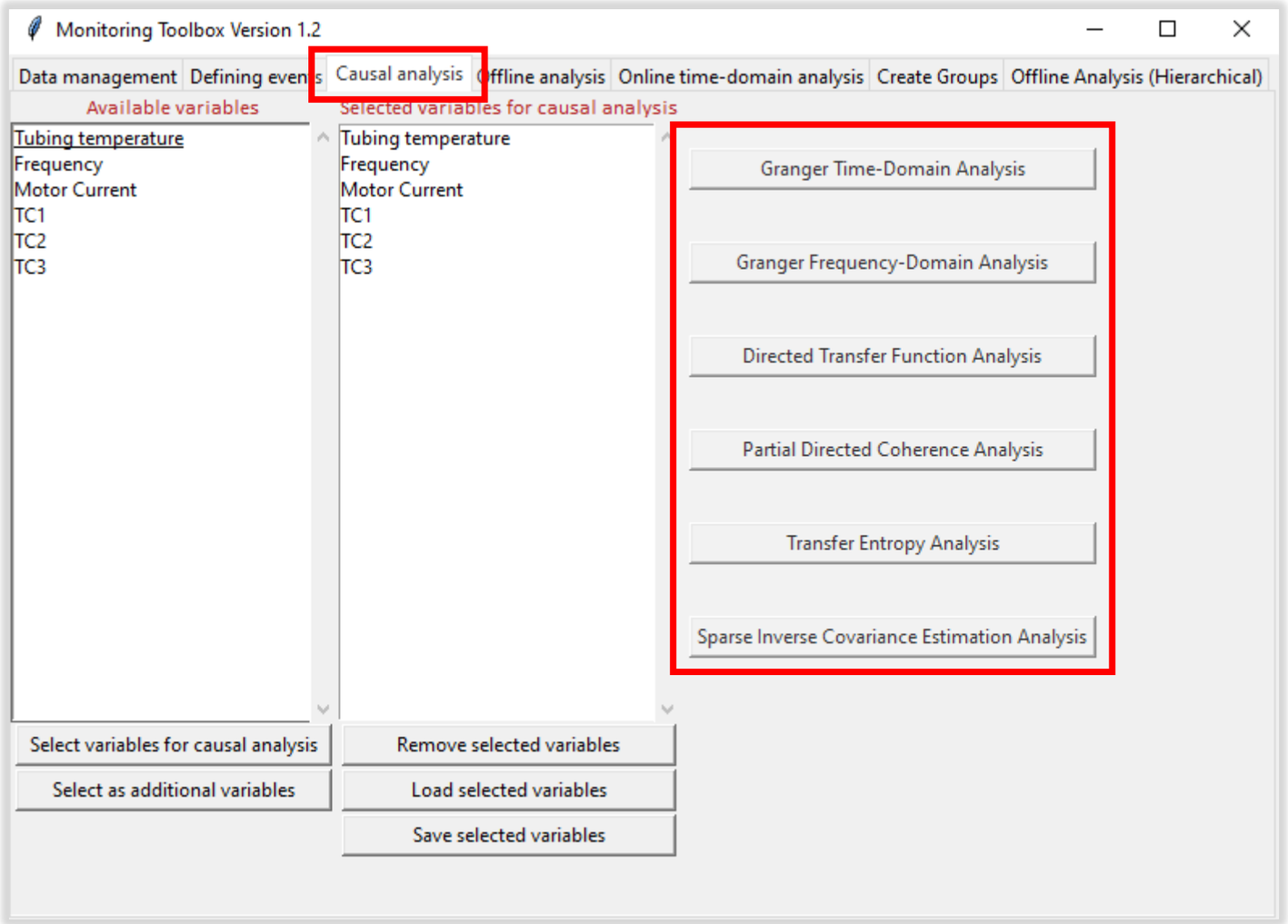




Predictive Monitoring Toolbox

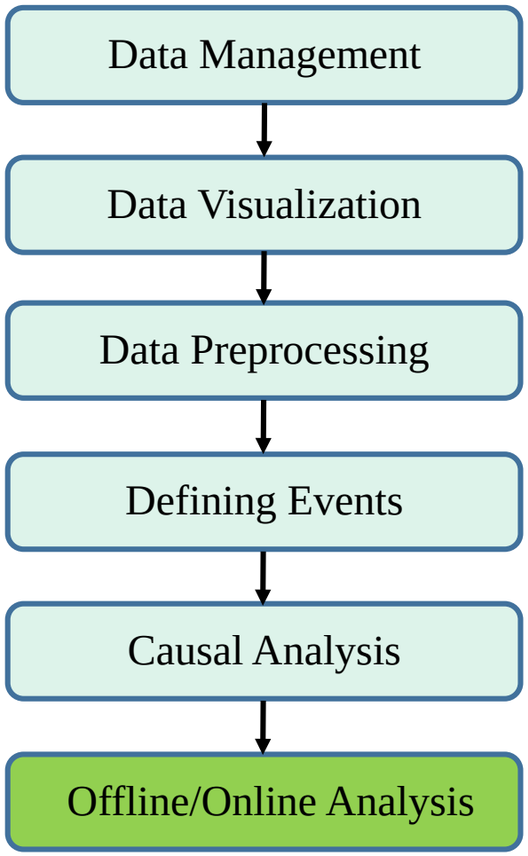
An overview of the toolbox – Causal analysis

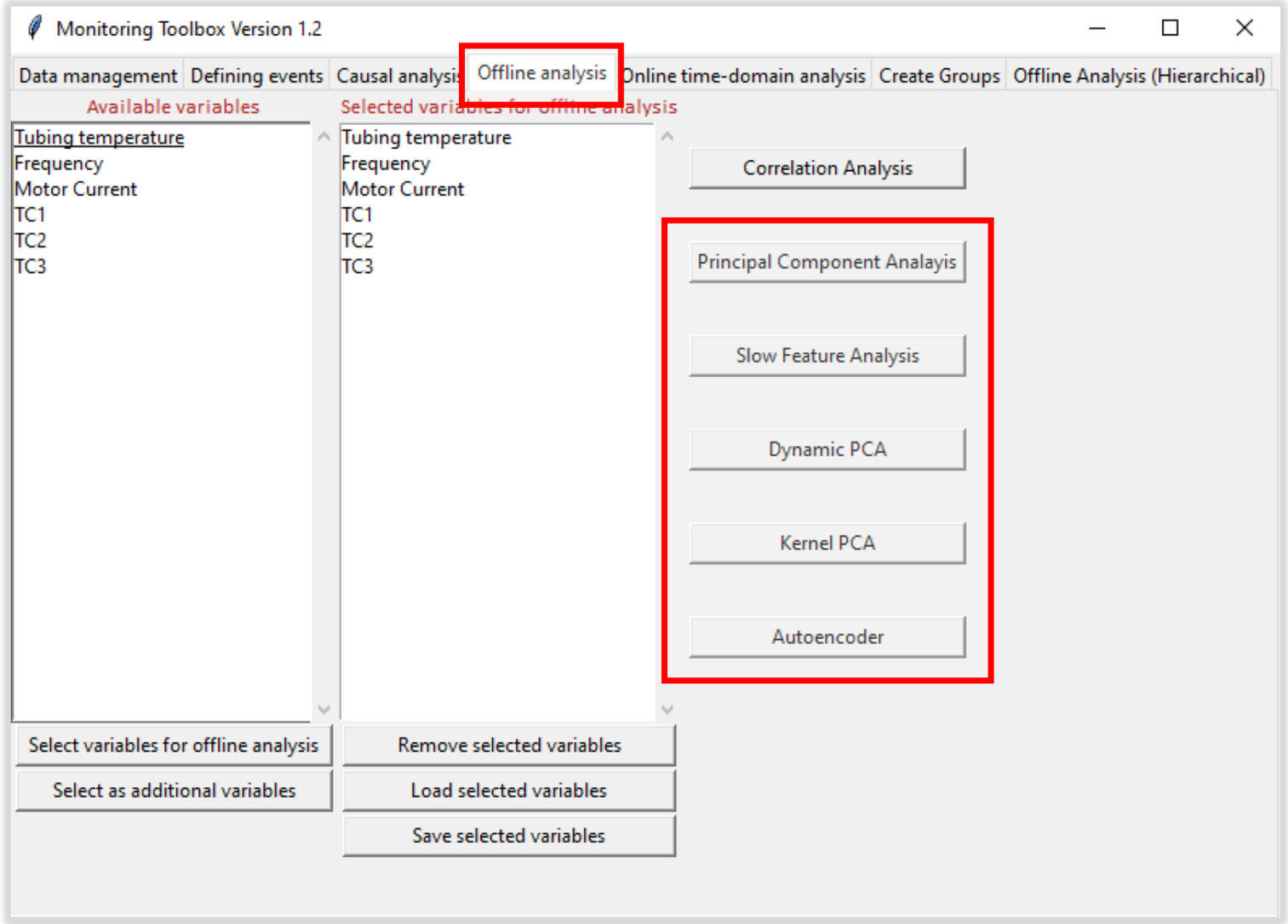




Predictive Monitoring Toolbox

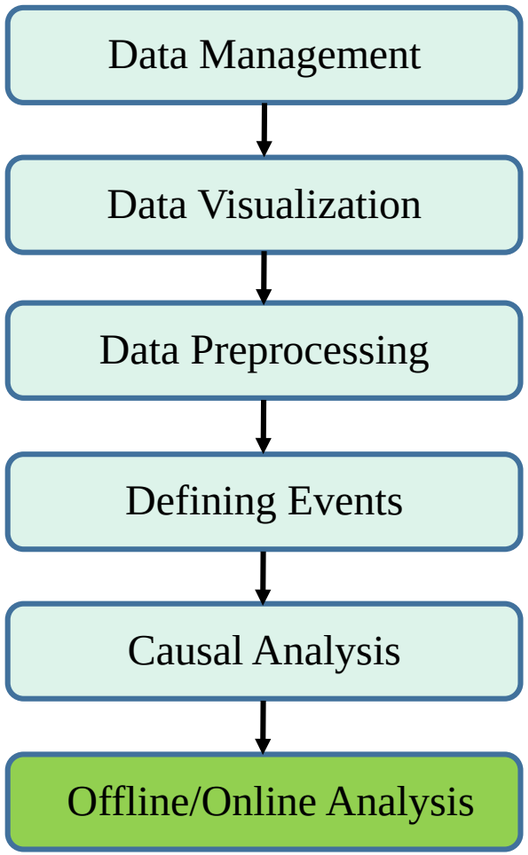
An overview of the toolbox – Offline analysis

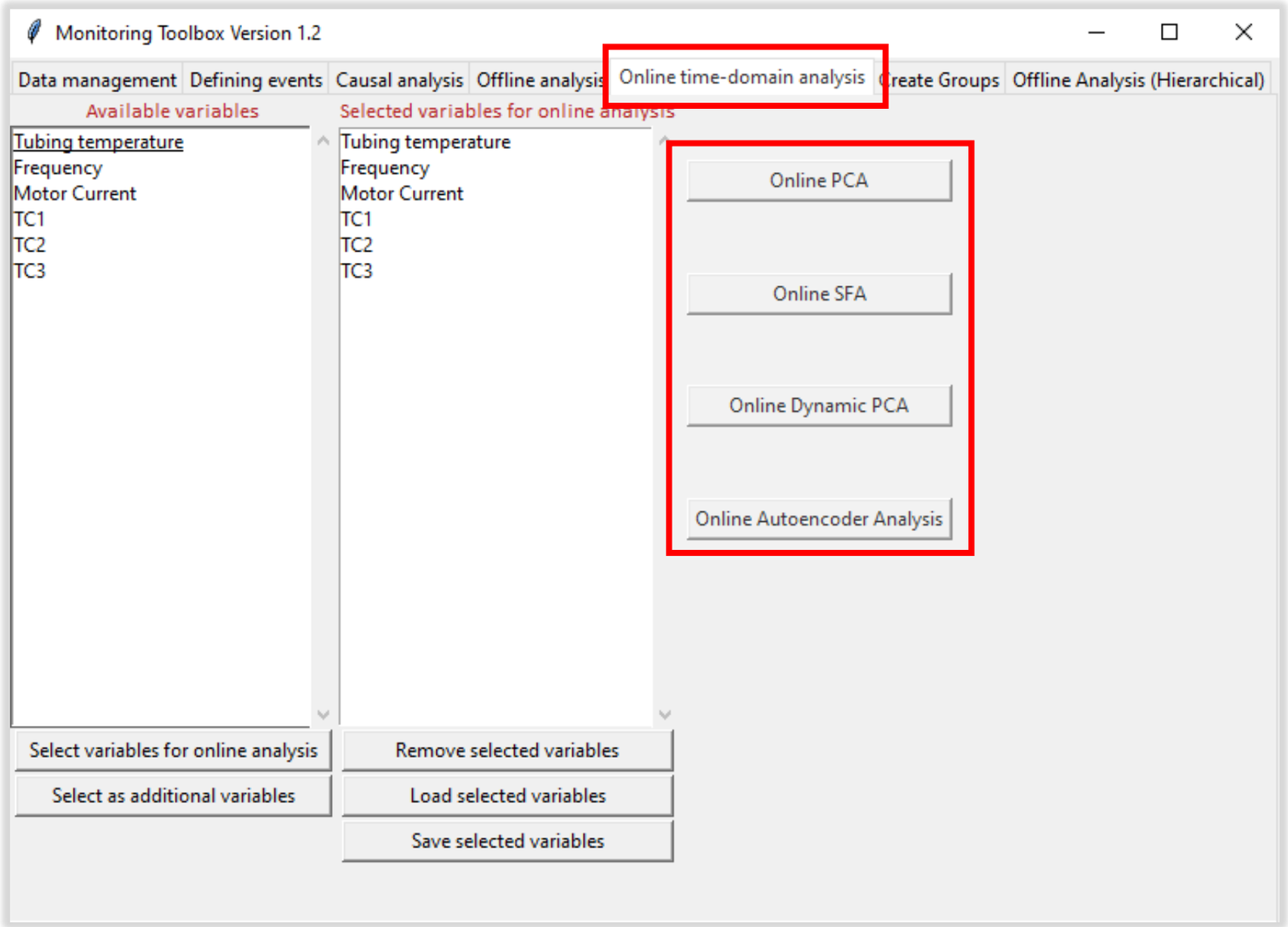




Predictive Monitoring Toolbox

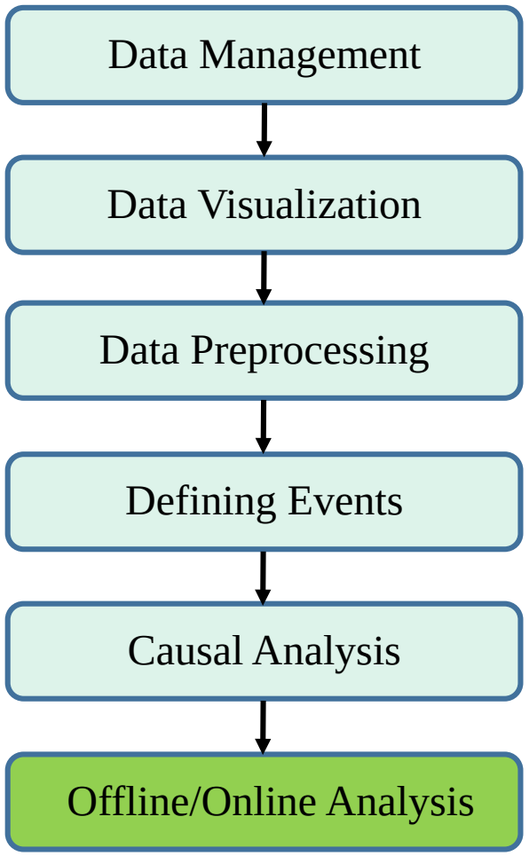
An overview of the toolbox – Online analysis

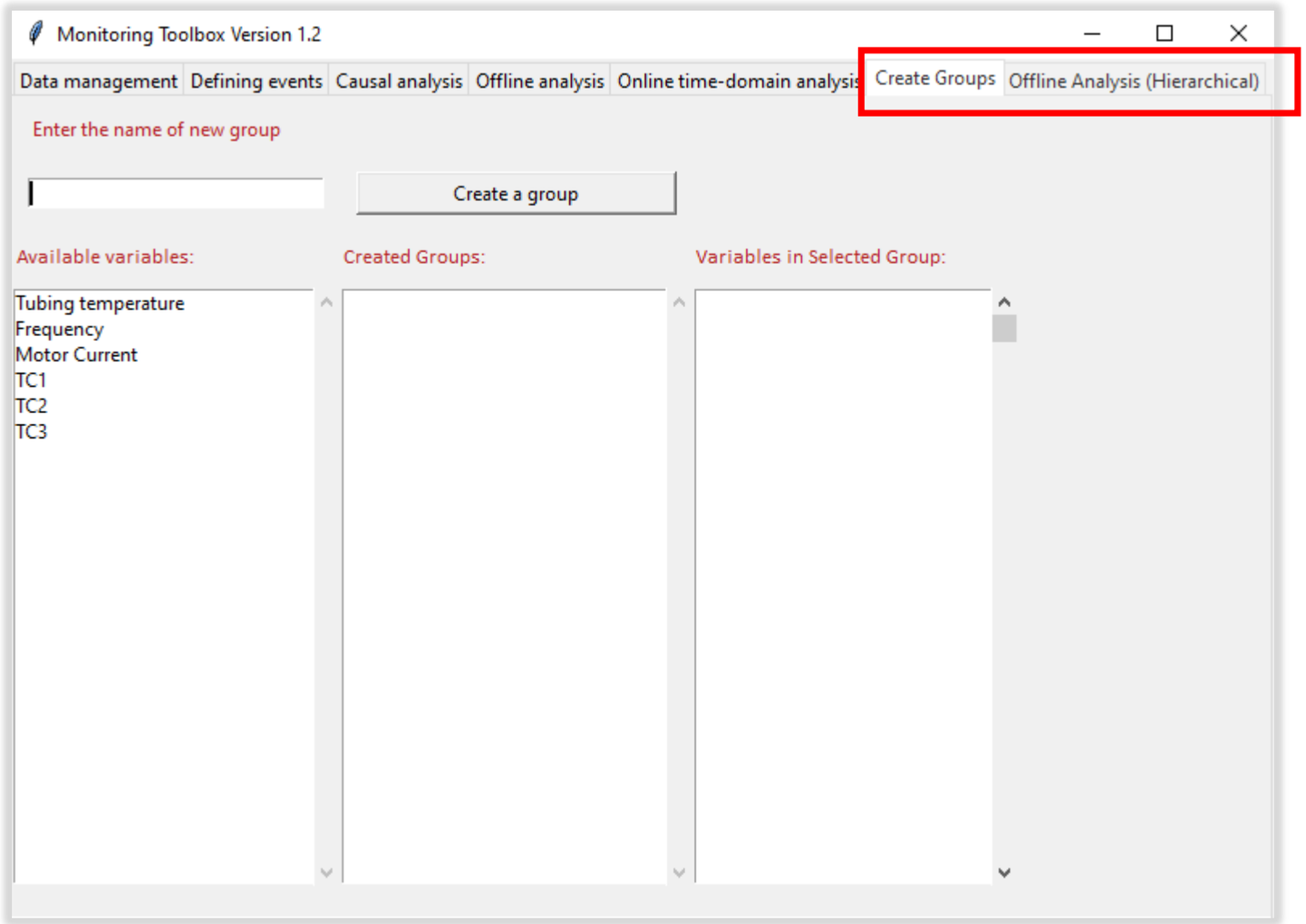




Predictive Monitoring Toolbox

An overview of the toolbox – Hierarchical distributed analysis



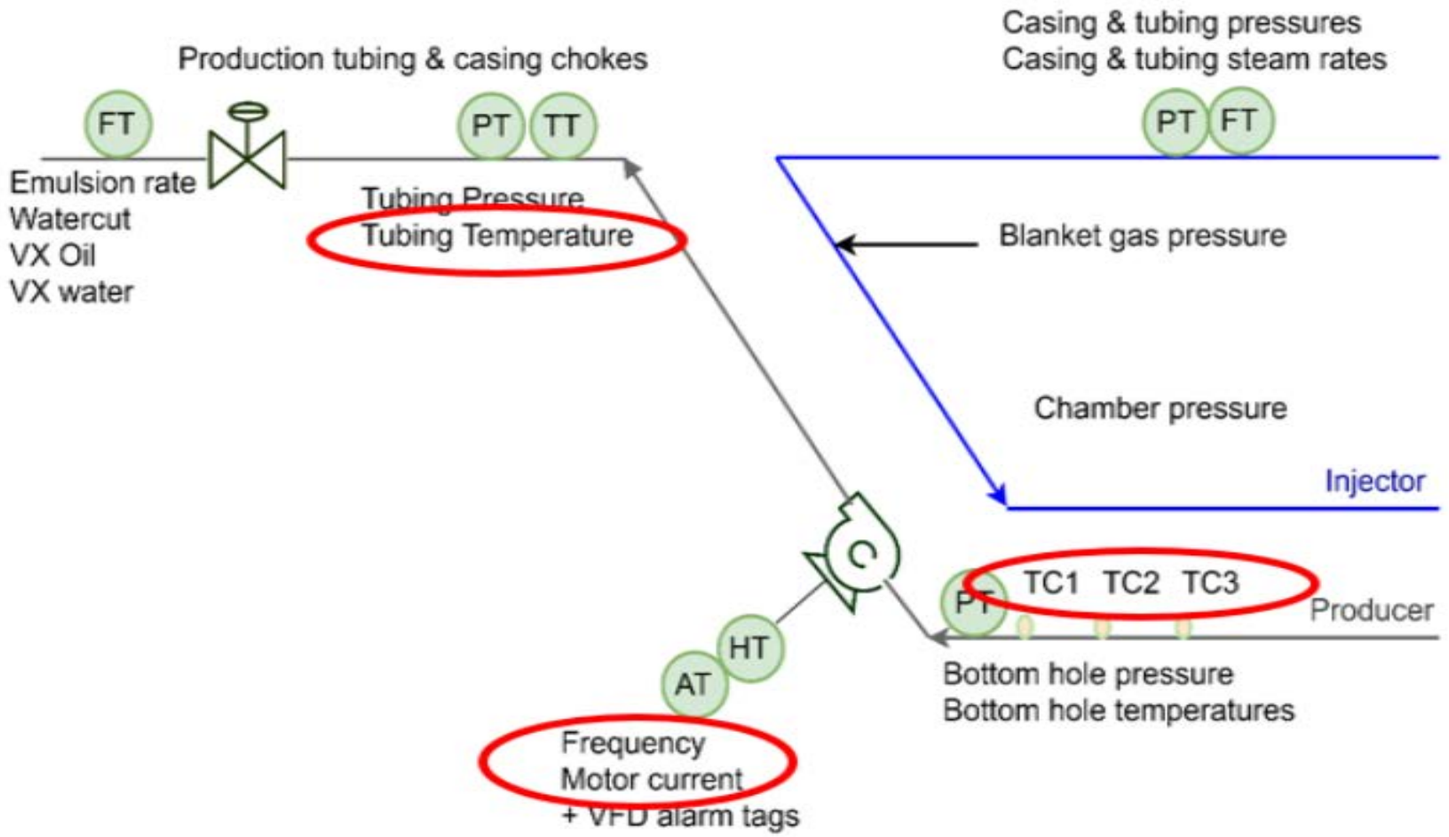


Start Guide based on a Case Study

Predictive Monitoring Toolbox

Quick Start Guide – A Case Study on Electrical Submersible Pump (ESP)

- Abnormal events took place in this process.
- The objective is to detect the occurrence of these events based on the process data.
- This case study is based on normalized data for ESP.



Predictive Monitoring Toolbox

Data preparation

Make sure the data are prepared following the guidelines below:

- Prepare the dataset using Excel
- the first column of a spreadsheet always contains timestamps
- the first row of a spreadsheet always contains the name of variables
- The data for each variable are recorded in each corresponding column
- For the created spreadsheet, select the entire timestamp column, right-click on this column and go to **“format cells”, then** go to **“Time” or “Custom”,** change the format of time information to a format similar to: **“mm/dd/yyyy h:mm” or “yyyy-mm-dd h:mm:ss”.**
- Save the dataset in **.xlsx** format

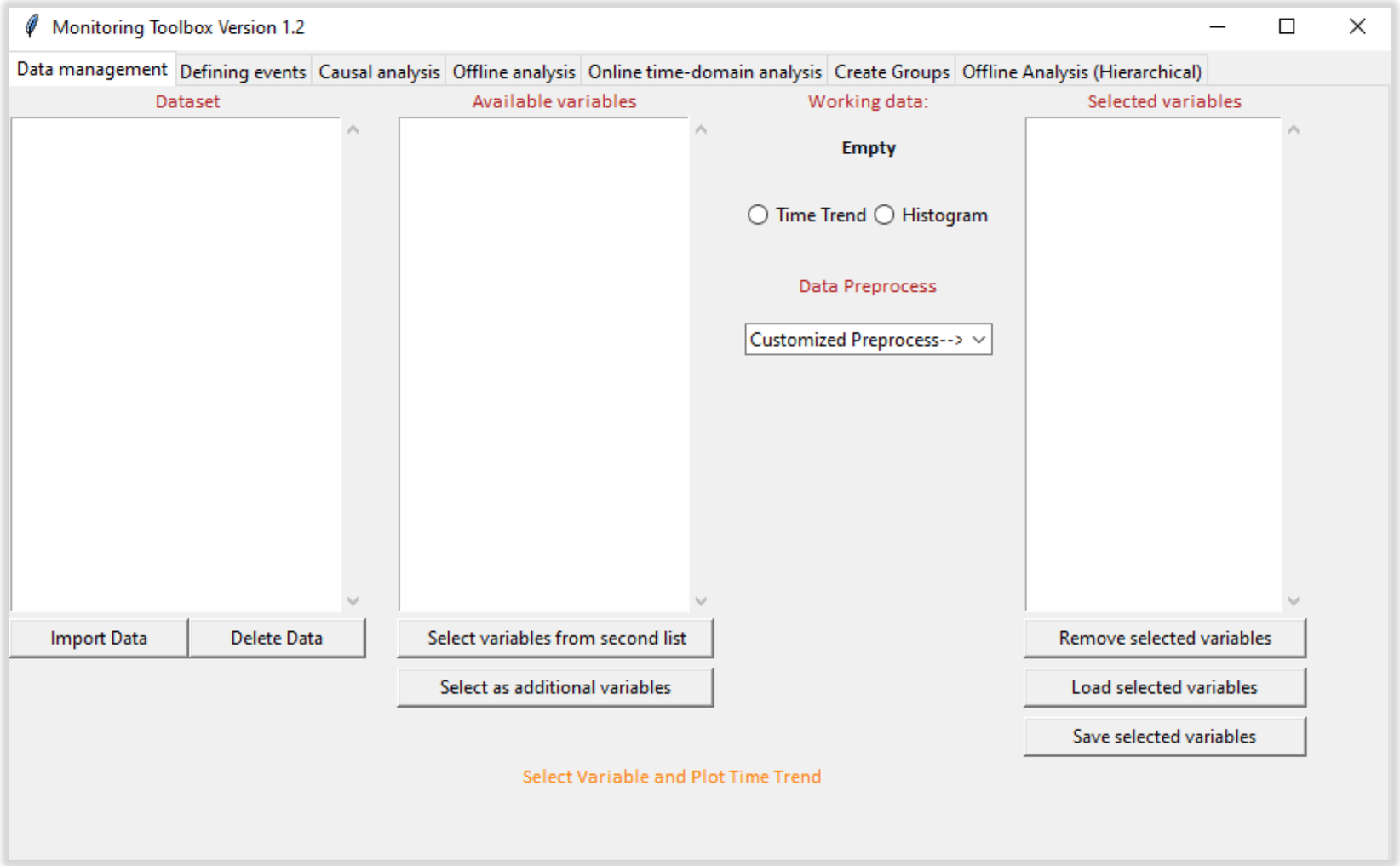
	A	B	C	D	E	F	G
1		Tubing temperature	Frequency	Motor Current	TC1	TC2	TC3
2	3/9/12 22:15	-1.364190189	0.675286424	0.49189377	-3.652629904	-3.478063792	-3.34500166
3	3/9/12 22:20	-1.498604965	0.675286424	0.587950416	-3.424377513	-3.2551307	-3.136641334
4	3/9/12 22:25	-1.302343733	0.675286424	0.49189377	-3.078778886	-2.977938815	-2.831244138
5	3/9/12 22:30	-1.184504911	0.675286424	0.49189377	-2.731496714	-2.494617748	-2.220262236
6	3/9/12 22:35	-1.118941918	0.675286424	0.324671085	-2.365448872	-2.020655317	-1.881266767
7	3/9/12 22:40	-1.053378925	0.675286424	0.49189377	-1.999401031	-1.690395652	-1.573328509
8	3/9/12 22:45	-0.987815931	0.675286424	0.65817481	-1.63335319	-1.556523086	-1.397895669
9	3/9/12 22:50	-0.894946846	0.675286424	0.49189377	-1.267305349	-1.474119634	-1.308935299
10	3/9/12 22:55	-0.67836475	0.675286424	0.230444806	-1.020548617	-1.359993647	-1.335386895
11	3/9/12 23:00	-0.594235033	0.675286424	0.322946051	-0.859315487	-1.276574446	-1.370401165
12	3/9/12 23:05	-0.738220634	0.675286424	0.230444806	-0.698082356	-1.269349947	-1.343913806
13	3/9/12 23:10	-0.550686357	0.675286424	0.230444806	-0.630755193	-1.262124479	-1.295956454
14	3/9/12 23:15	-0.552768548	0.675286424	0.230444806	-0.572068276	-1.254899011	-1.238107421
15	3/9/12 23:20	-0.554850738	0.675286424	0.49189377	-0.513381358	-1.230904492	-1.211772778
16	3/9/12 23:25	-0.527884322	0.675286424	0.401066418	-0.454694441	-1.202520661	-1.169750048

Predictive Monitoring Toolbox

To start

Make sure the data sets to be used are in the same folder

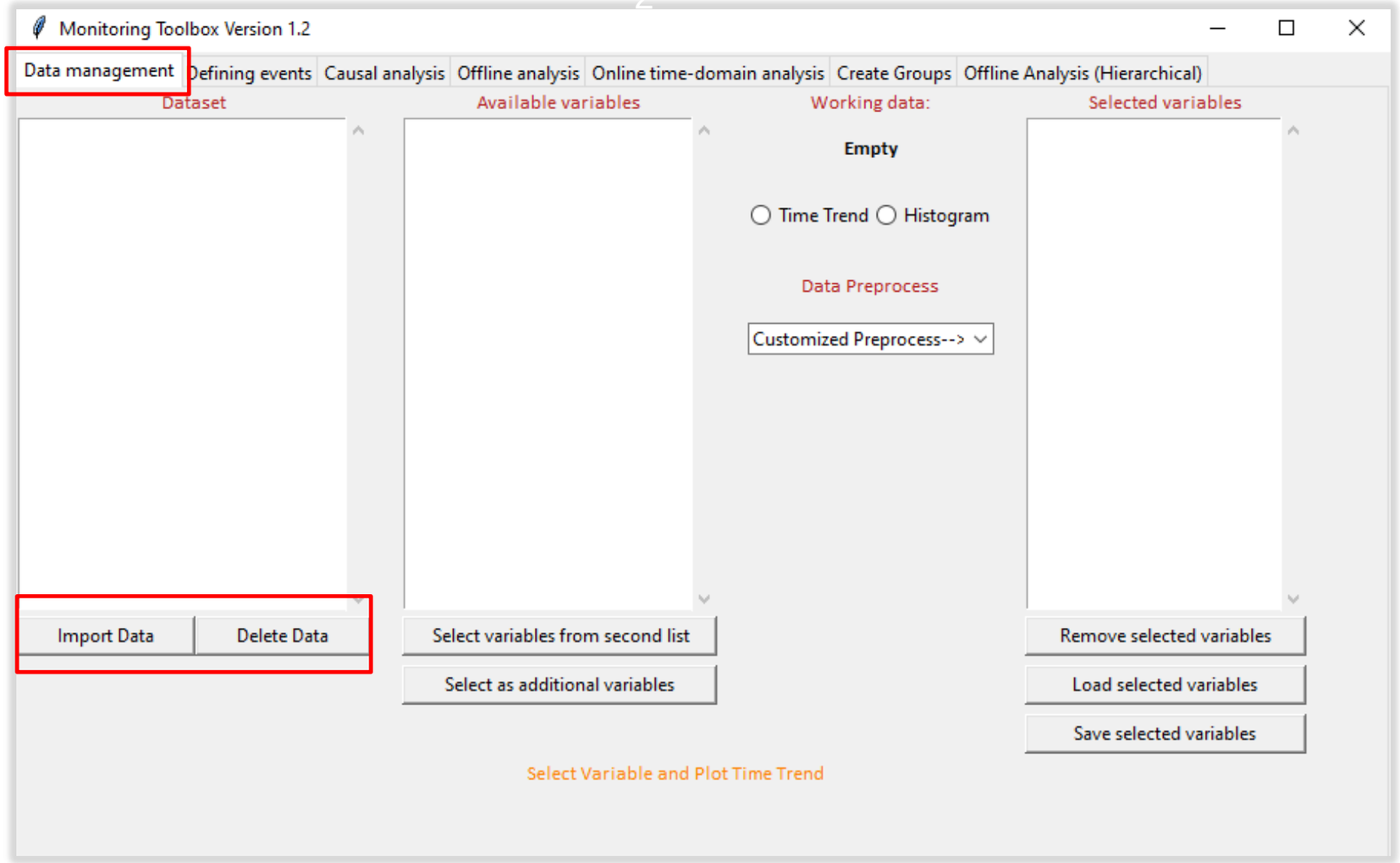
Click “Monitoring_v1.2.exe” to start the program

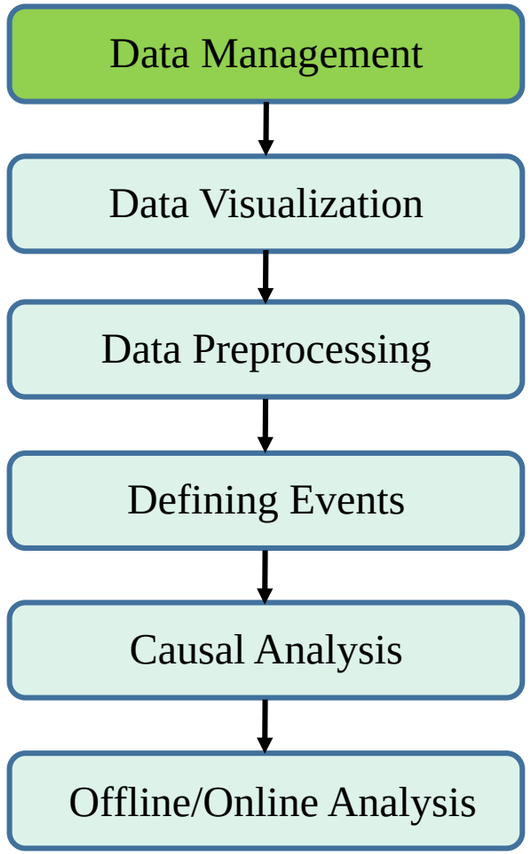


Predictive Monitoring Toolbox

Data management – Import data sets

Make sure the data sets to be used are in the same folder





Predictive Monitoring Toolbox

Data management – Import data sets

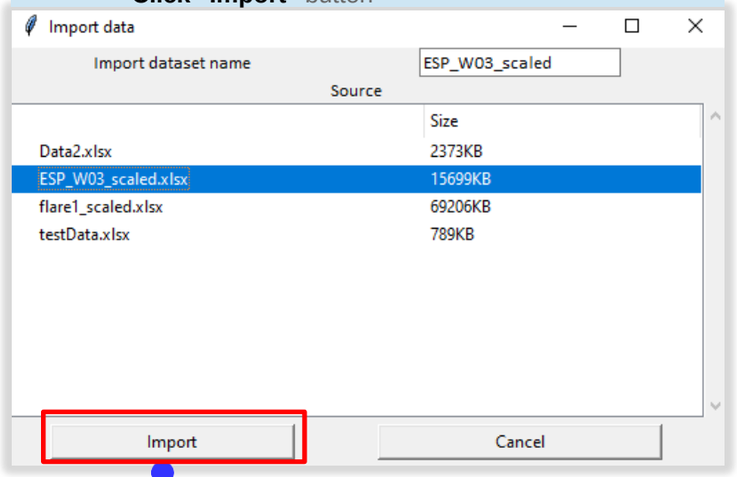
Make sure the data sets to be used are in the same folder

1

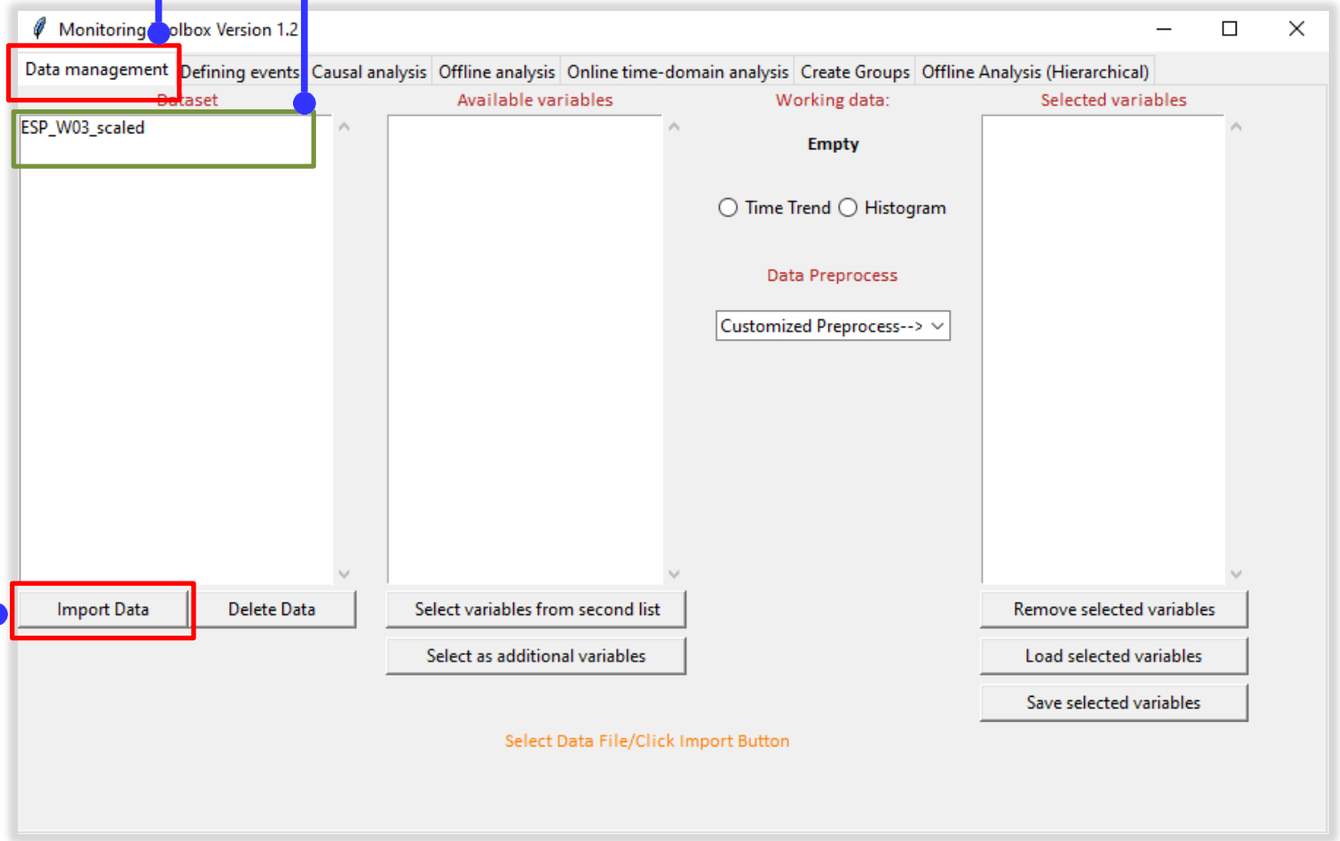
- **Go to "Data management"** tab

2

- **Click "Import Data"** button
- **Select** data file **"EP_W03_scaled.xlsx"**
- **Click "Import"** button



"**ESP_W03_Scaled**" has been imported

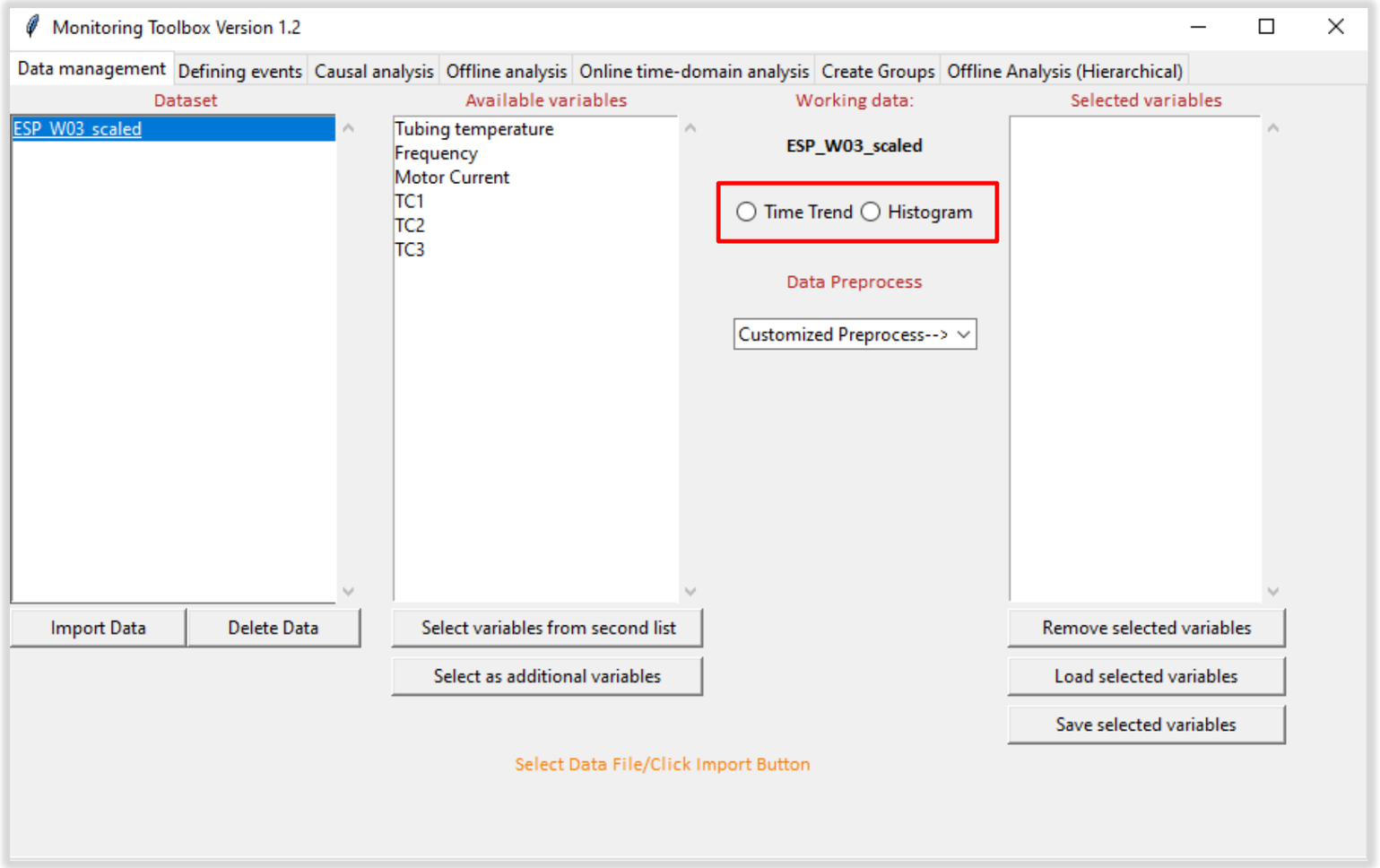


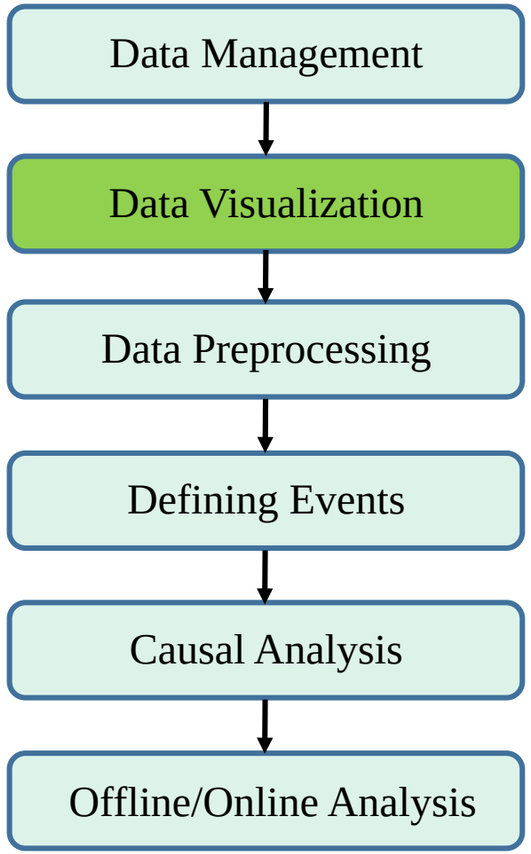
Note: Please save data sets in the ".xlsx" format

Predictive Monitoring Toolbox

Data visualization

Under the "**Data management**" tab, the variables can be visualized



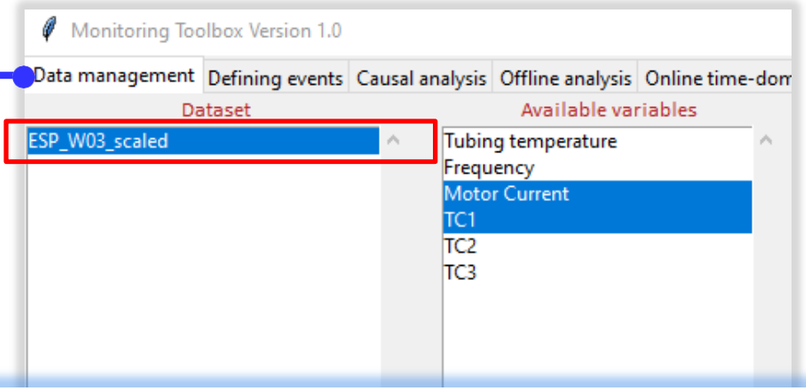


Predictive Monitoring Toolbox

Data visualization

3

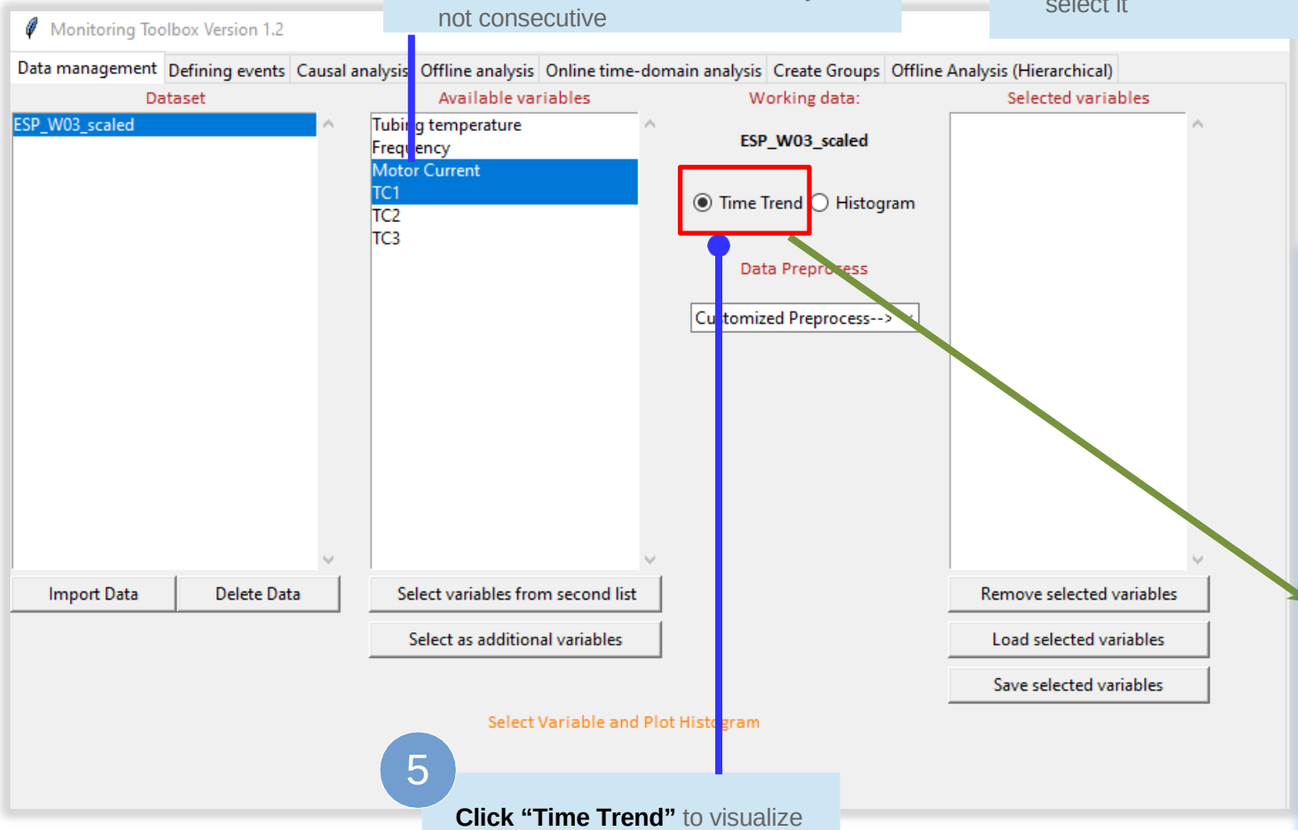
- **Click on the “Data management” tab.**
- **Click on “ESP_W03_scaled”** from the left workspace (i.e., “Dataset” column) to select it



4

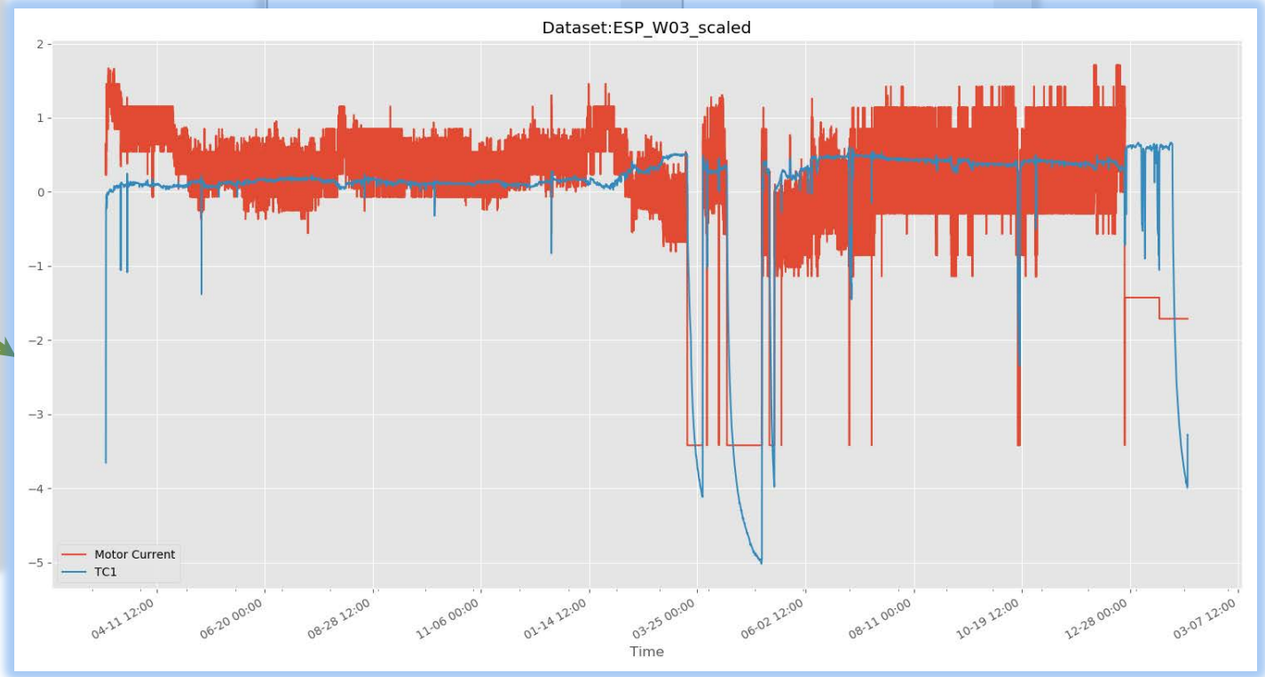
Drag-select to select variables for visualization

Hold “CTRL” to multi-select if they are not consecutive



5

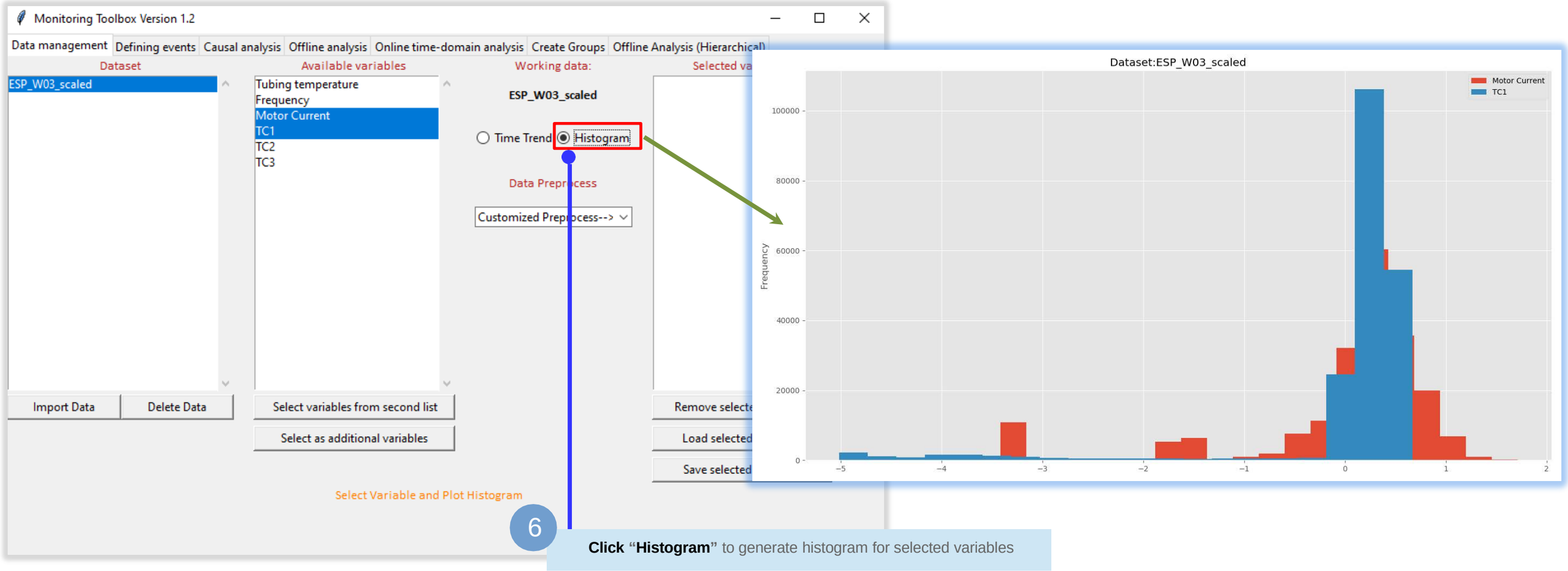
Click “Time Trend” to visualize the selected variables



Note: Please close the current popup figures before taking next action

Predictive Monitoring Toolbox

Data visualization

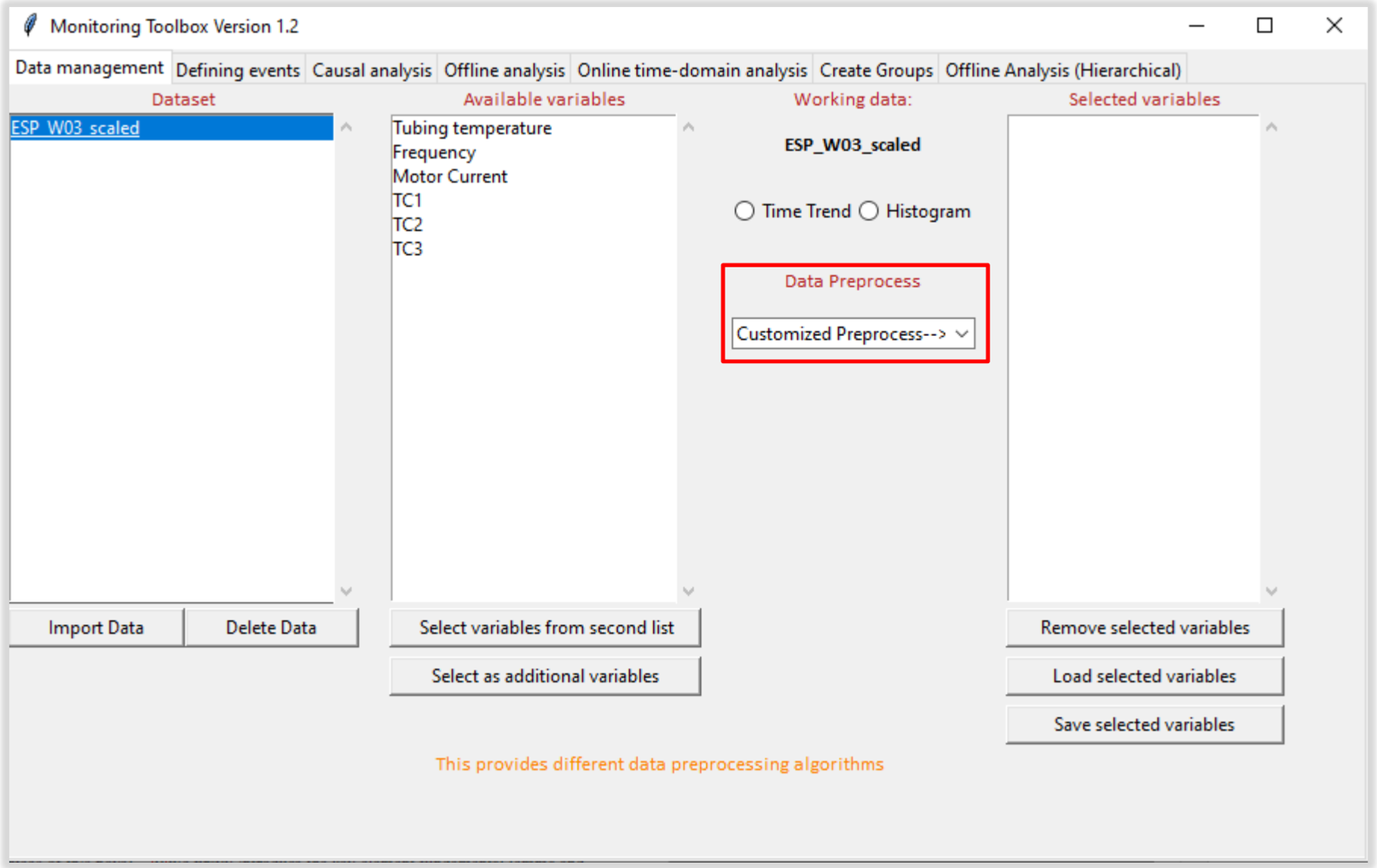


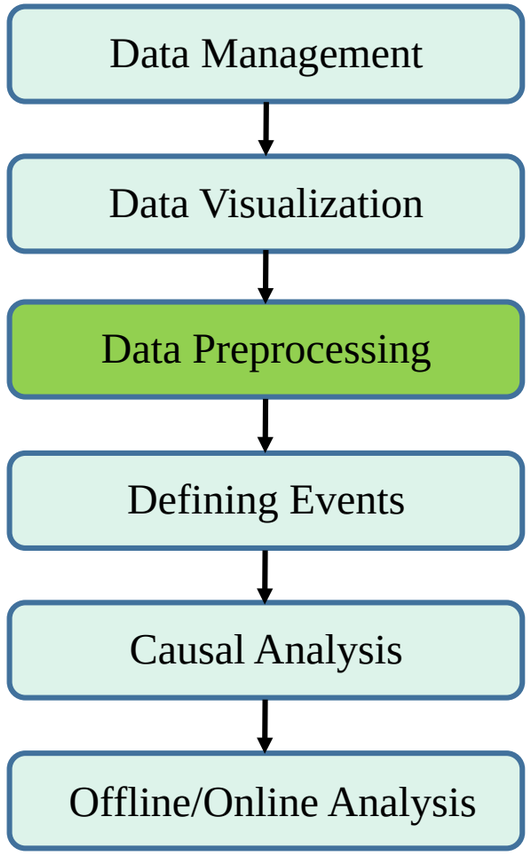
Click "Histogram" to generate histogram for selected variables

Predictive Monitoring Toolbox

Data preprocessing

Under the "**Data management**" tab, the data can be preprocessed before further analysis is conducted





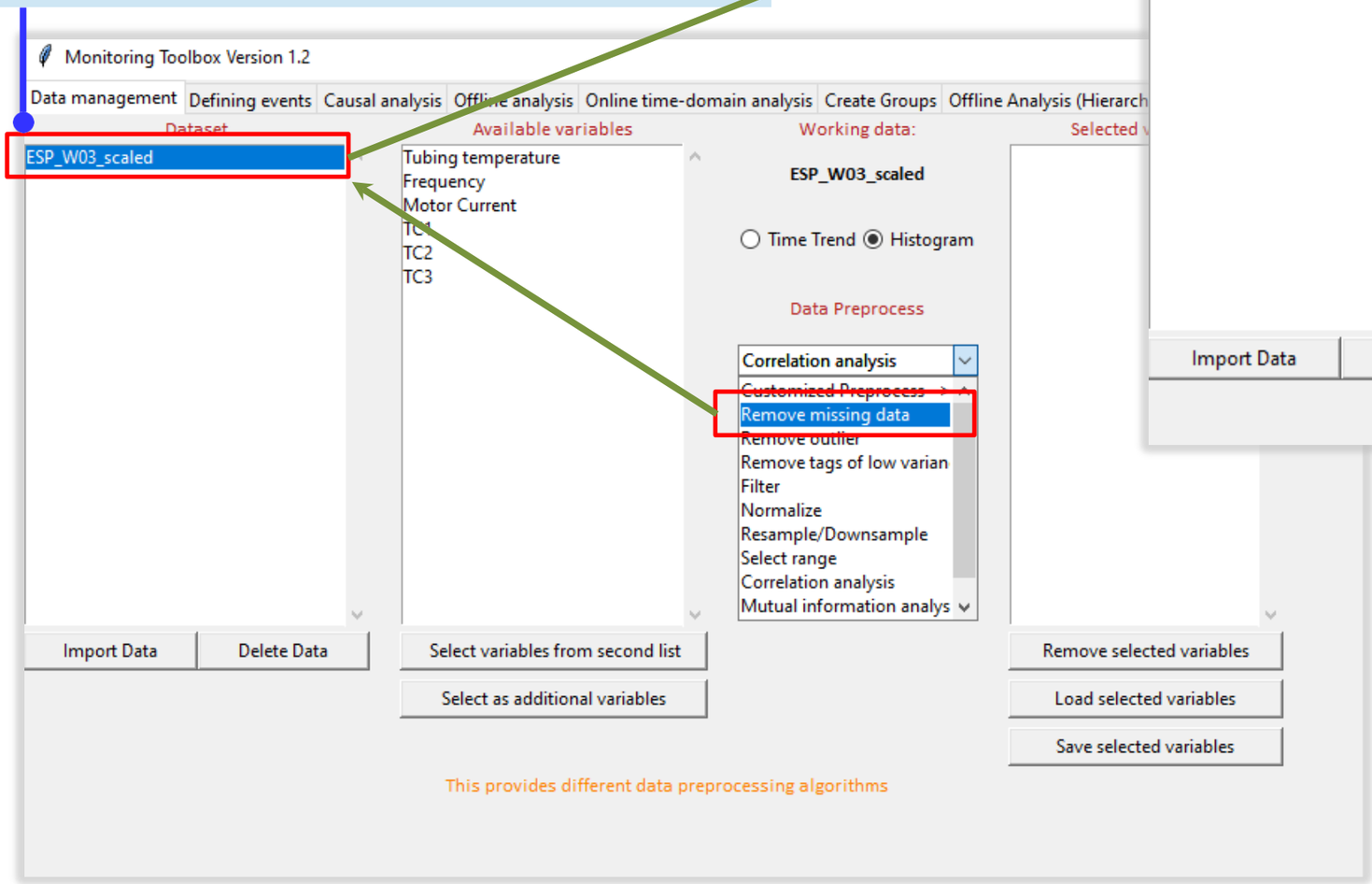
Predictive Monitoring Toolbox

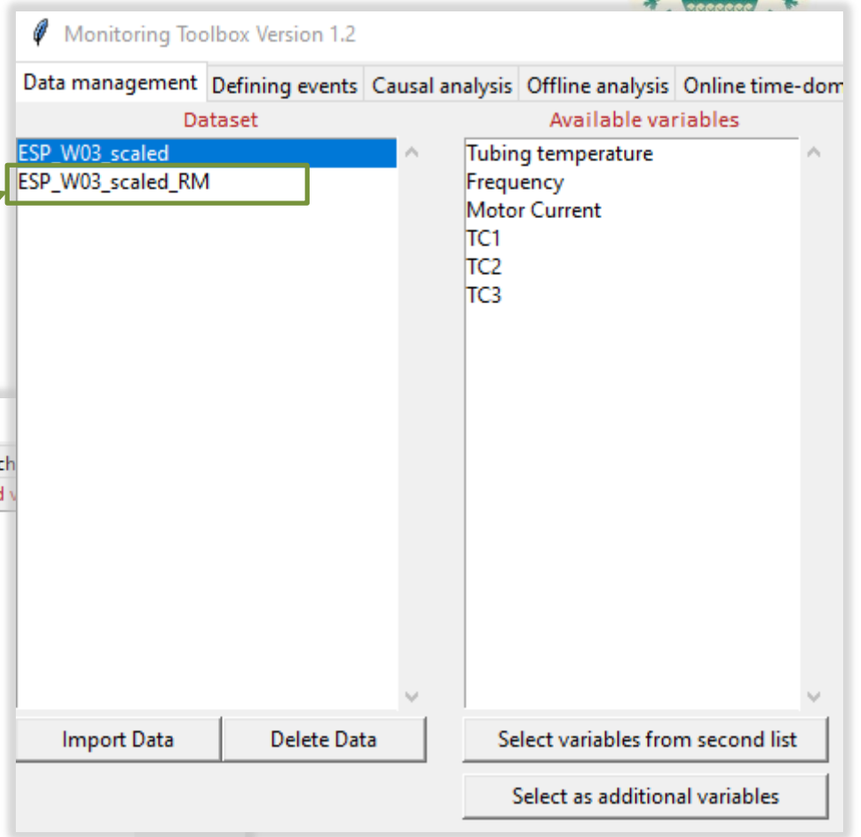
Data preprocessing – Remove missing data

1

This step is to remove missing data from the dataset

- **Select** "**ESP_W03_scaled**" from Dataset column
- **Click "Customized Preprocess -->"** on the right
- **Click** "**Remove missing data"**; new data is "**ESP_W03_scaled_RM**"





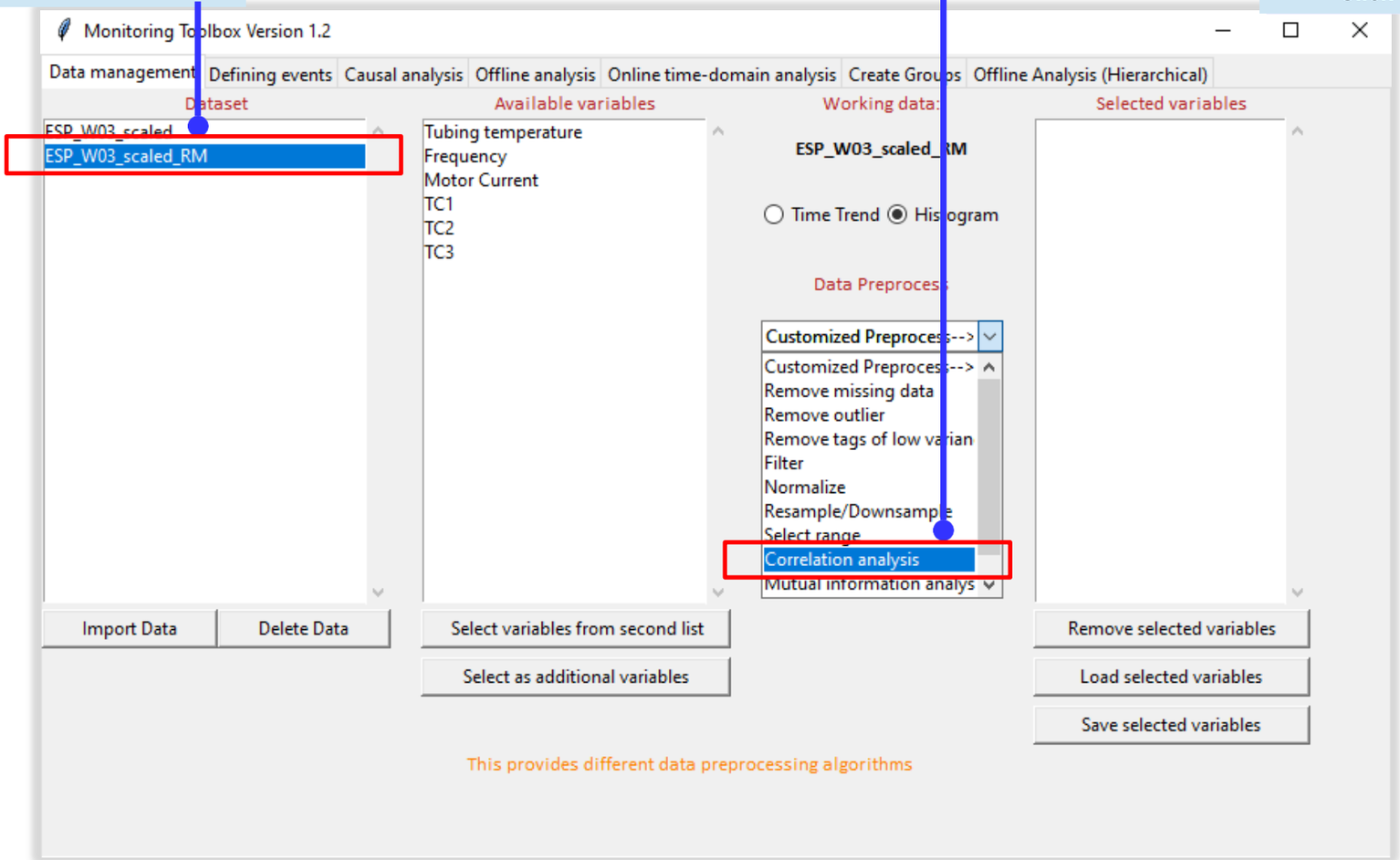
Predictive Monitoring Toolbox

Data preprocessing – Correlation analysis

Conduct correlation analysis to check the influential variables as well as the co-linearity of the data-set.

2 **Select data set for correlation analysis**
- **Click “Data management”** tab
- **Select “ESP_W03_scaled_RM”**

3 **Choose a method**
- **Click “Customized Preprocessing-->”** on the right
- **Click “Correlation Analysis”,** a window will pop up



Predictive Monitoring Toolbox

Data preprocessing – Correlation analysis

Conduct correlation analysis to check the influential variables as well as the co-linearity of the data-set.

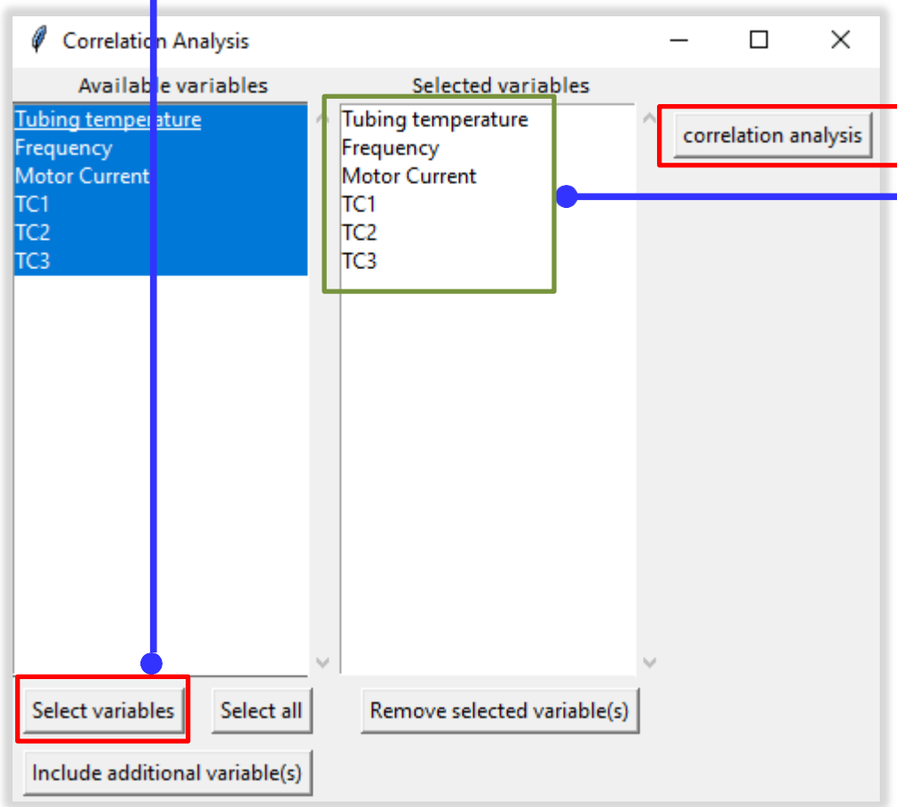
4

Choose variables to analyze

- **Drag-select (**can hold "**CTRL**"**)** in the left list to select variables **"Tubing temperature; Frequency; … TC3"**
- **Click "Select variables"**

Selected variables

The variables that are selected manually for correlation analysis are shown in the right column



5

Correlation analysis

Click "correlation analysis", a heatmap will be available, as given in the next slide

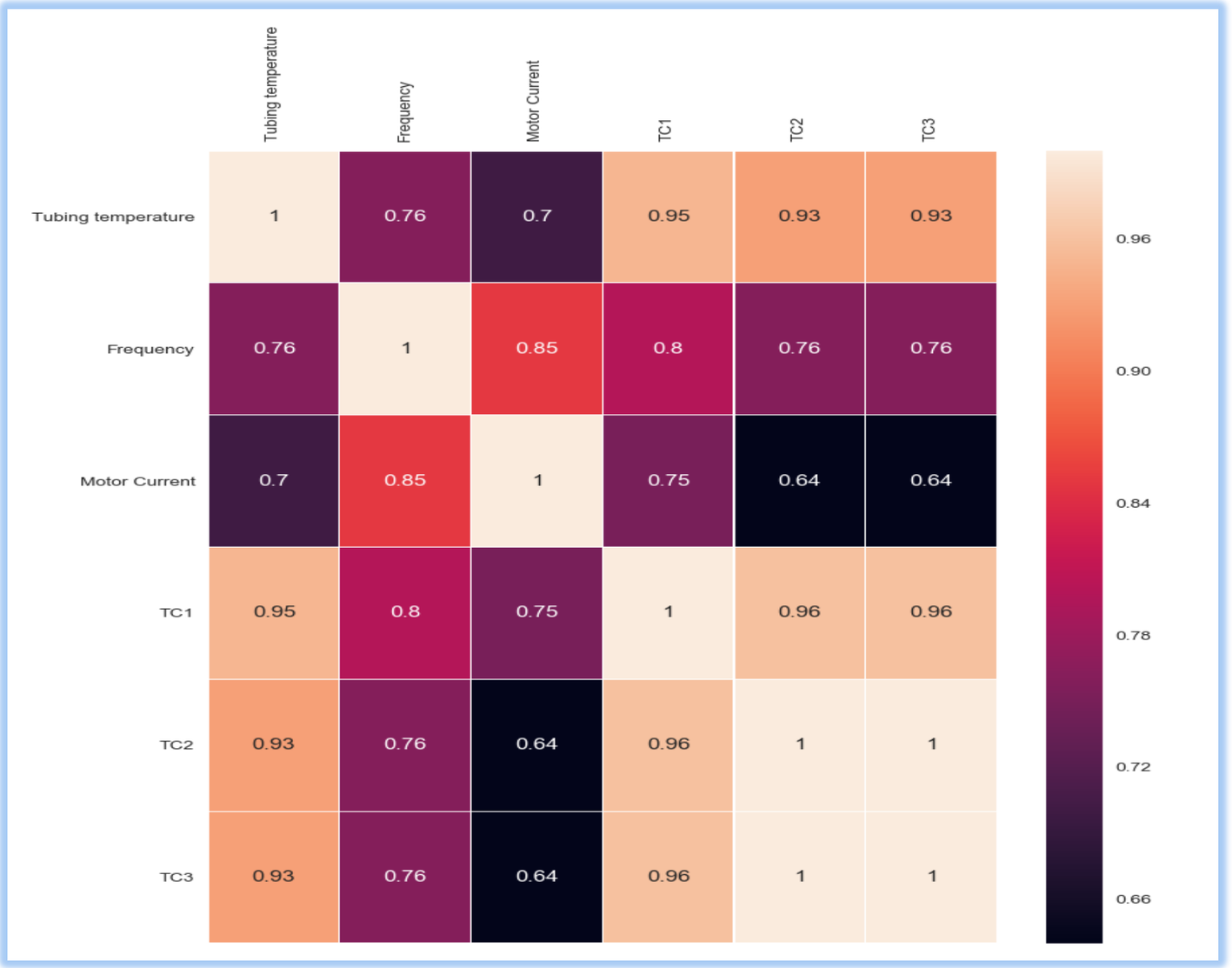
Predictive Monitoring Toolbox

Correlation analysis – Result

Results

Variables are highly correlated to each other if the absolute value of the corresponding index is large.

One can select more variables and get more information from the correlation test results.



Note: Please close the current popup figures before taking next action

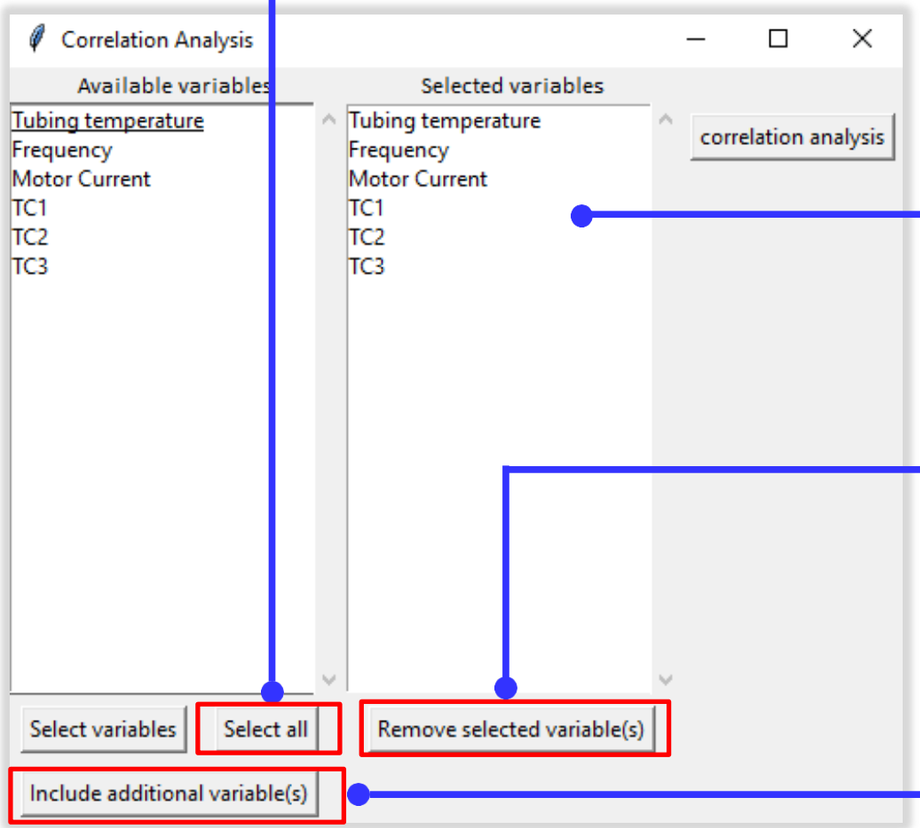
Predictive Monitoring Toolbox

Data preprocessing – Correlation analysis

If one would like to choose all variables (the number of the variables should not be too large), then can click **"Select all"**

Selected variables

The variables that are selected manually for correlation analysis are shown in the right column



If one would like to deselect a few variables, then select in the right list, then click **"Remove selected variable(s)"** to remove the ones that you would like to deselect.

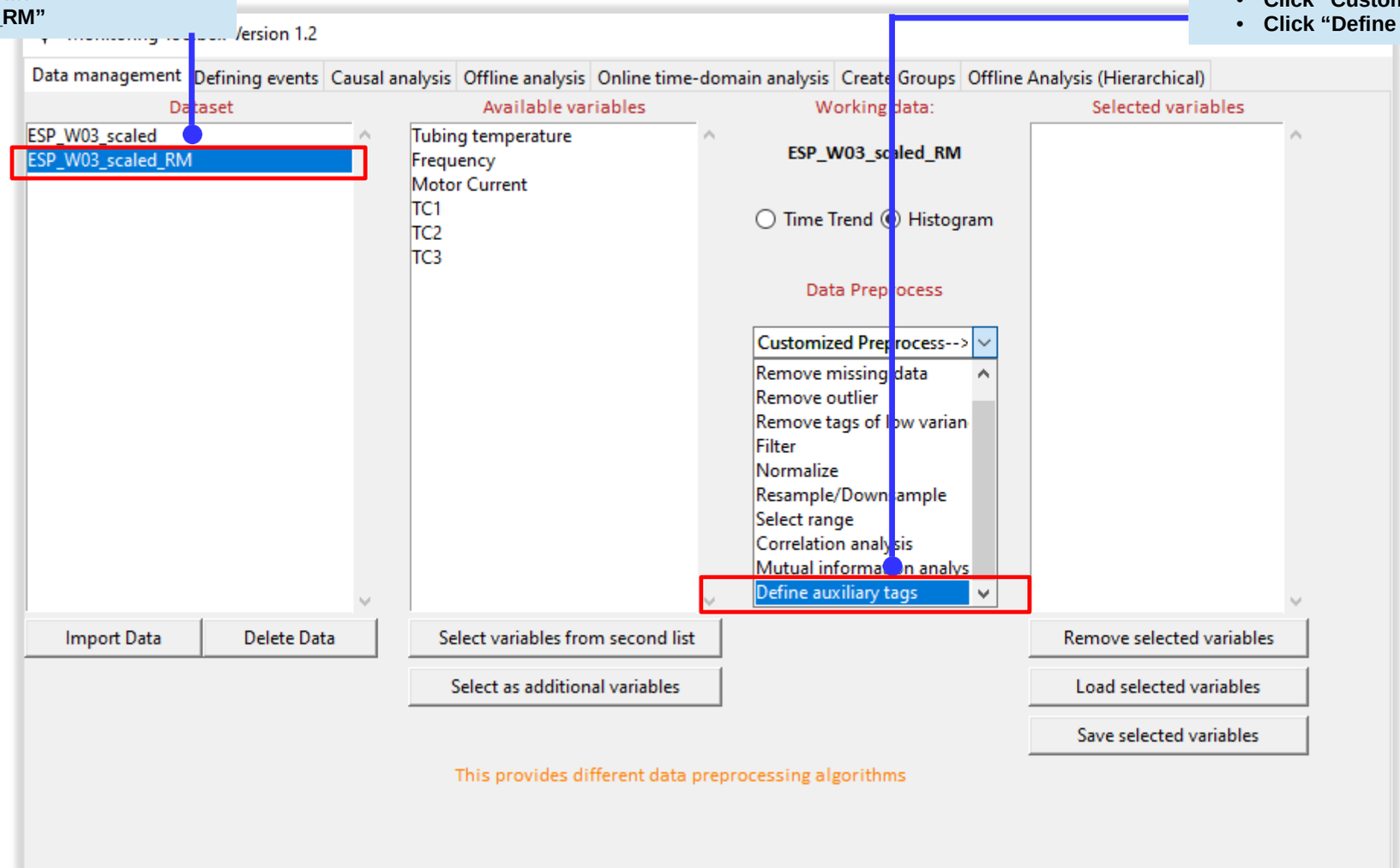
If one would like to add additional variables, then select in the right list, then click **"Included additional variable(s)"** to remove the ones that you would like to deselect.

Predictive Monitoring Toolbox

Data preprocessing – Defining auxiliary tags

1 **Select data set for correlation analysis**
- **Click “Data management”** tab
- **Select “ESP_W03_scaled_RM”**

2 **Choose method**
- **Click “Customized Preprocessing-->”** on the right
- **Click “Define auxiliary tags”,** a window will pop up



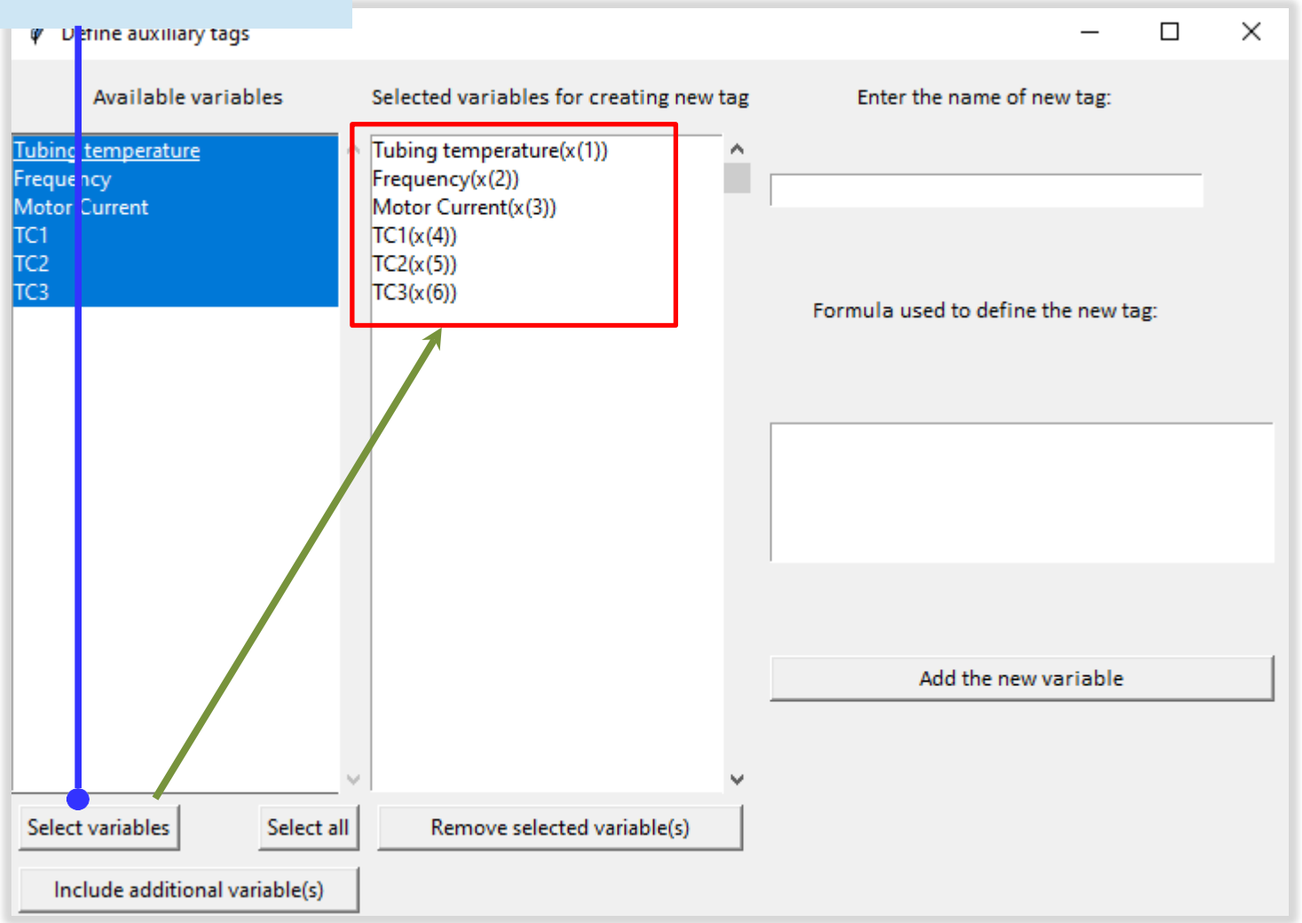
Predictive Monitoring Toolbox

Data preprocessing – Defining auxiliary tags

3

Choose variables to analyze

- **Drag-select** (can hold "**CTRL**") in the left list to select variables **"Tubing temperature; Frequency; … TC3"**
- **Click "Select variables"**



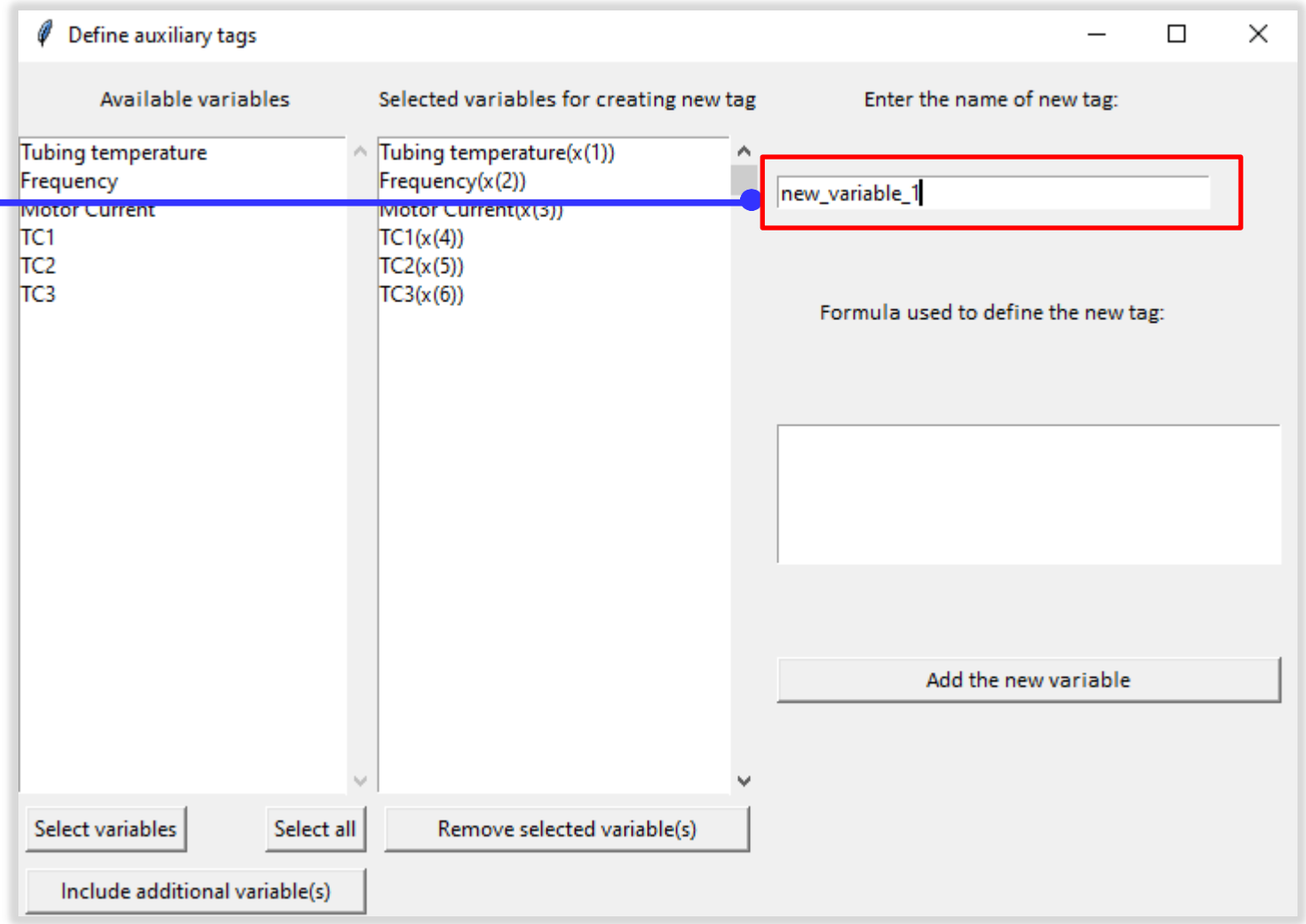
Predictive Monitoring Toolbox

Data preprocessing – Defining auxiliary tags

4

Define the name of the new variable

- **Enter “new_variable_1”** as the name of the newly defined variable



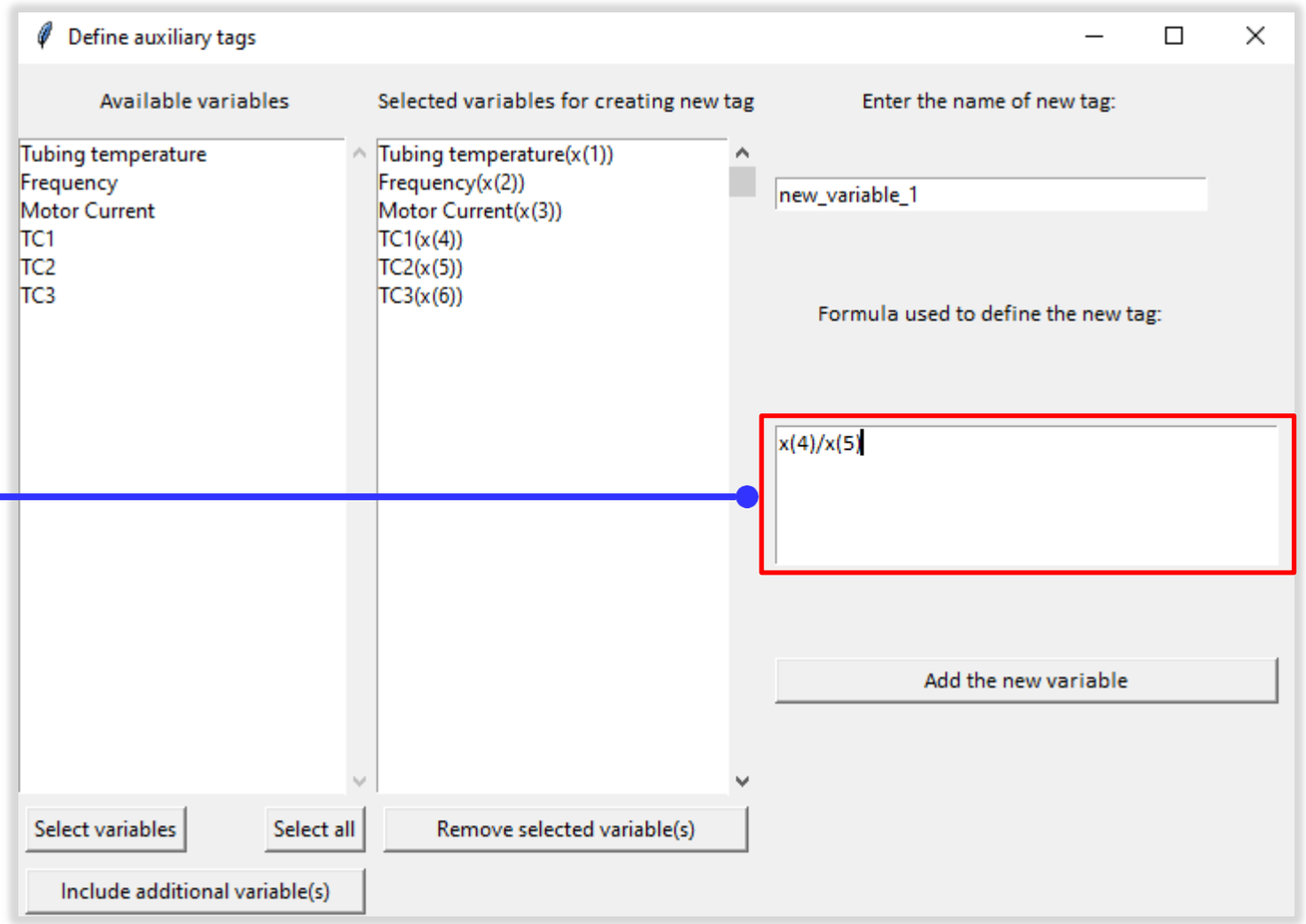
Predictive Monitoring Toolbox

Data preprocessing – Defining auxiliary tags

5

Enter the formula used for defining the new variable

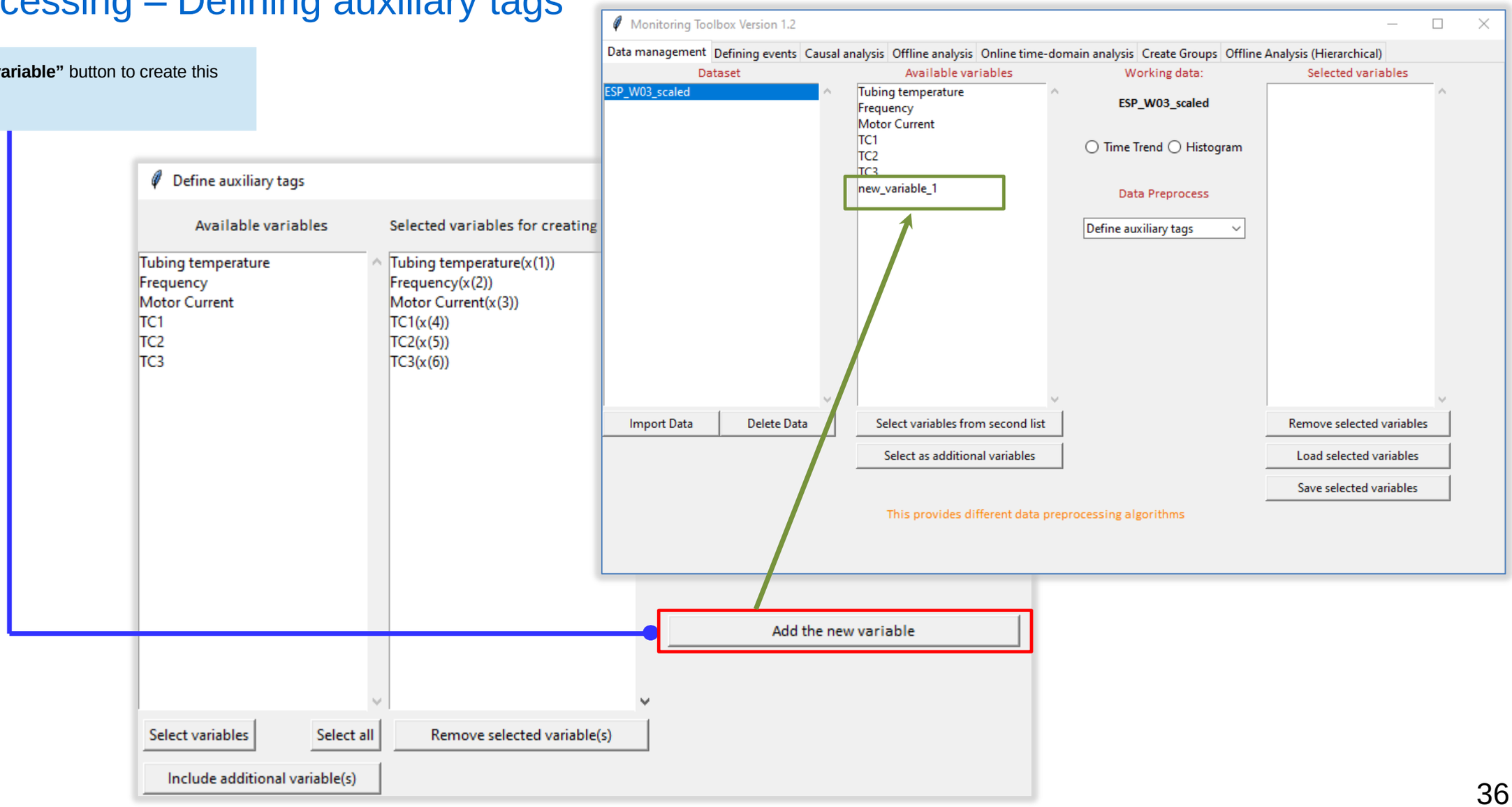
- **Enter "x(4)/x(5)"** - the new variable is calculated in a way that TC1 is divided by TC2



Predictive Monitoring Toolbox

Data preprocessing – Defining auxiliary tags

6 **Click the "Add the new variable"** button to create this new variable



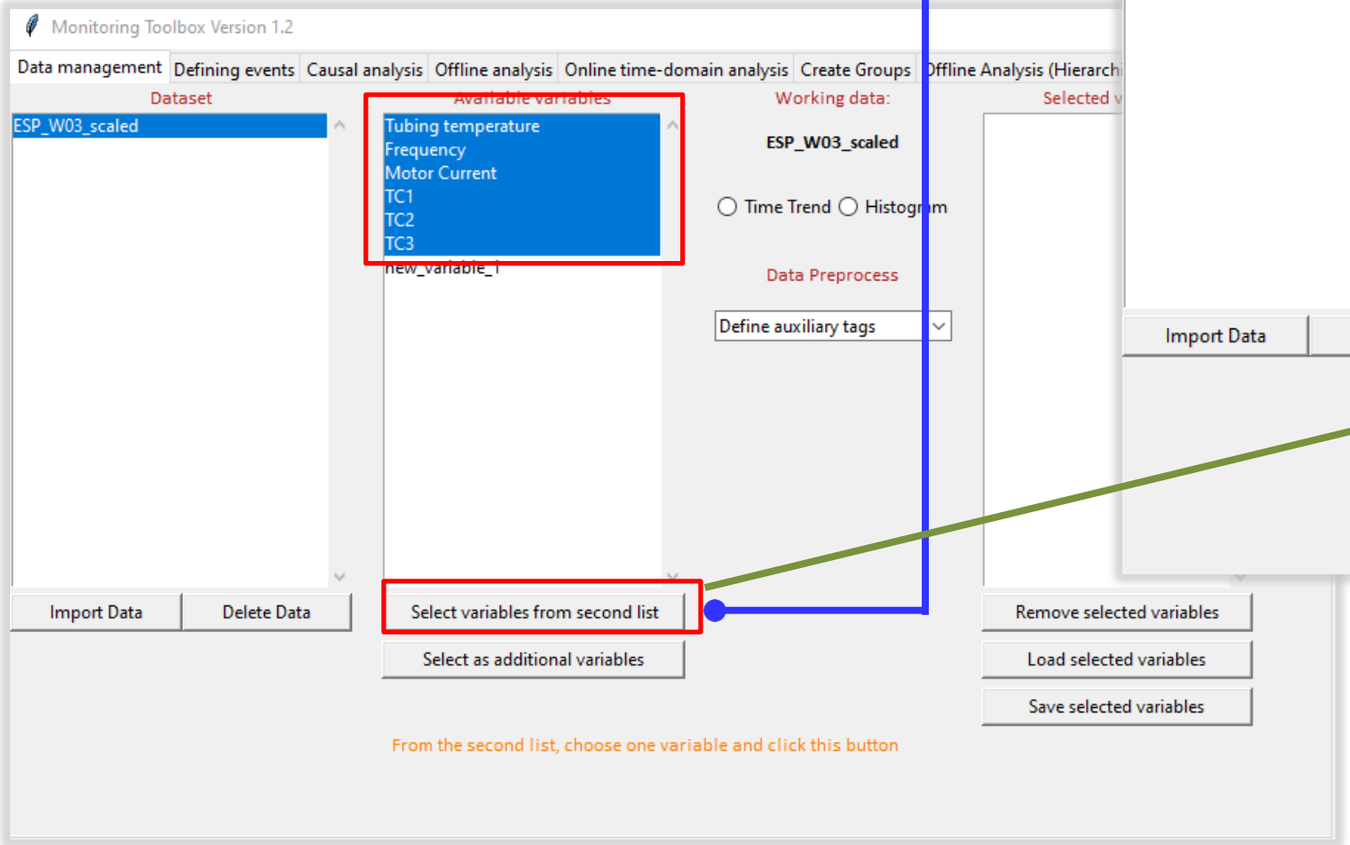
Predictive Monitoring Toolbox

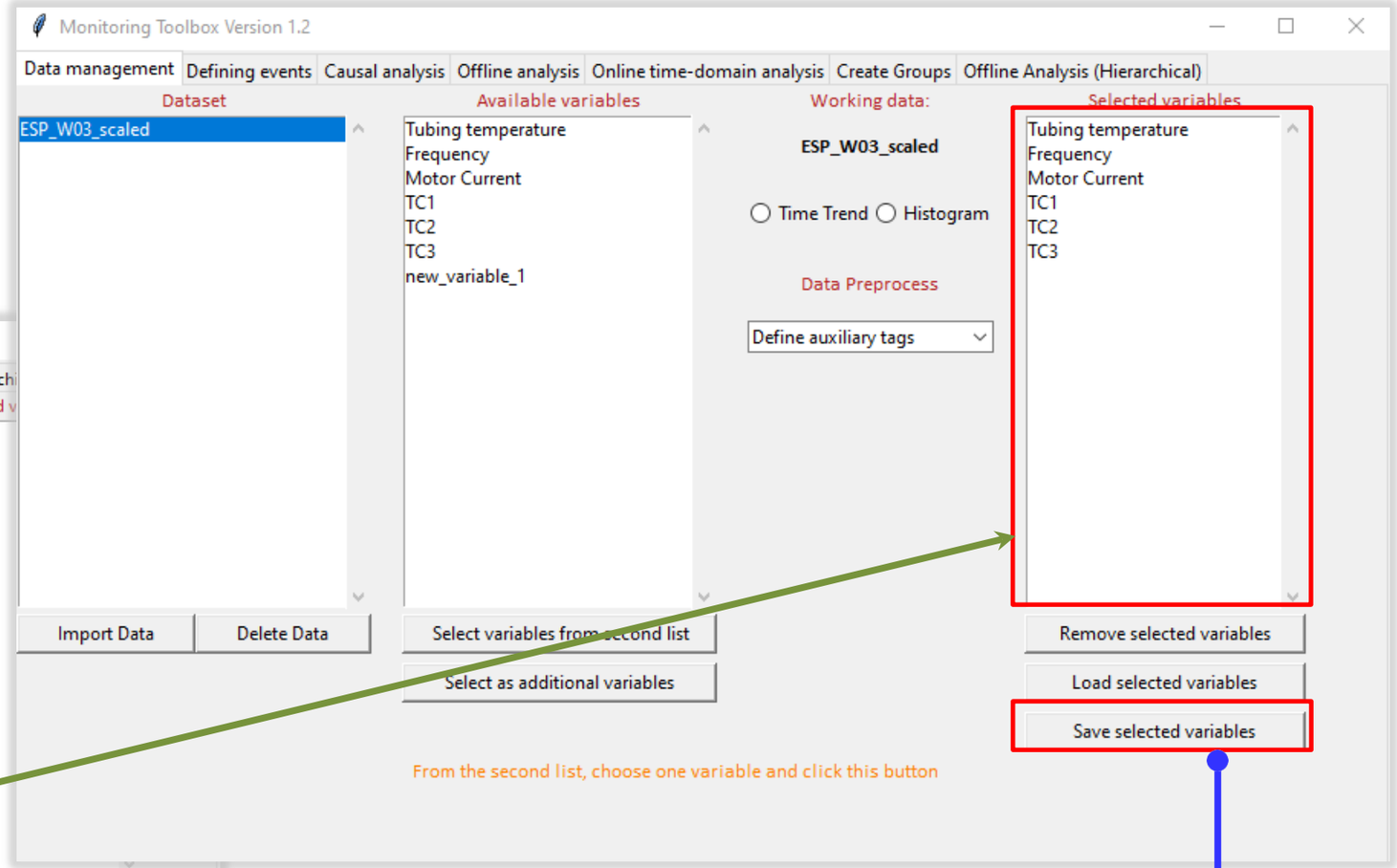
Select variables for analysis

1

Select variables from second list

- **Choose variables from the second list**
- **Click "Select variables from second list"** button to select variables for further analysis





2

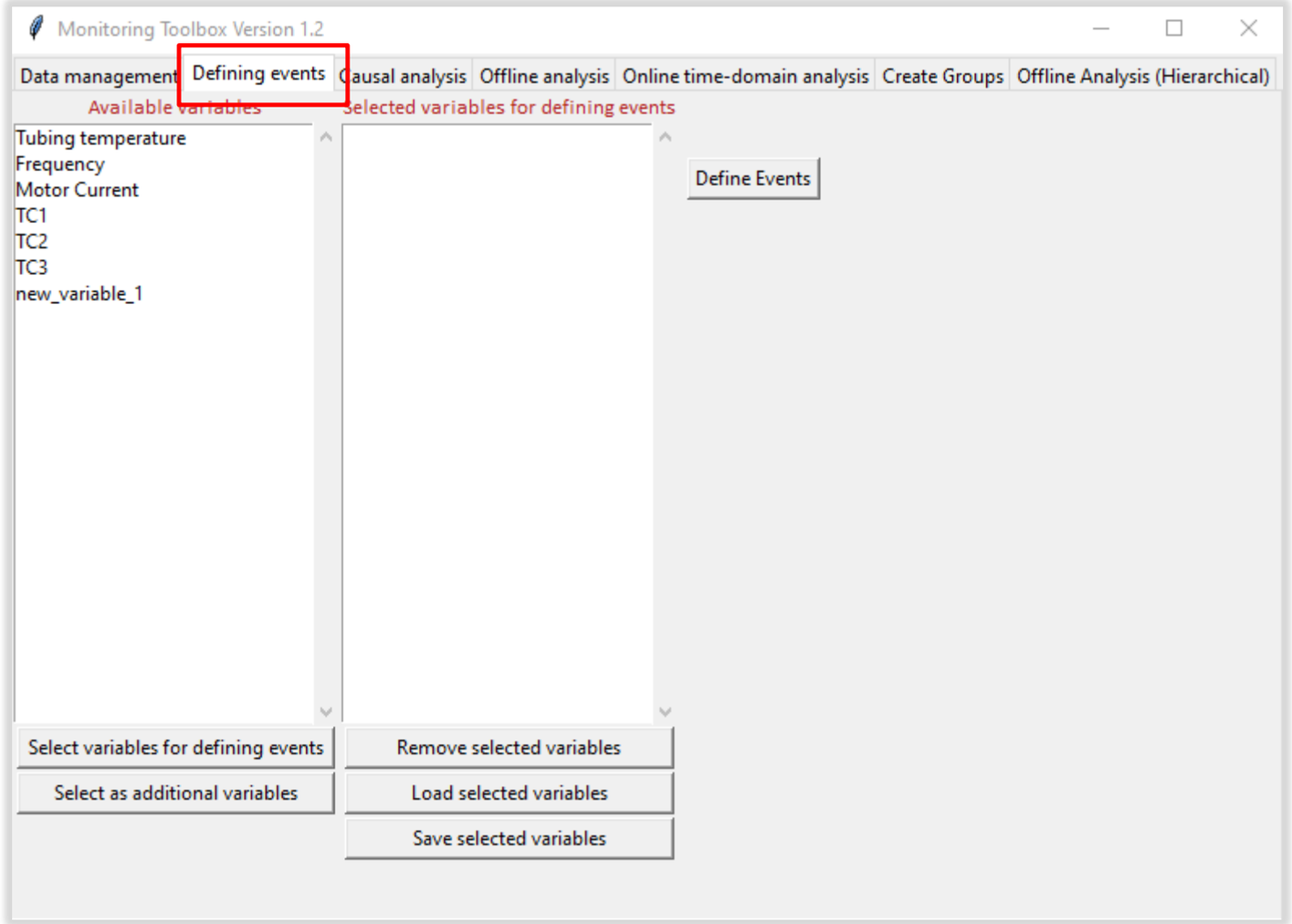
Save selected variables

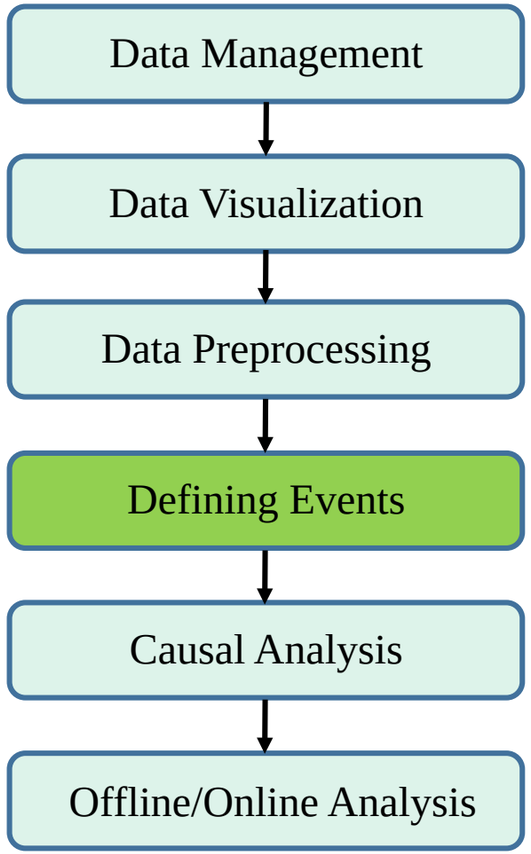
- **The variables selected are shown in the third column**
- **Click "Save selected variables"** button to save this list of variables for further analysis

Predictive Monitoring Toolbox

Defining Events

Under the "**Defining events**" tab, different types of events that took place can be defined





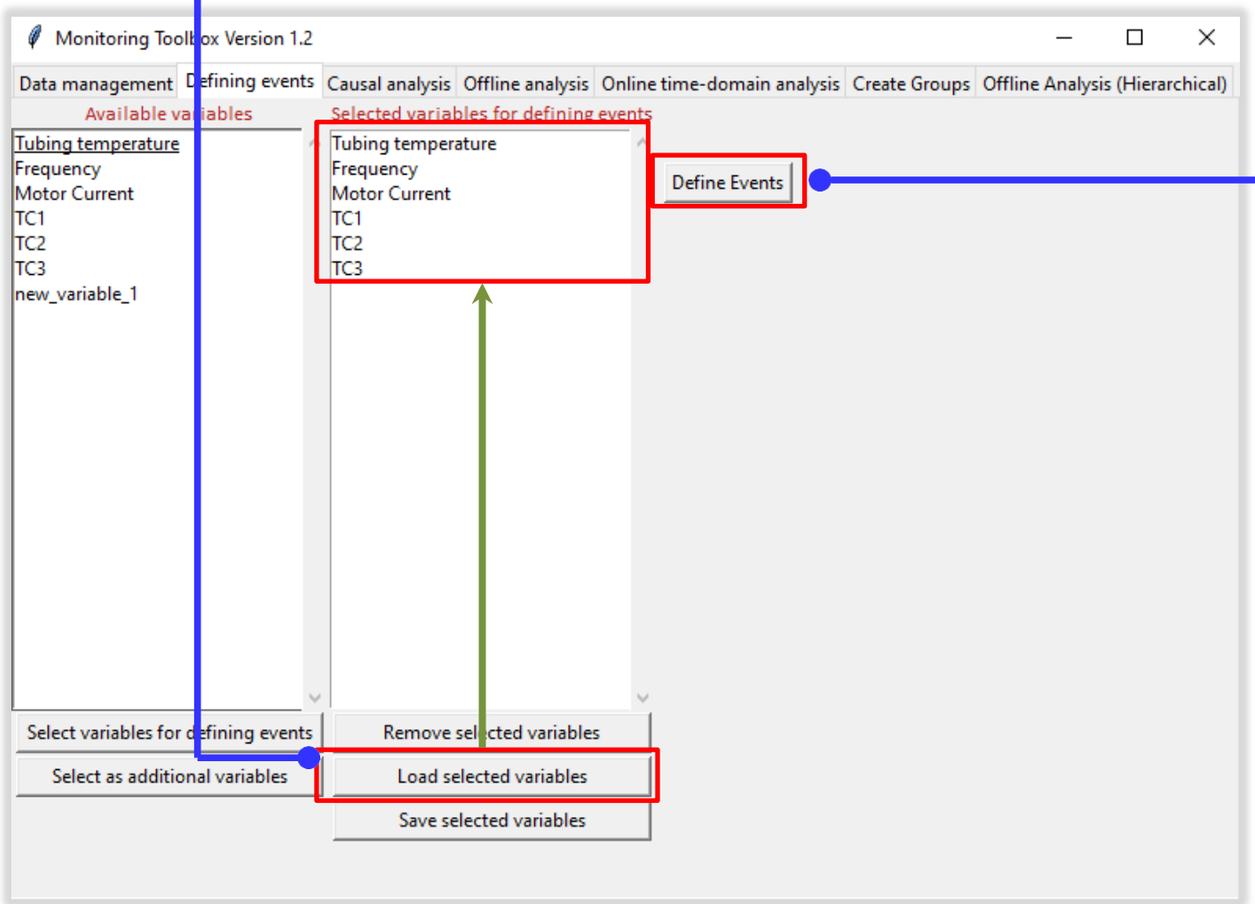
Predictive Monitoring Toolbox

Defining Events

Under the "**Defining events**" tab, different types of events that took place can be defined

1 **Load pre-selected variables for defining events**
- **Click "Load selected variables" button**

2 **Go to "Define Events" window**
- **Click "Define Events" button**



Predictive Monitoring Toolbox

Defining Events

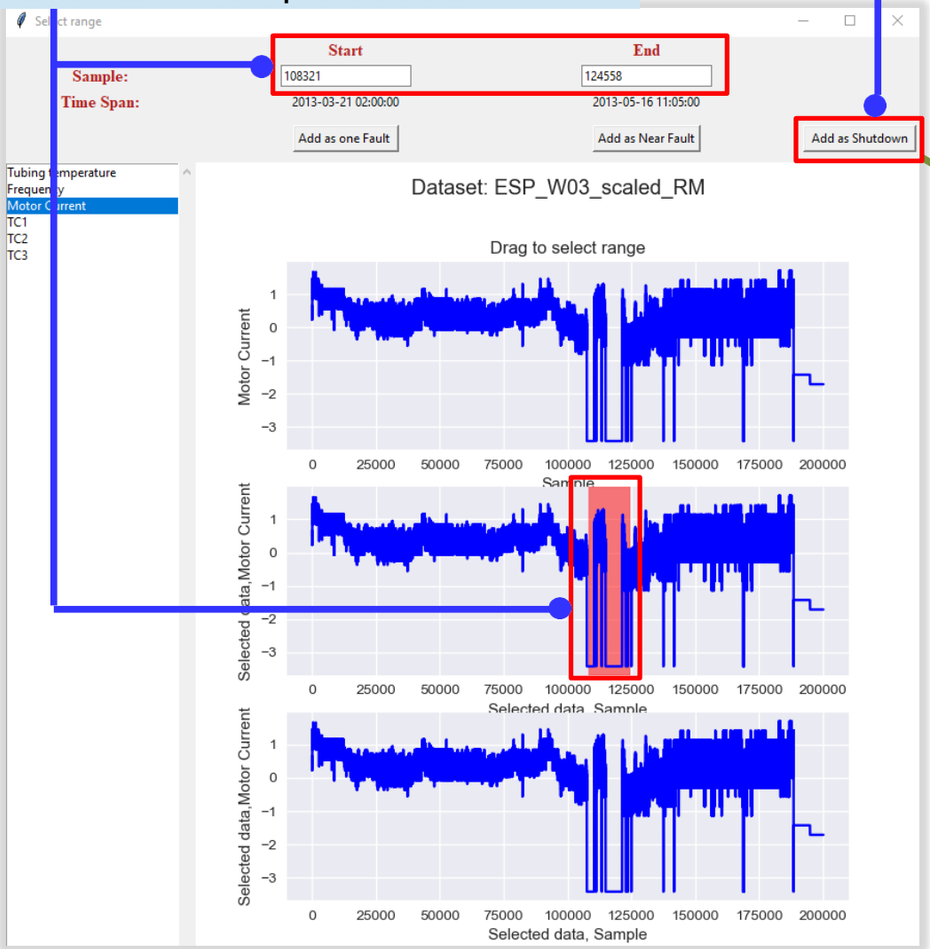
Under the "**Defining events**" tab, different types of events that took place can be defined

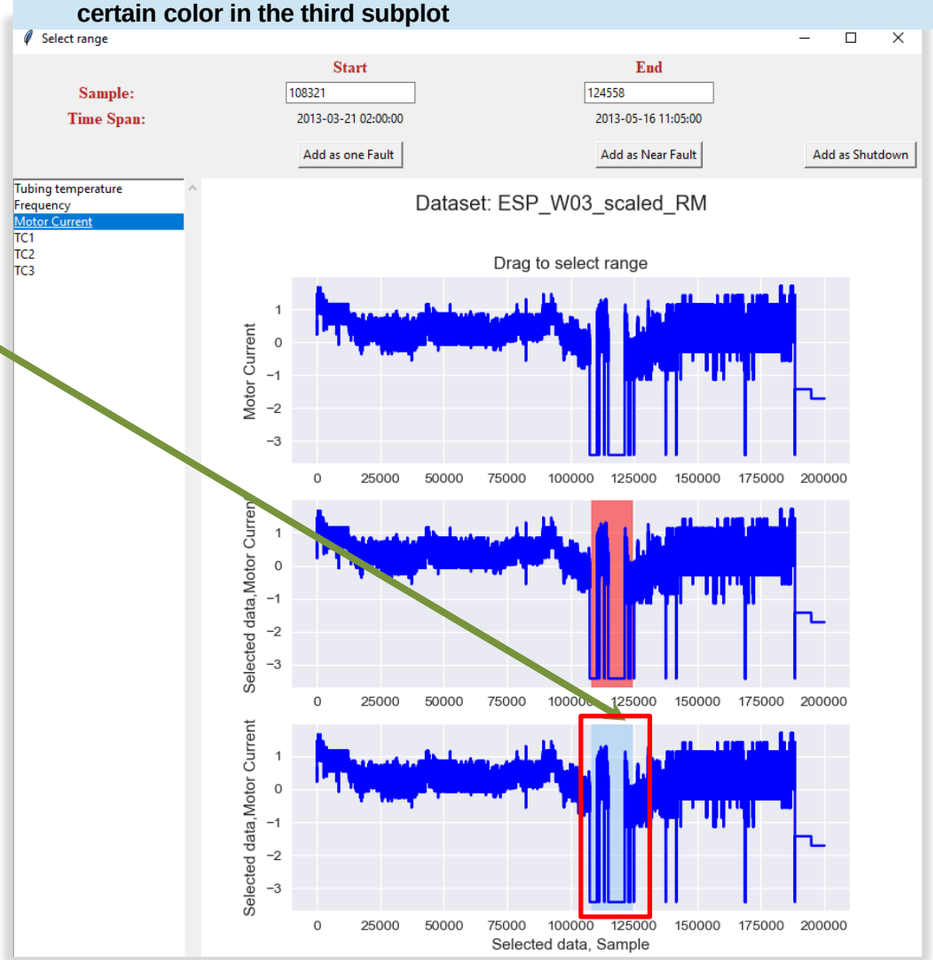
1 **Select a time window corresponding to an event**

- **Drag and select on the second subplot or specify the "Start" and the "End" on the top of the window**

2 **Add an event**

- **Click any of the three buttons "Add as one fault" or "Add as a near fault" or "Add as shutdown" to add an event; the event will be highlighted in a certain color in the third subplot**





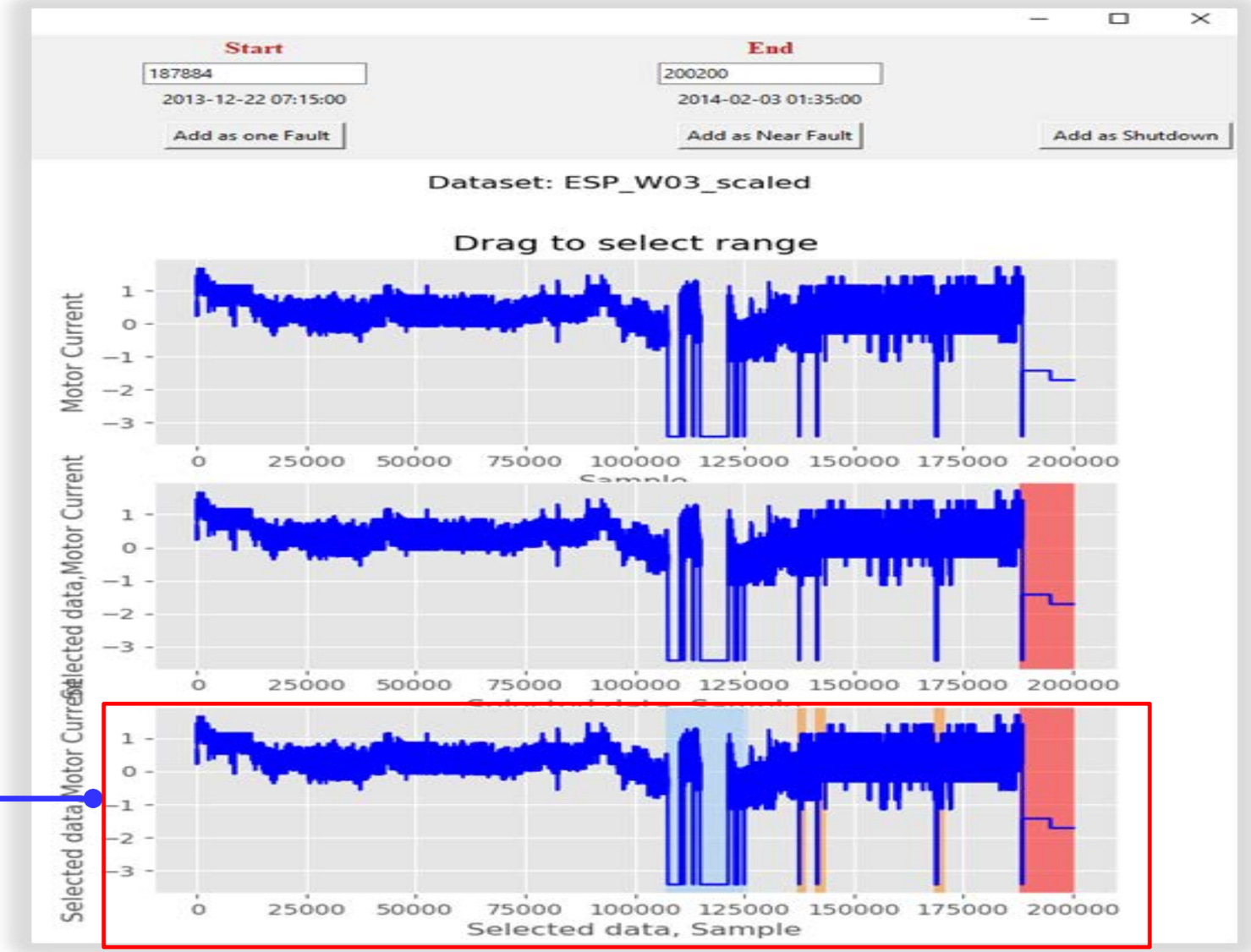
Predictive Monitoring Toolbox

Defining Events

Three types of events have been added

Select a time window corresponding to an event

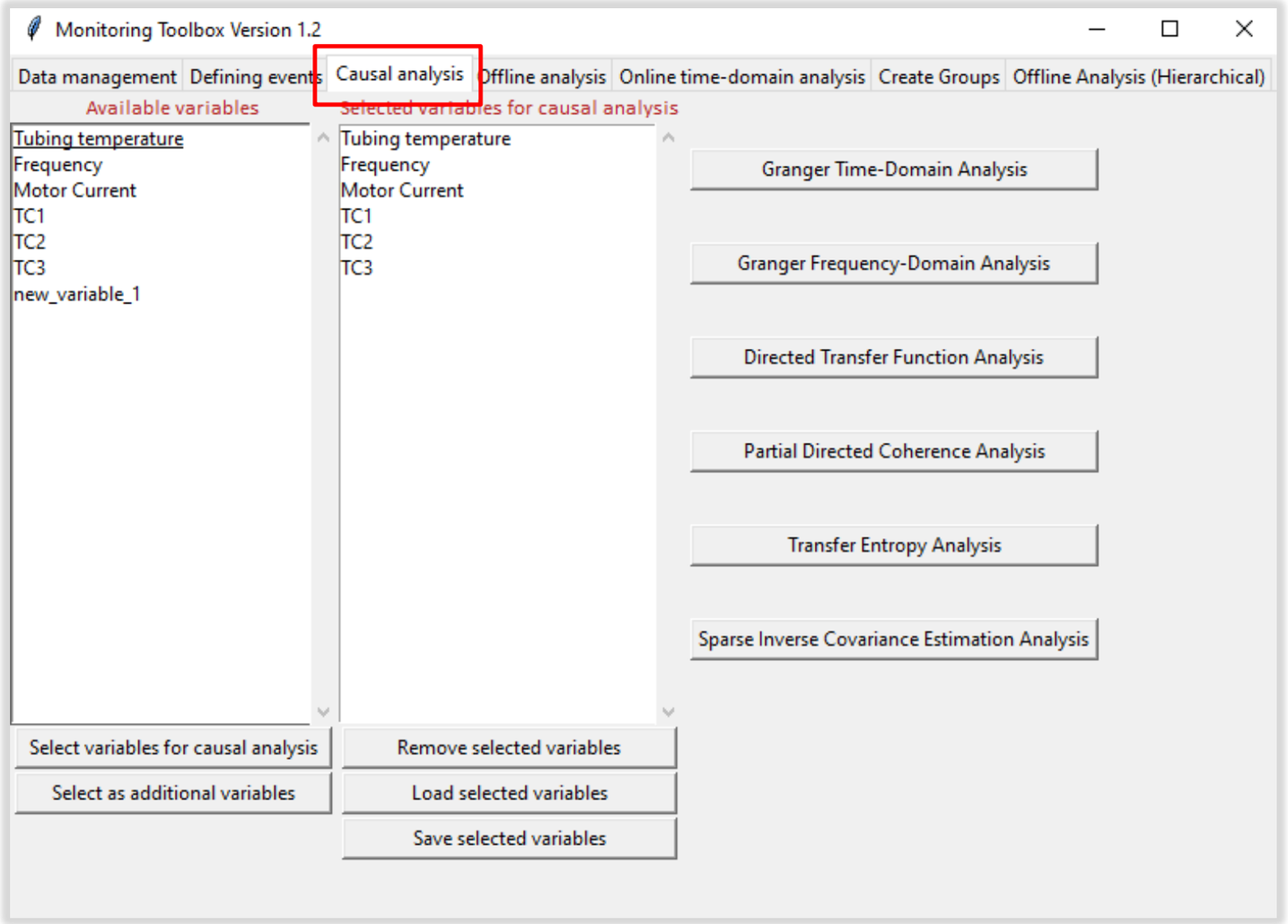
- **Shutdown is highlighted in blue; near faults are highlighted in yellow; faults are highlighted in red in the third subplot**

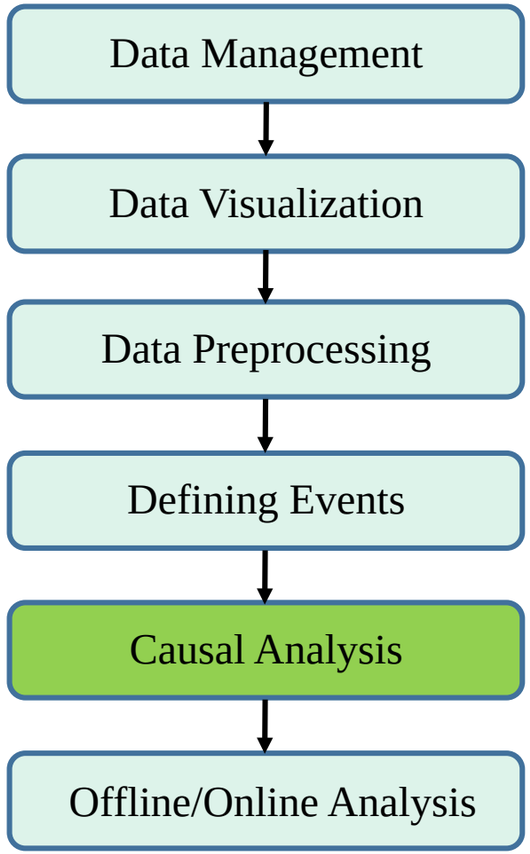


Predictive Monitoring Toolbox

Causal analysis

Under the "**Causal analysis**" tab, six algorithms are provided and can be used to investigate cause-effect relationships among variables

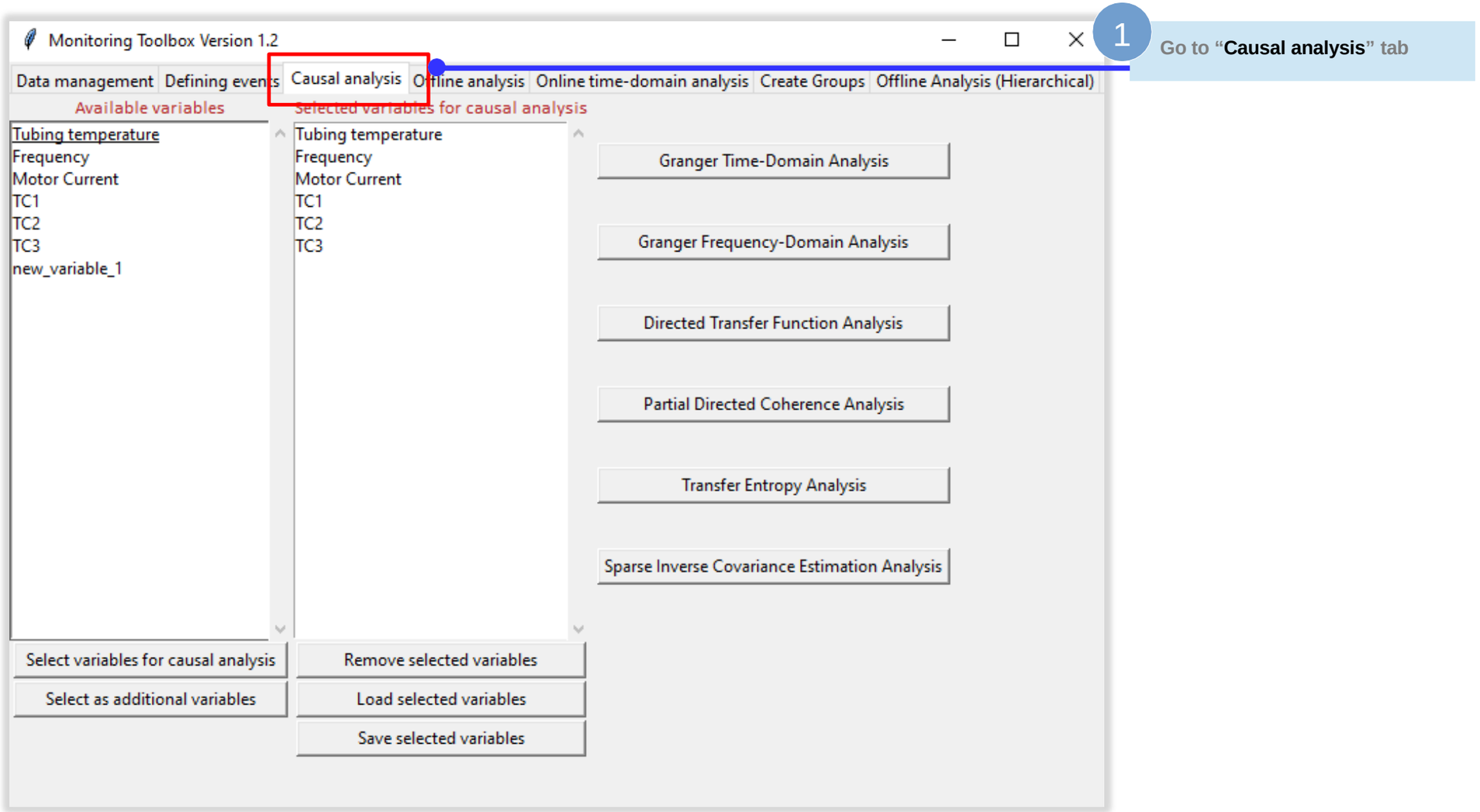




Predictive Monitoring Toolbox

Causal analysis

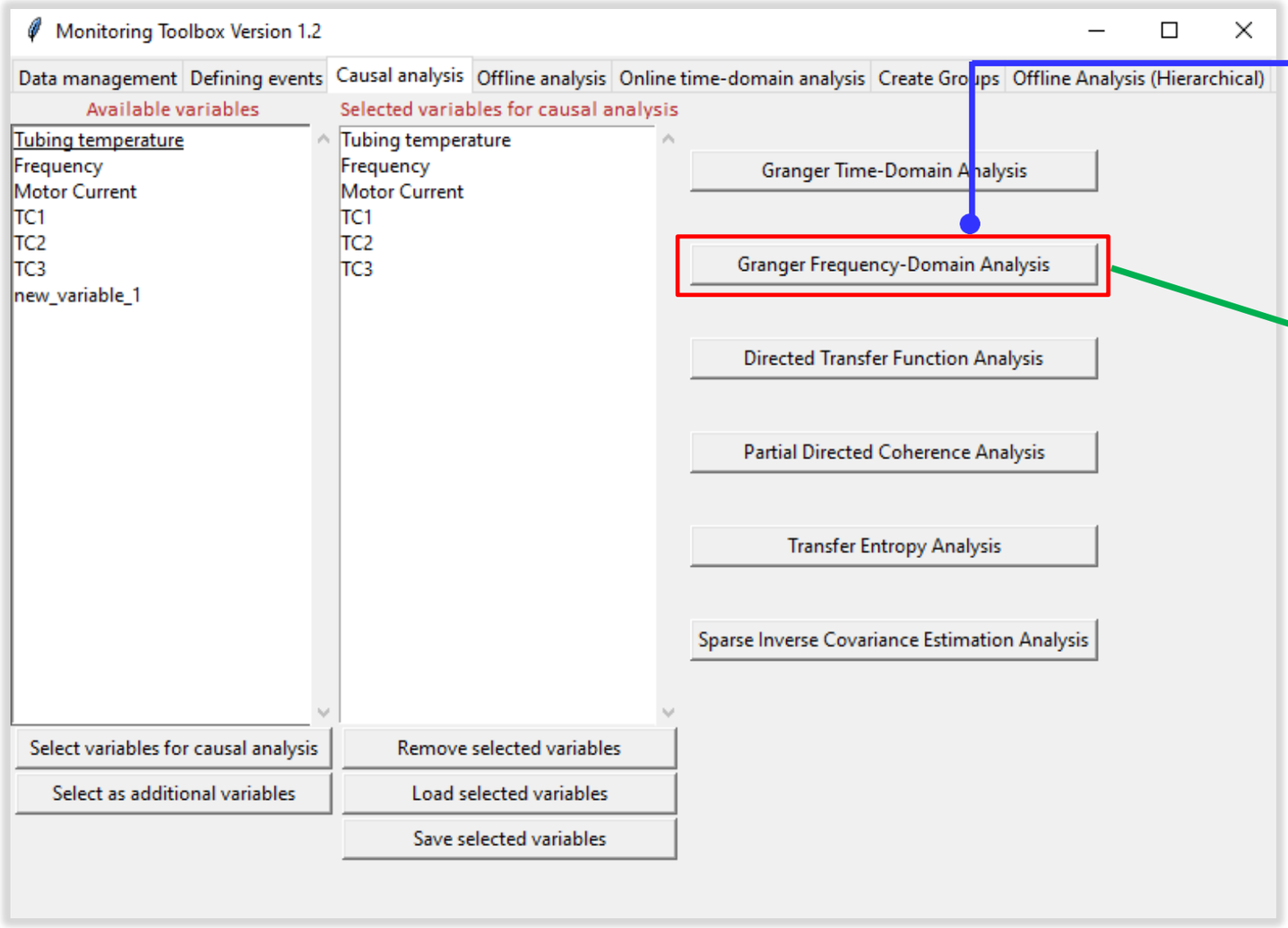
Under the "**Causal analysis**" tab, six algorithms are provided and can be used to investigate cause-effect relationships among variables



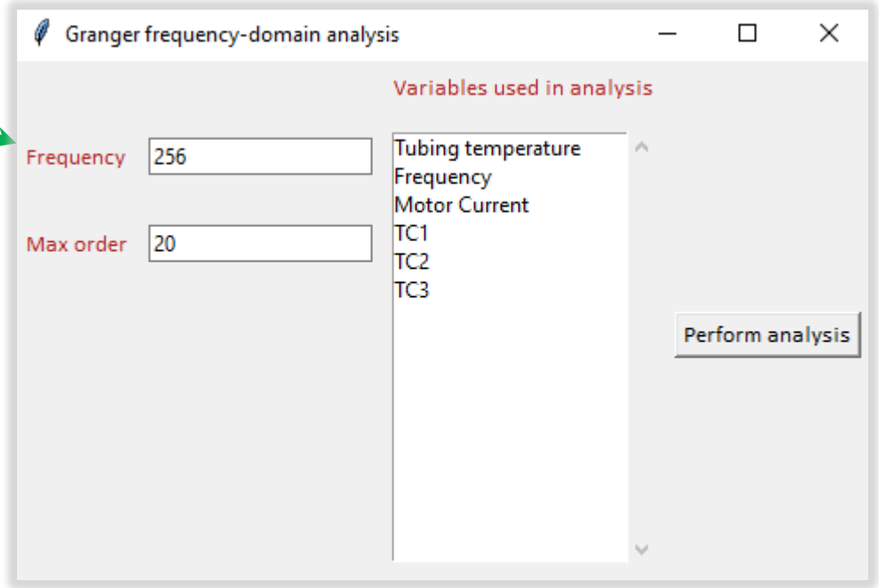
Predictive Monitoring Toolbox

Causal analysis

Under the "**Causal analysis**" tab, six algorithms are provided and can be used to investigate cause-effect relationships among variables



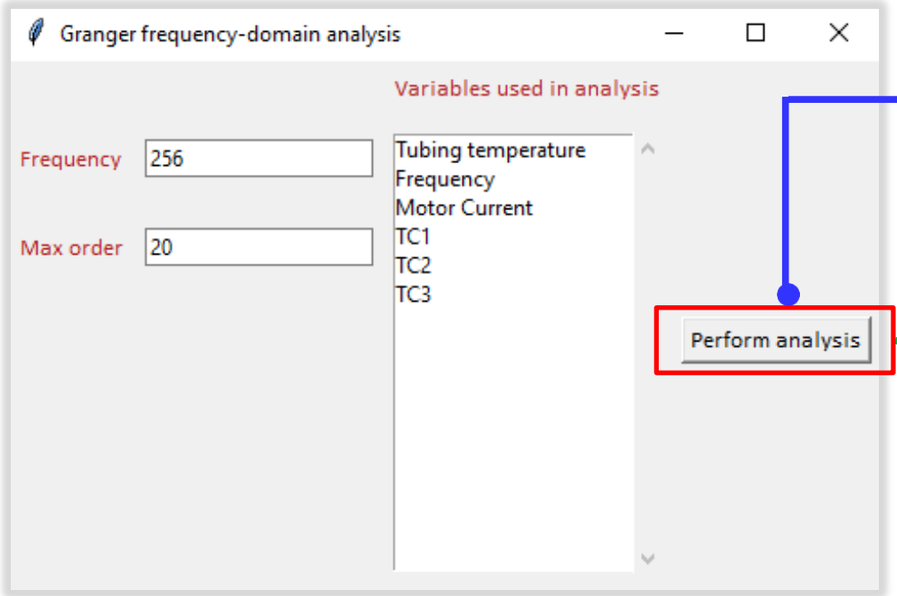
2 Click the "Granger **Frequency-Domain Analysis**" button, and a window will pop-up



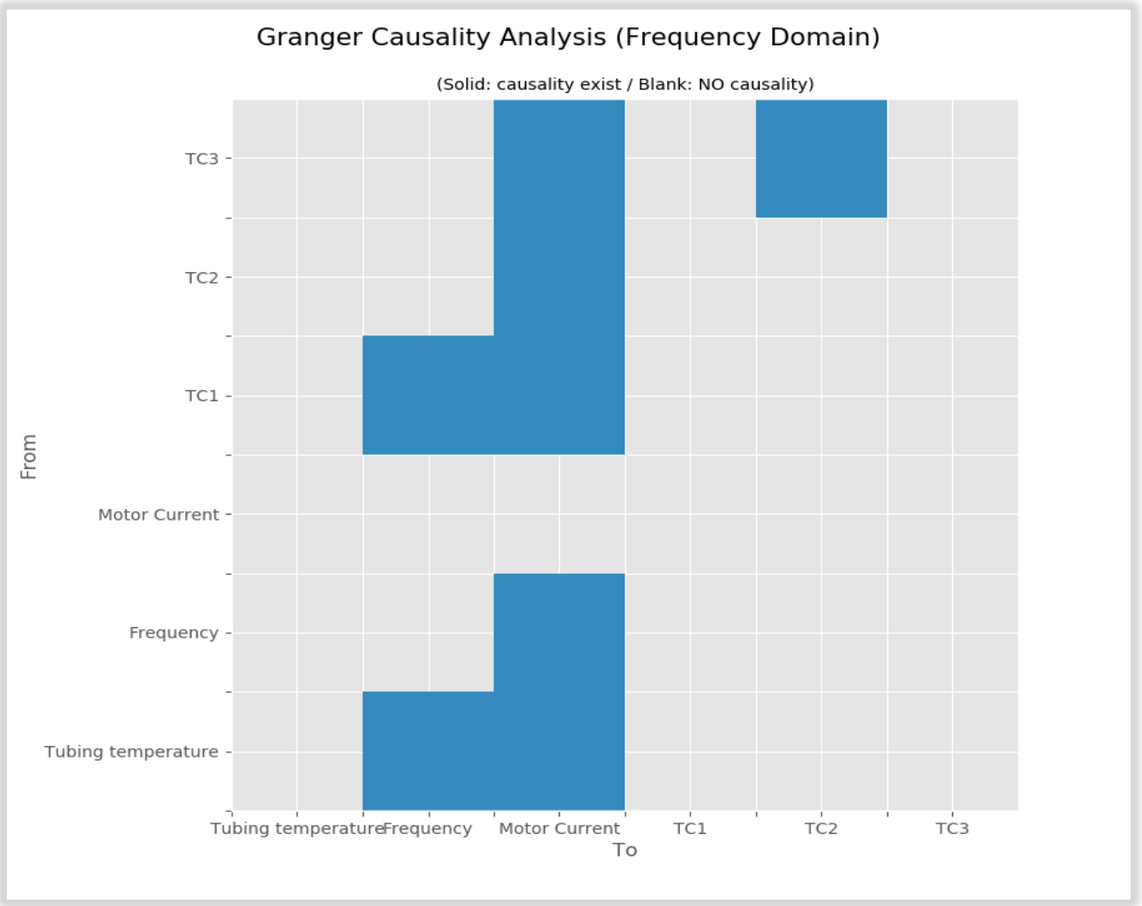
Predictive Monitoring Toolbox

Causal analysis

Under the "**Causal analysis**" tab, six algorithms are provided and can be used to investigate cause-effect relationships among variables

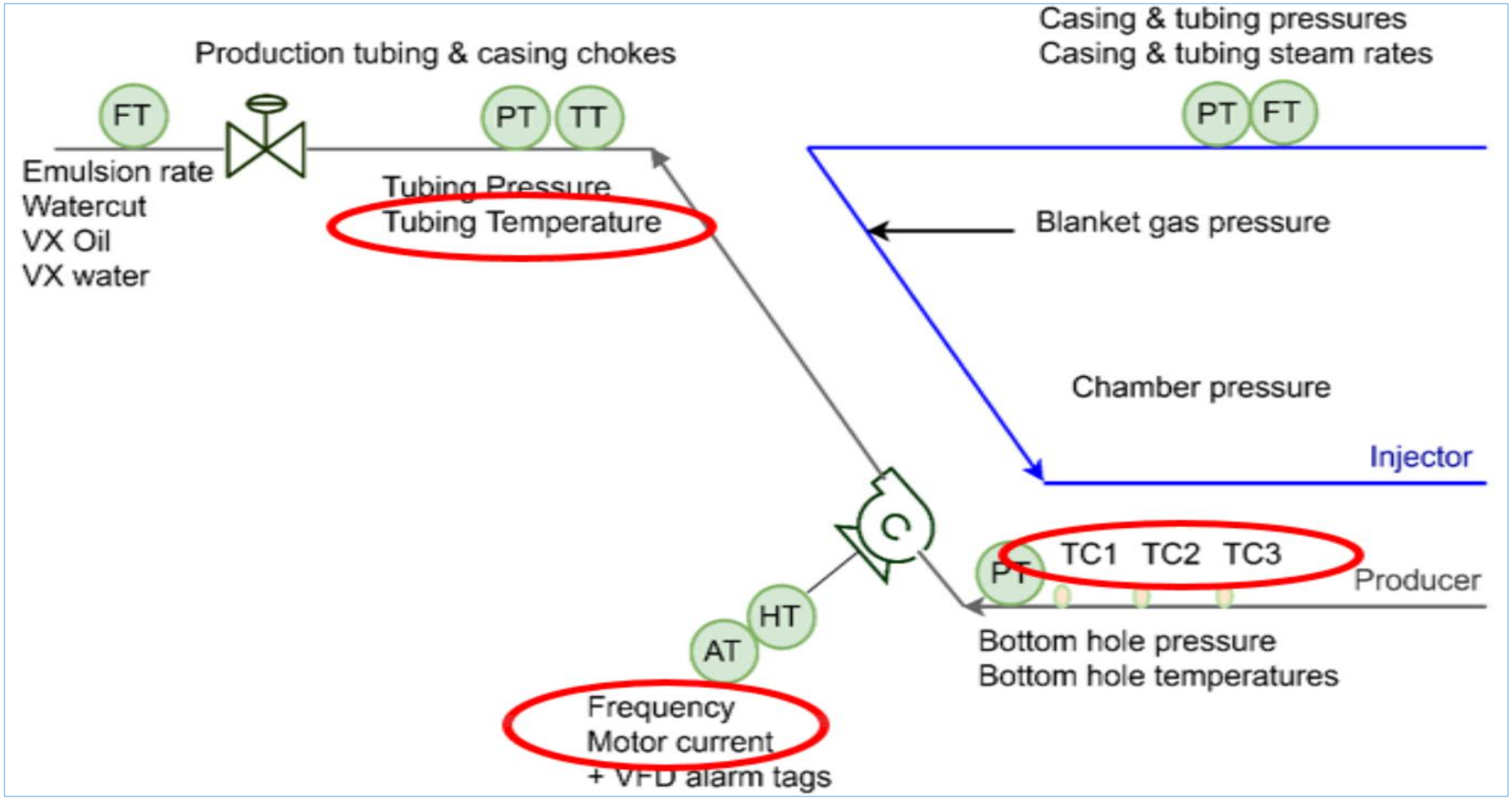


3 Click the "Granger **Frequency-Domain Analysis**" button, the results will be shown



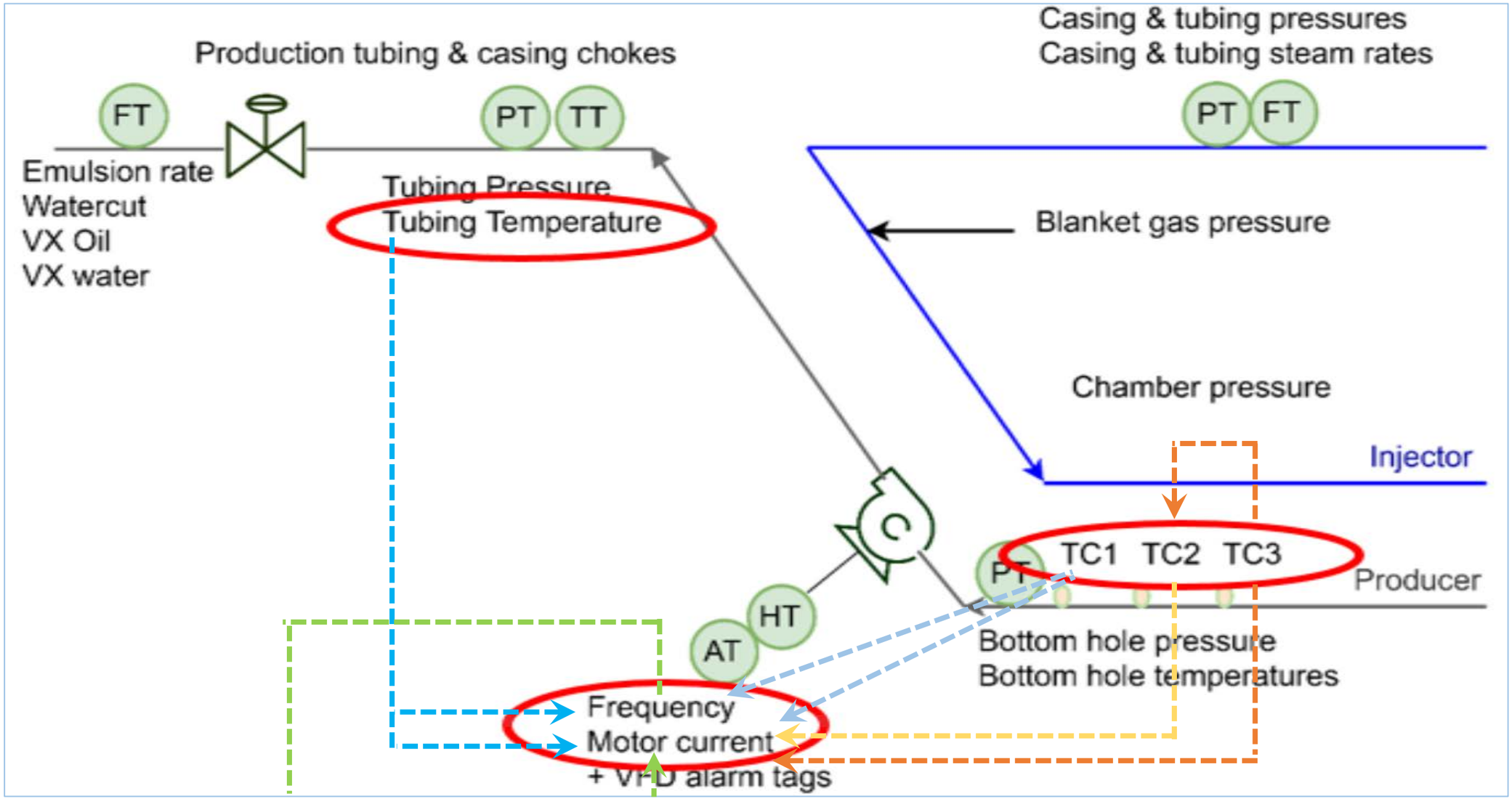
Predictive Monitoring Toolbox

Causal analysis – interpreting the results: **Each dashed line indicates a cause-effect relationship**



Predictive Monitoring Toolbox

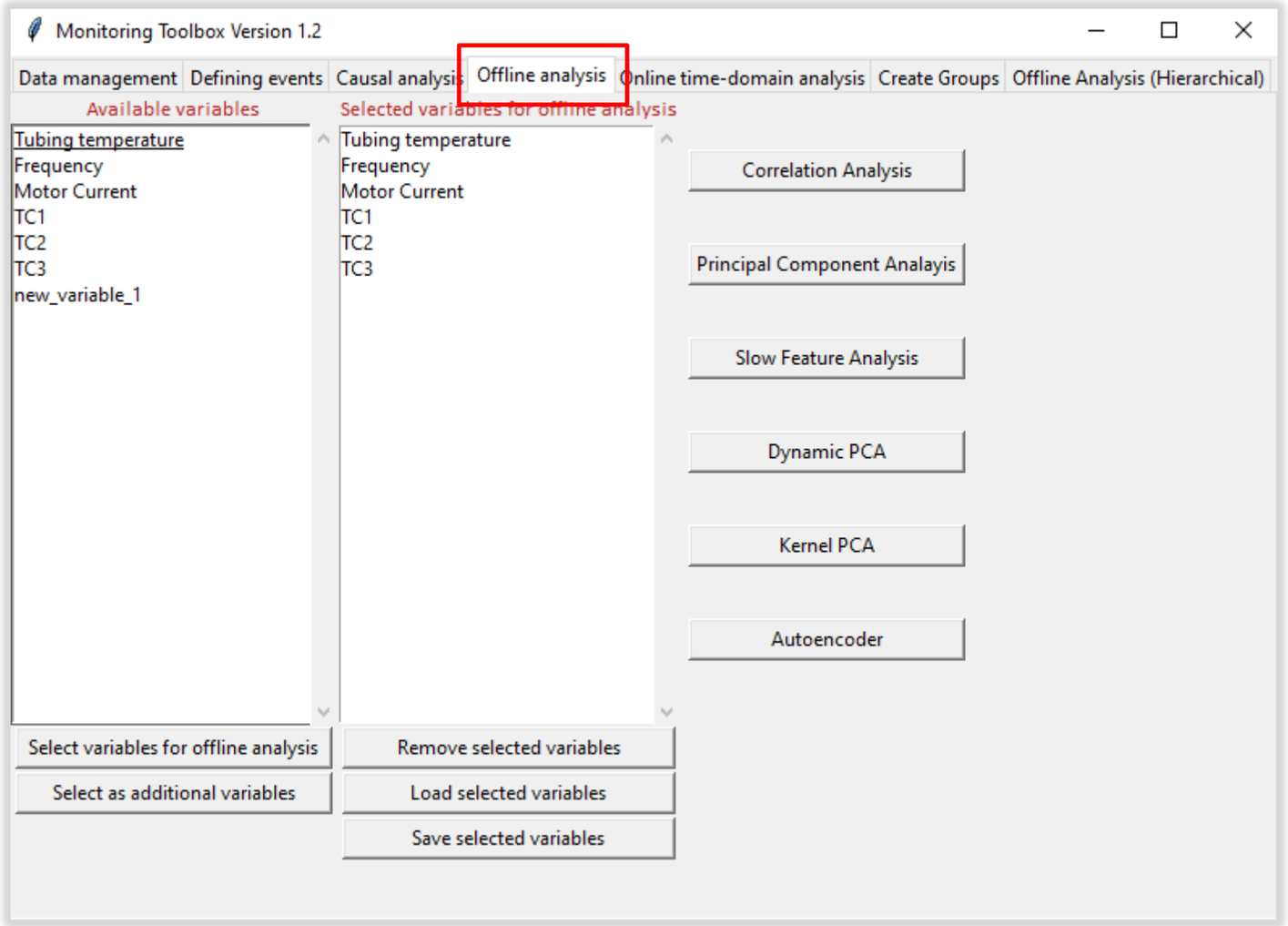
Causal analysis – interpreting the results: **Each dashed line indicates a cause-effect relationship**

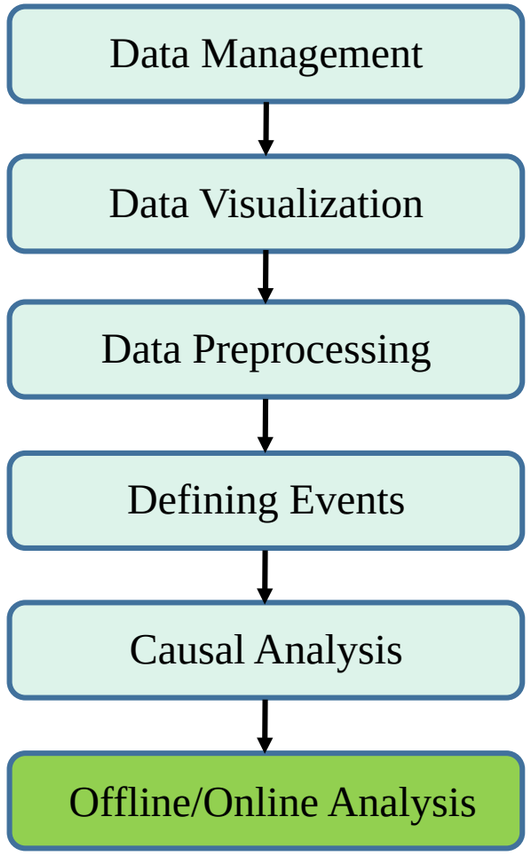


Predictive Monitoring Toolbox

Offline monitoring

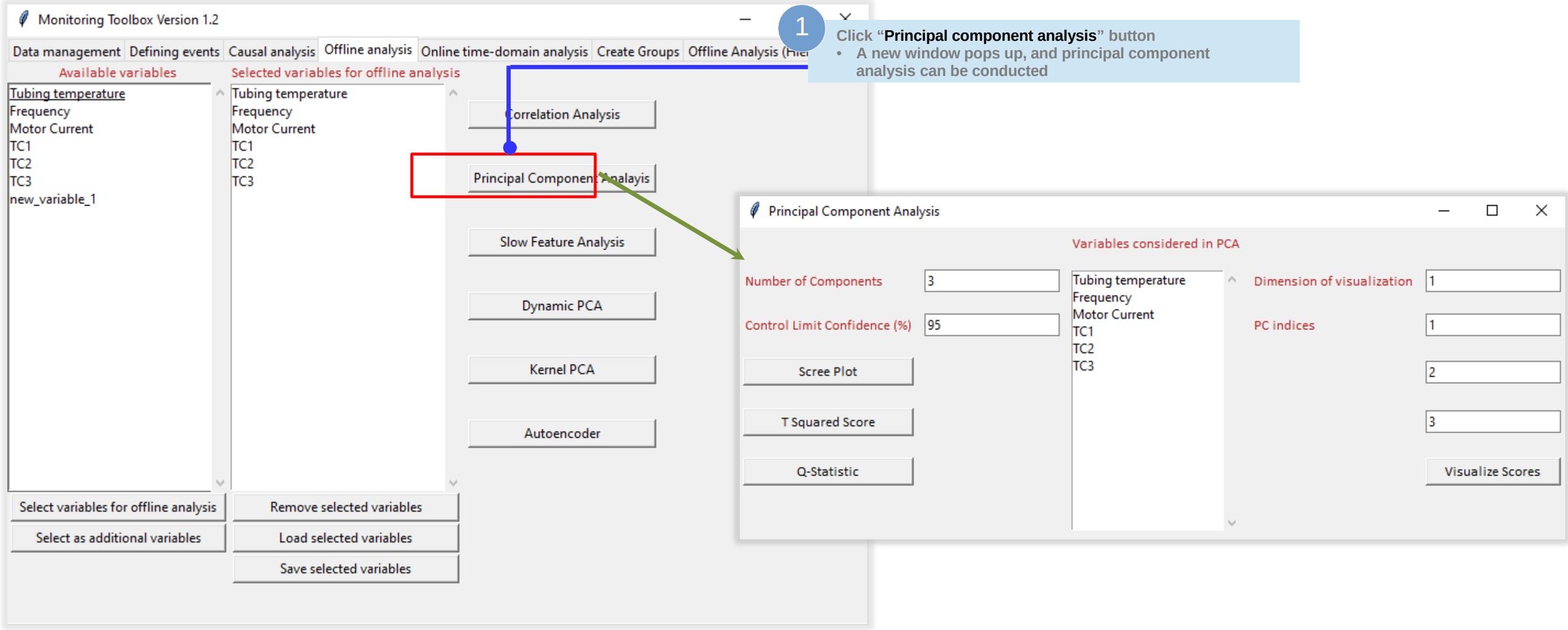
Under the "**Offline analysis**" tab, five algorithms are provided and can be used to monitor the process and detect events





Predictive Monitoring Toolbox

Offline monitoring – Principal component analysis

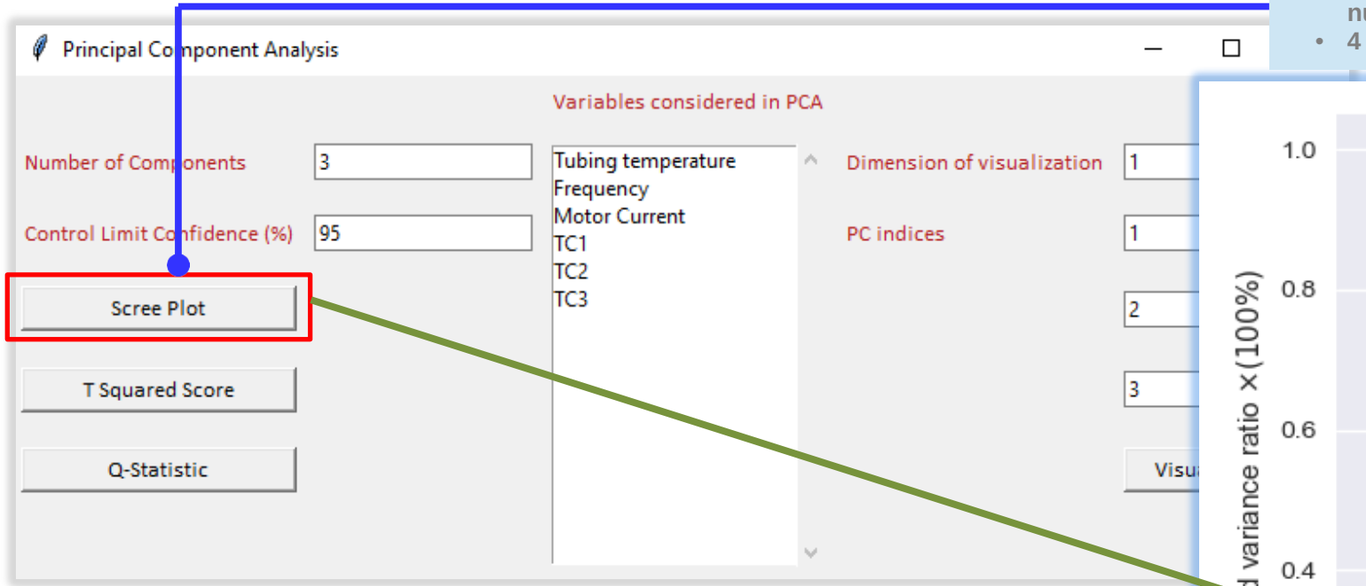


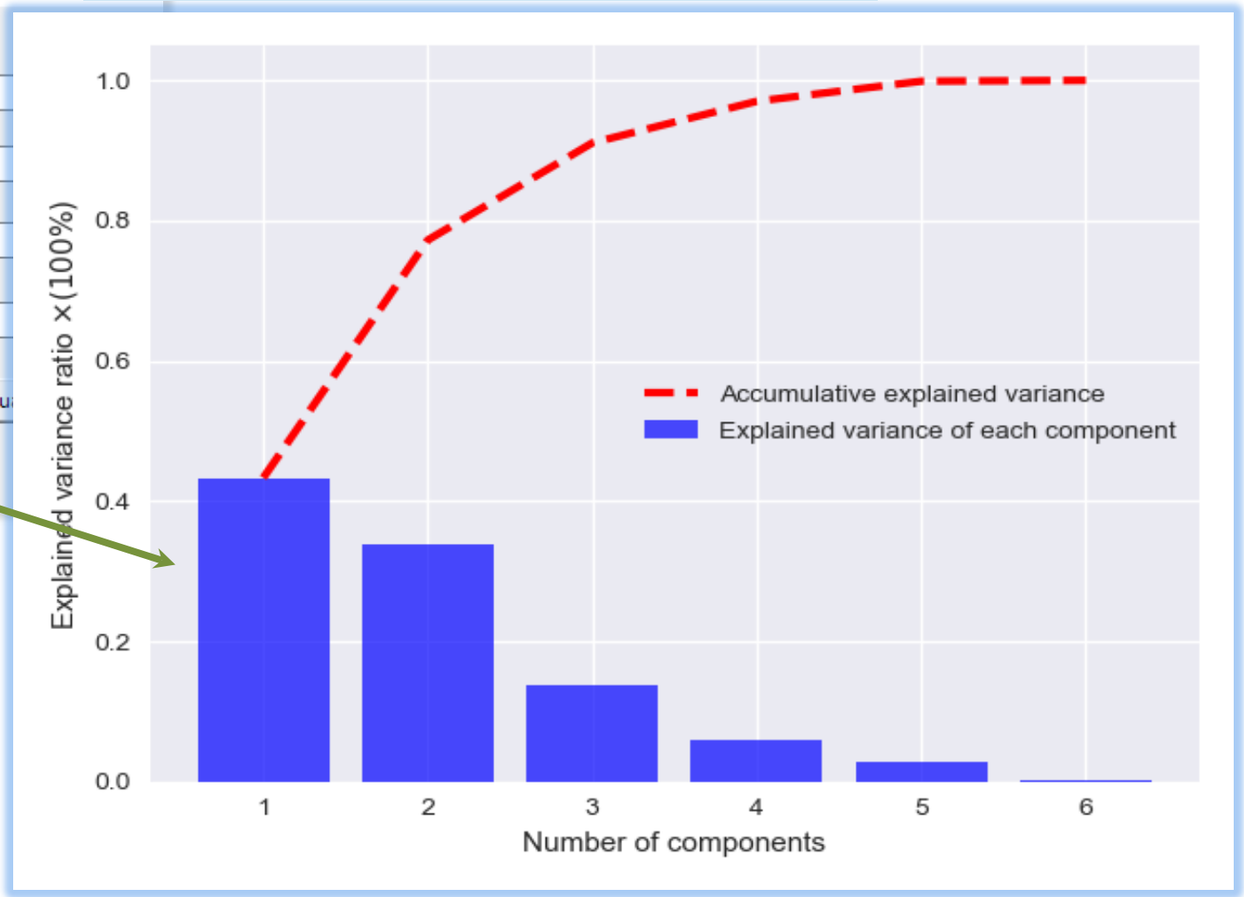
Predictive Monitoring Toolbox

Offline monitoring – Principal component analysis (PCA)

2

Click "**Scree plot**" button
- A figure will be generated and can help to select a good number of principal components for PCA.
- 4 is a good number of components based on the plot below.

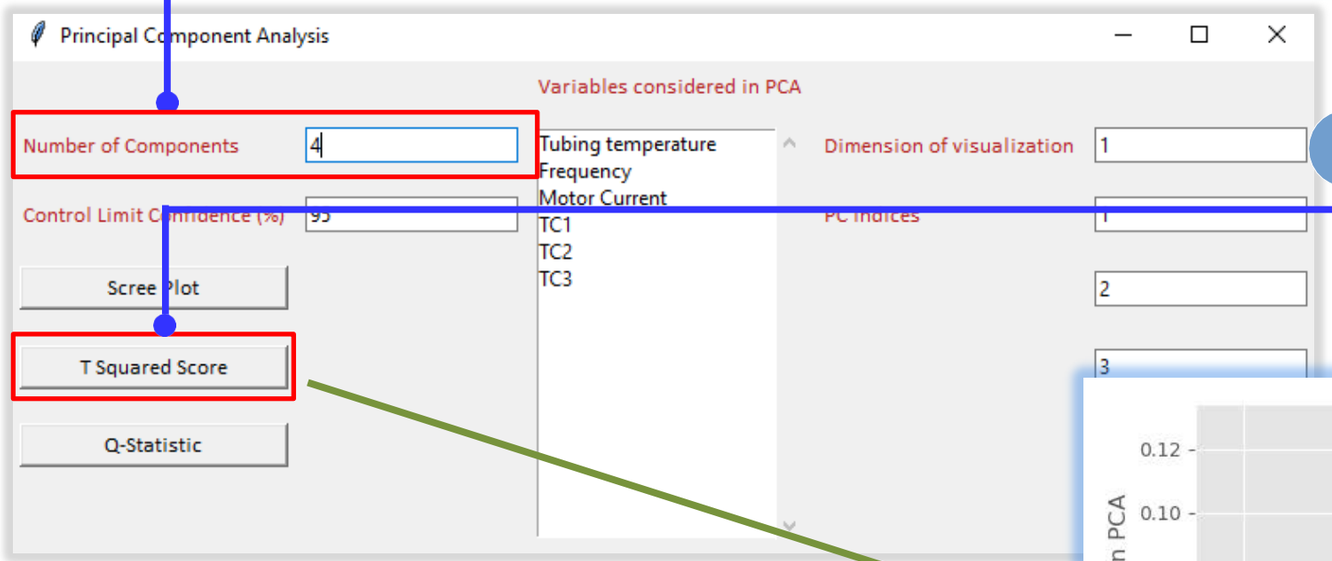




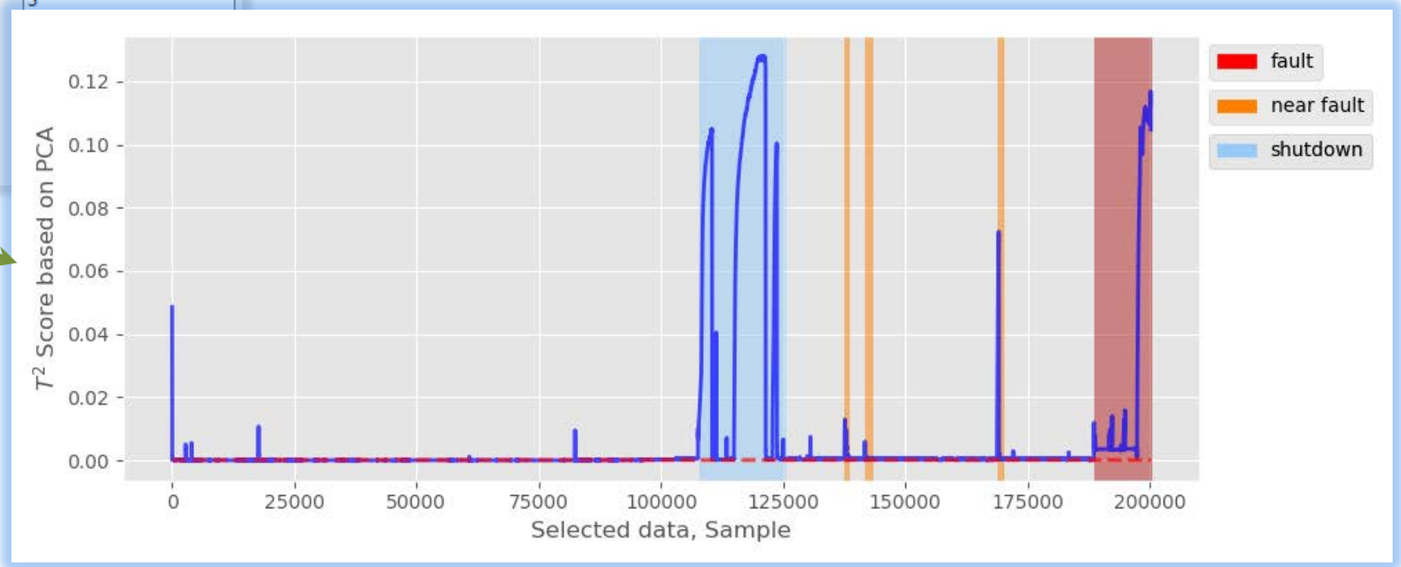
Predictive Monitoring Toolbox

Offline monitoring – Principal component analysis (PCA)

3 Enter "**4**" as the number of components.

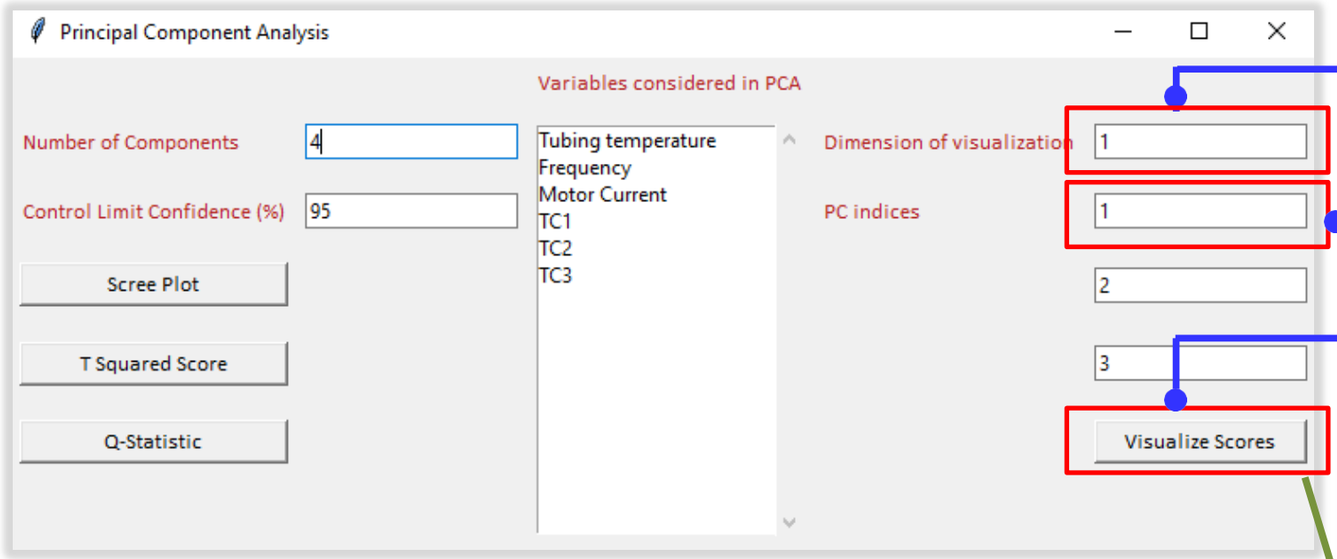


4 Click "**T Squared Score**" button
- PCA is conducted, and the trajectory of T-squared score is presented and compared with the threshold. If the score breaches the threshold, an event is detected.



Predictive Monitoring Toolbox

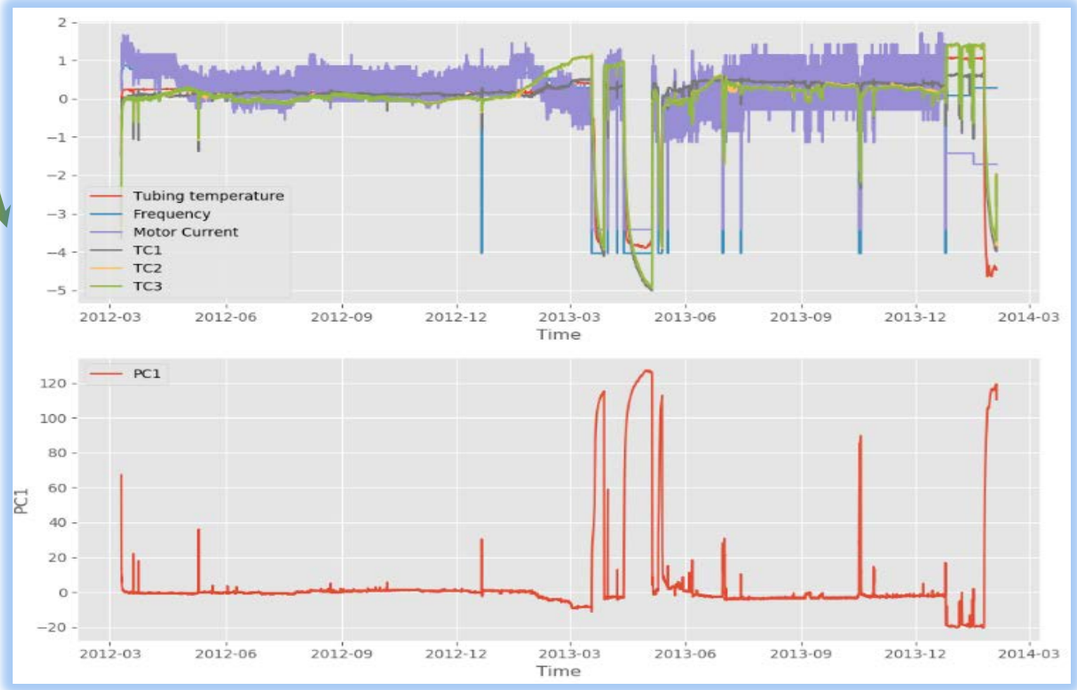
Offline monitoring using PCA – Visualizing principal components of PCA (1D)



5 Enter "**1**" as the dimension of visualization.

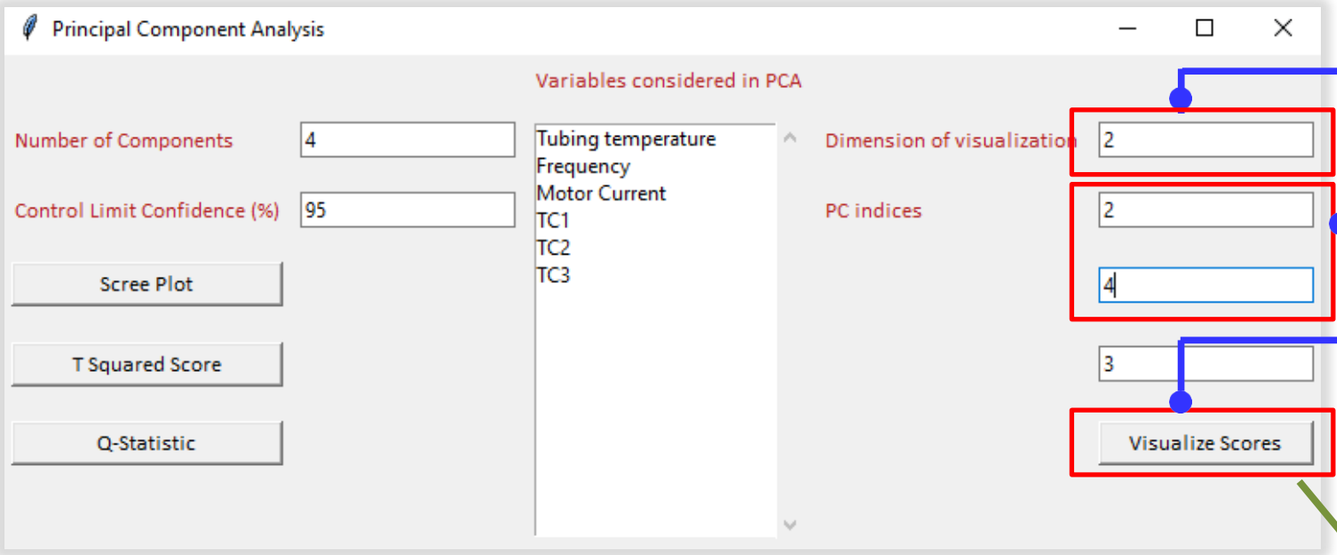
6 This means that the first component is to be visualized.

7 Clicking "**Visualize Scores**" shows the 1-dimensional visualization of the principal component.



Predictive Monitoring Toolbox

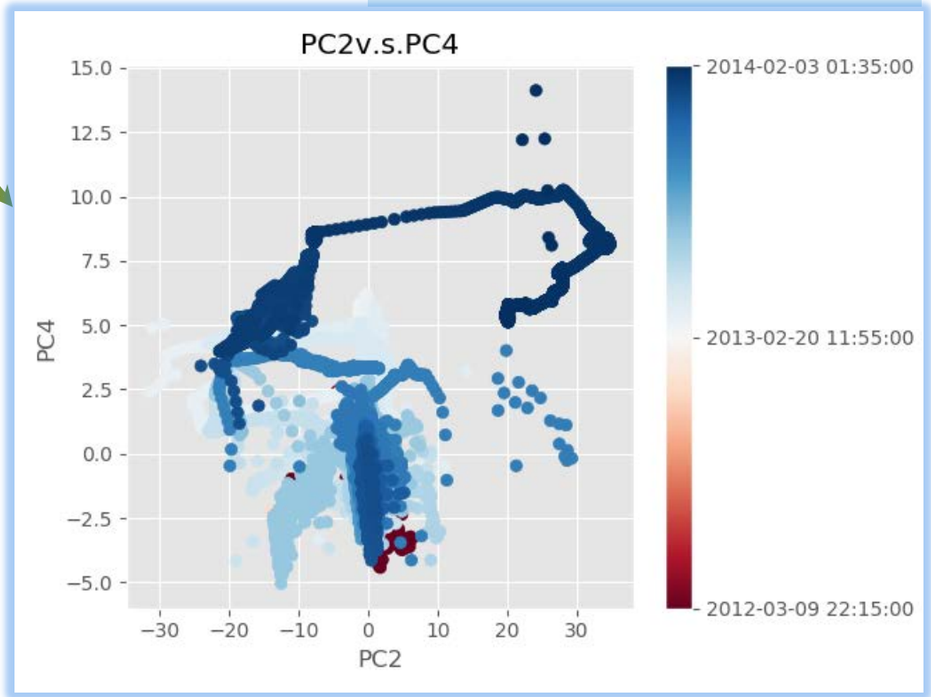
Offline monitoring using PCA – Visualizing principal components of PCA (2D)



8 Enter "**2**" as the dimension of visualization.

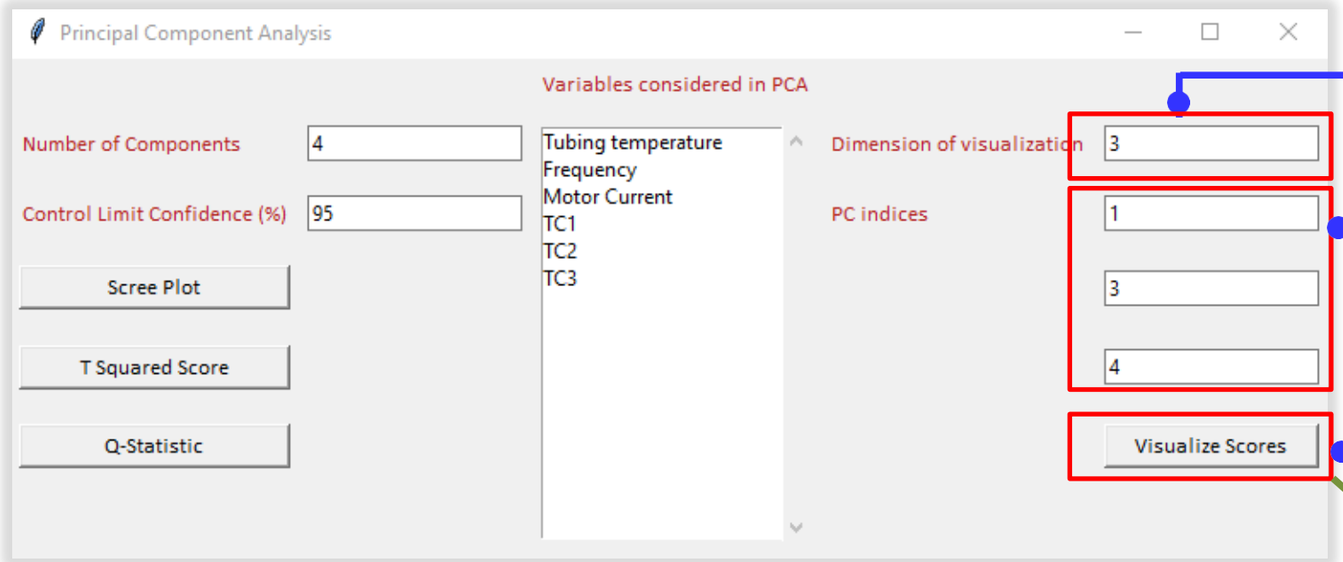
9 These two indices refer to the second and the fourth components.

10 Clicking "**Visualize Scores**" shows the 2-dimensional visualization of the principal component.



Predictive Monitoring Toolbox

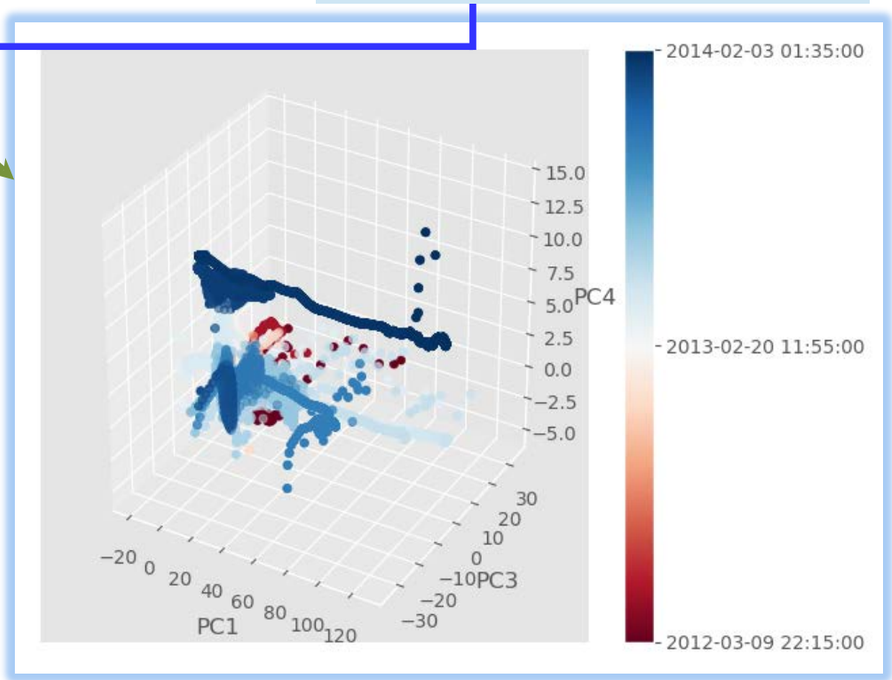
Offline monitoring using PCA – Visualizing principal components of PCA (3D)



11 Enter "**3**" as the dimension of visualization.

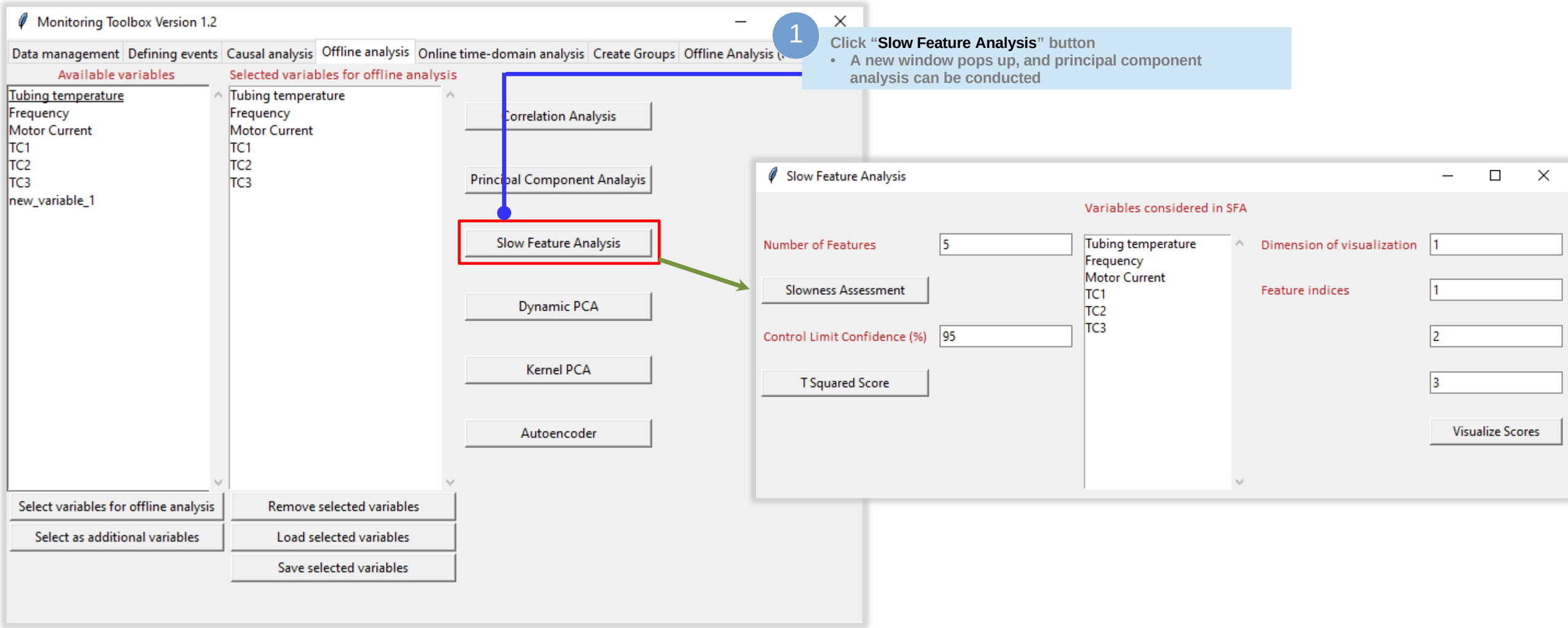
12 These three indices refer to the first, third and the fourth components.

13 Clicking "**Visualize Scores**" shows the 2-dimensional visualization of the principal component.



Predictive Monitoring Toolbox

Offline monitoring – Slow feature analysis

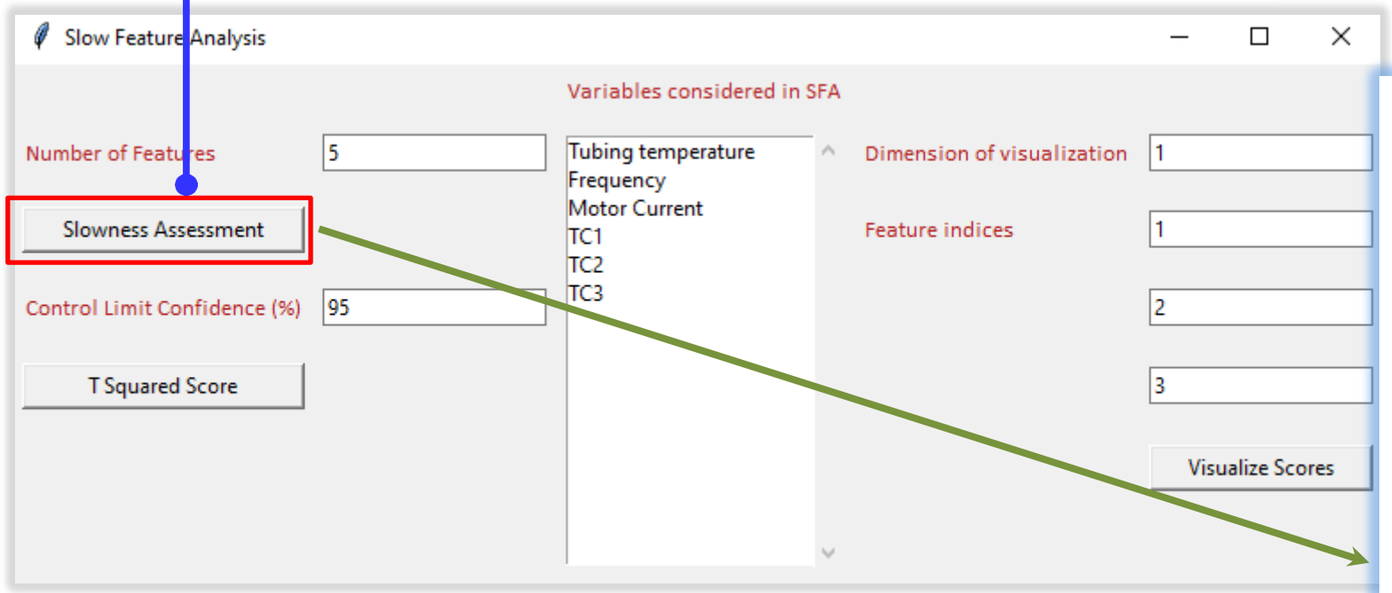


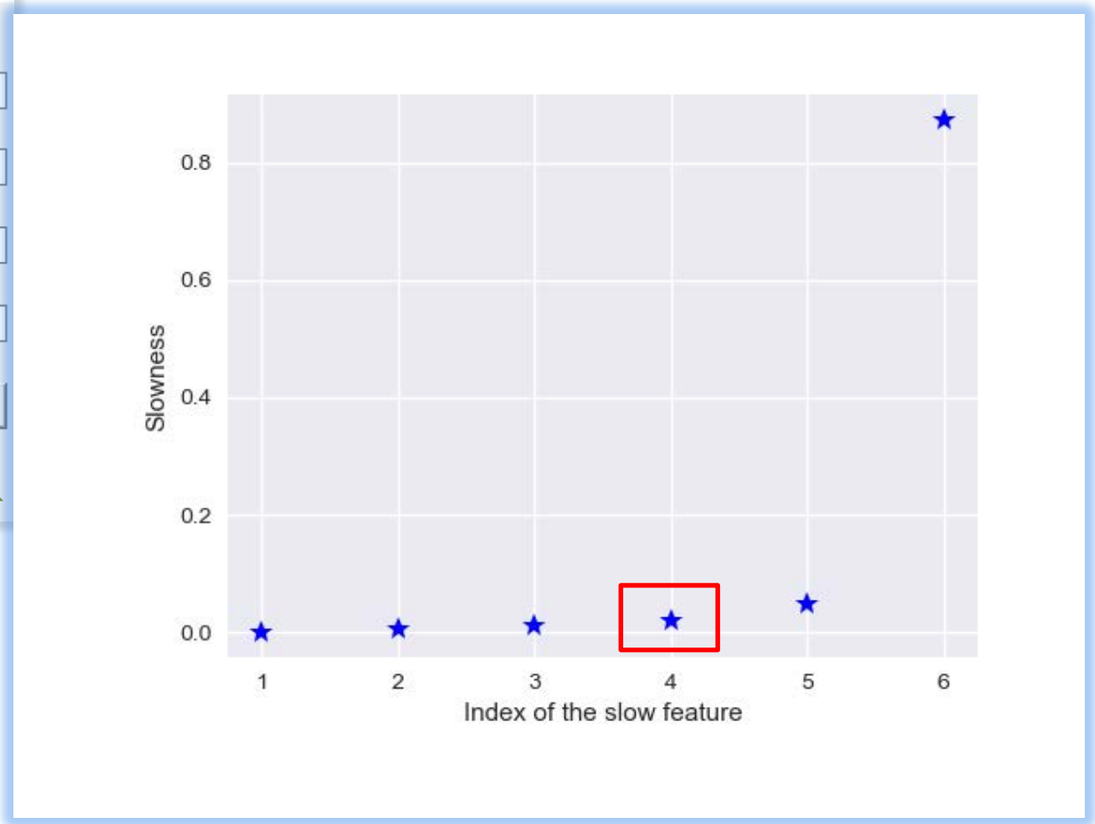
Predictive Monitoring Toolbox

Offline monitoring – Slow feature analysis (SFA)

2

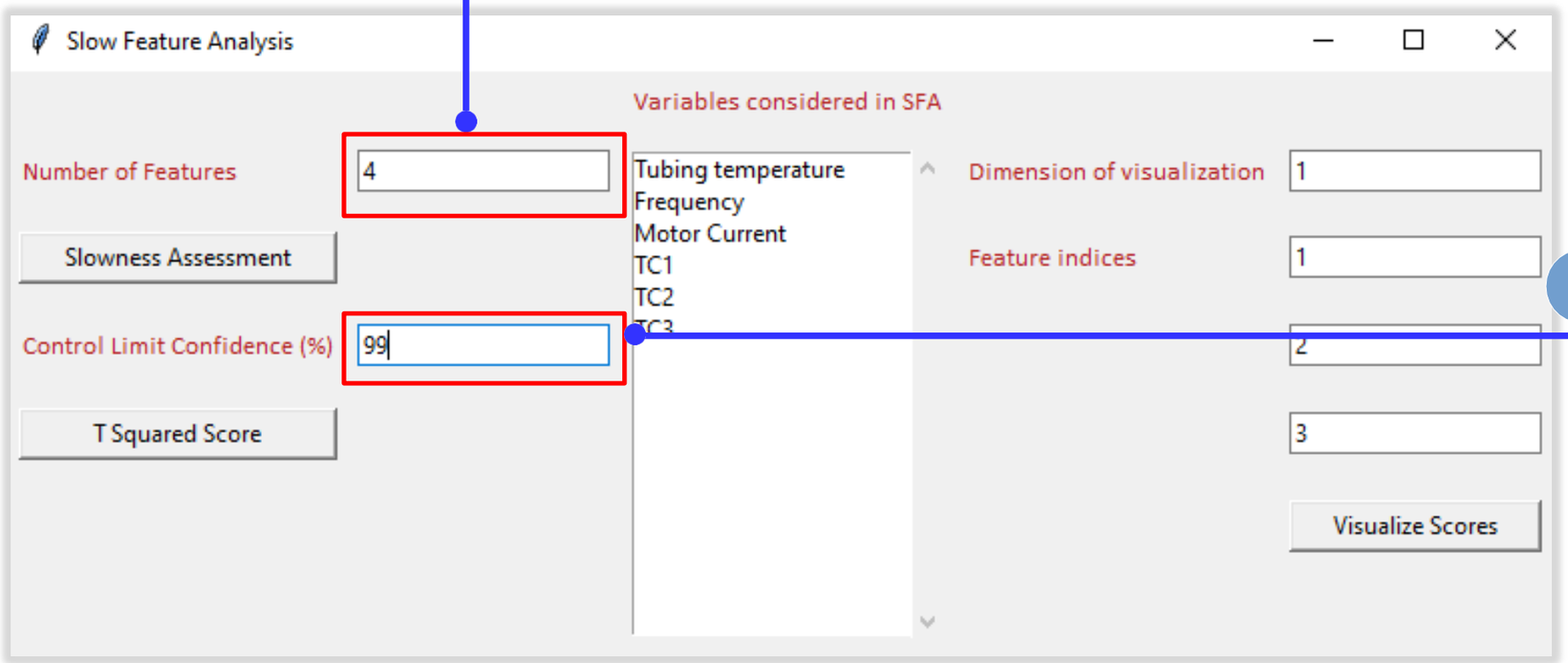
Click "**Slowness Assessment**" button
- A figure will be generated and can help to select a good number of slow features for SFA.
- 4 is a good number of components based on the plot below.





Predictive Monitoring Toolbox

Offline monitoring – Slow feature analysis (SFA)

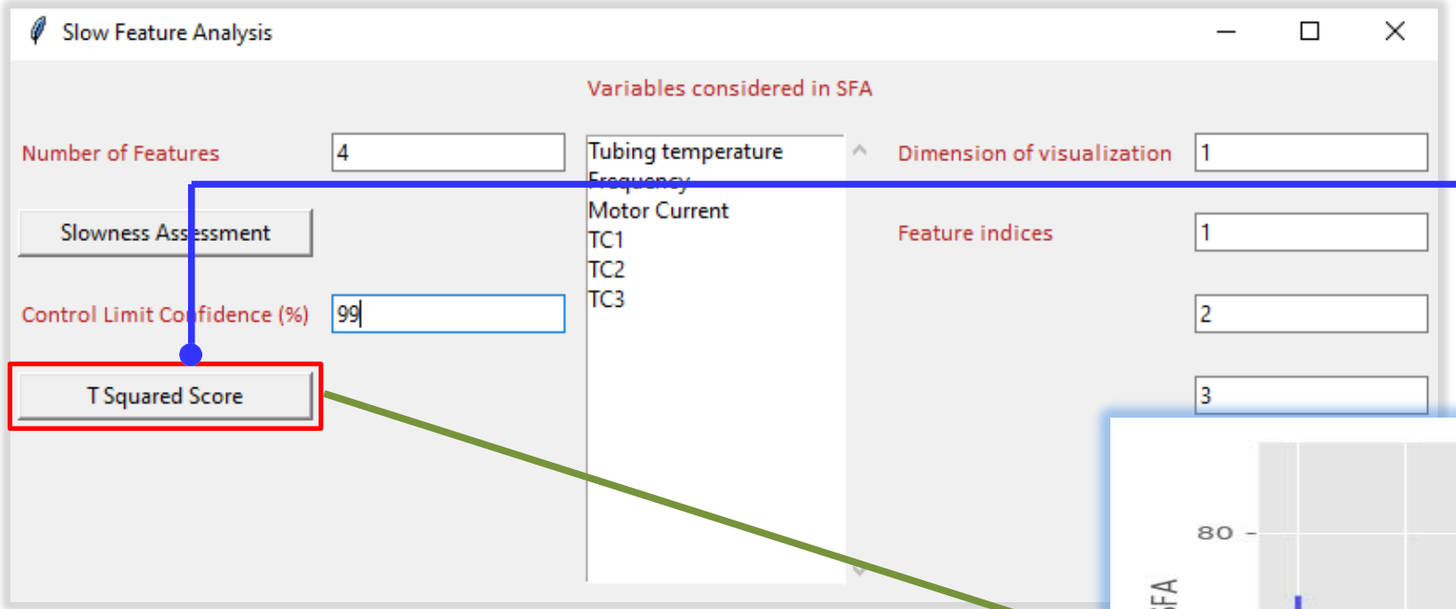


3 Enter "**4**" as the number of components.

4 Enter "**99**" so that the control limit confidence is made 99%.

Predictive Monitoring Toolbox

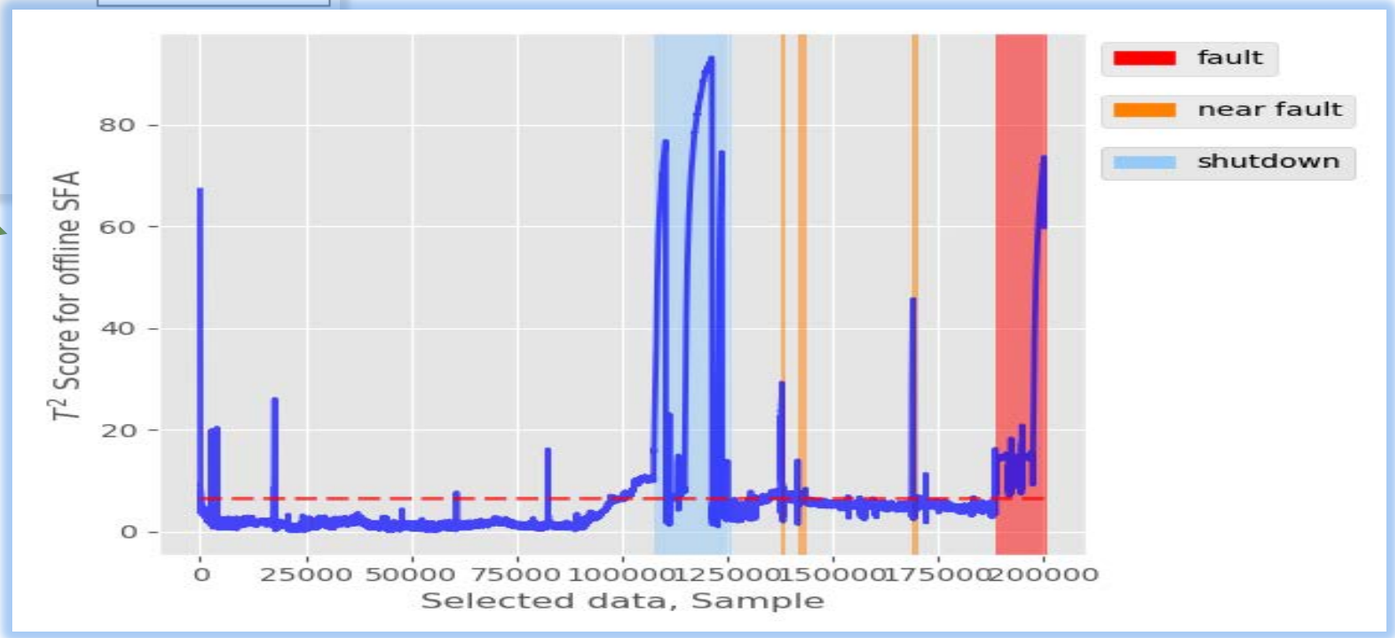
Offline monitoring – Slow feature analysis (SFA)



5

Click "**T Squared Score**" button

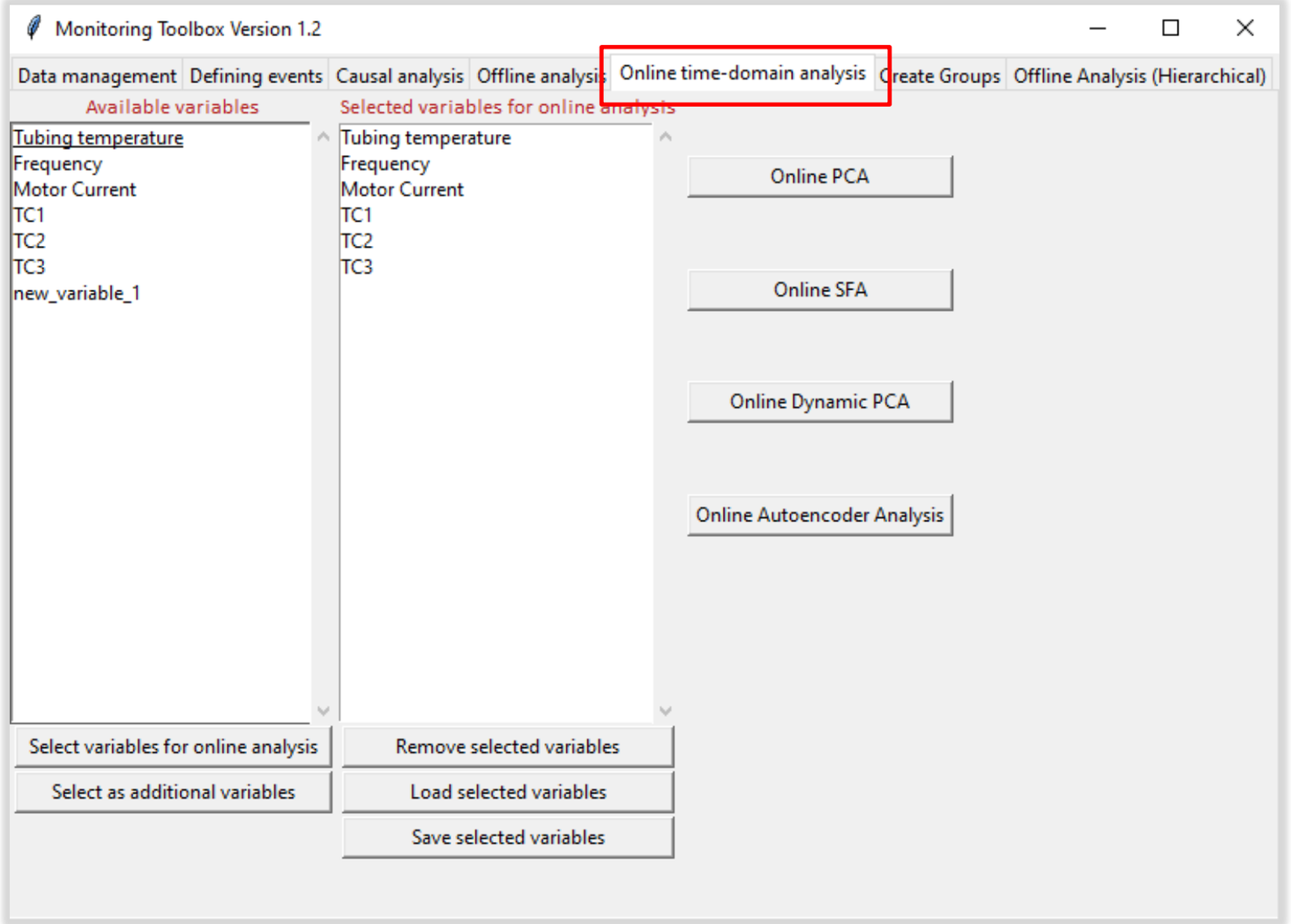
- SFA is conducted, and the trajectory of T-squared score is presented and compared with the threshold. If the score breaches the threshold, an event is detected.

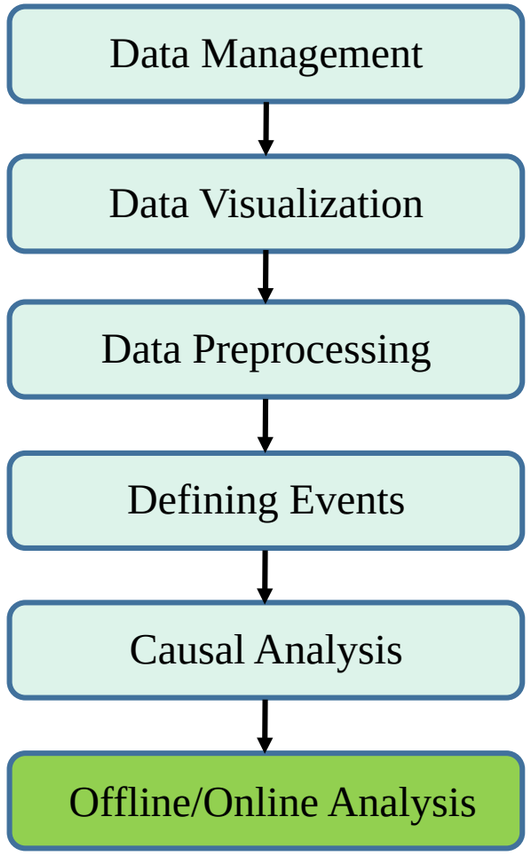


Predictive Monitoring Toolbox

Online monitoring (this simulates the implementation of online process monitoring)

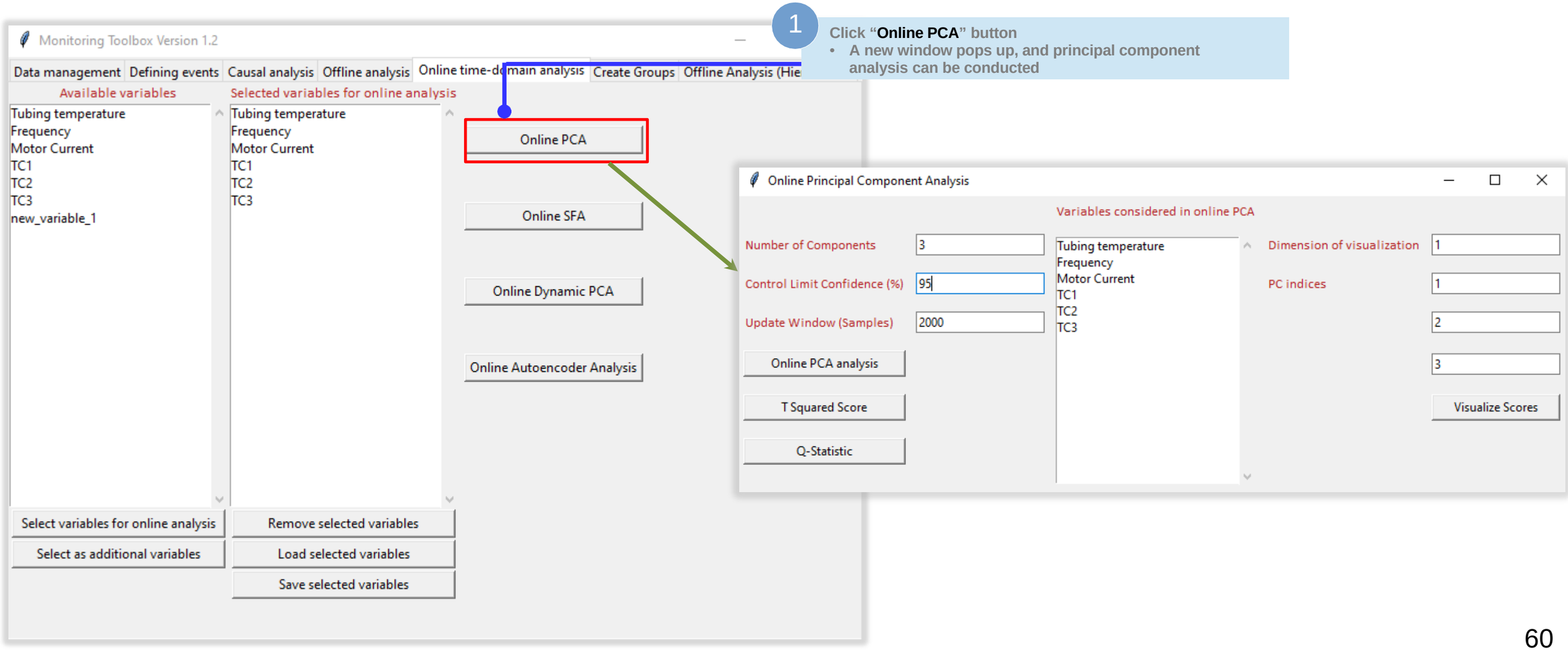
Under the "**Online analysis**" tab, four algorithms are provided and can be used to monitor the process and detect events





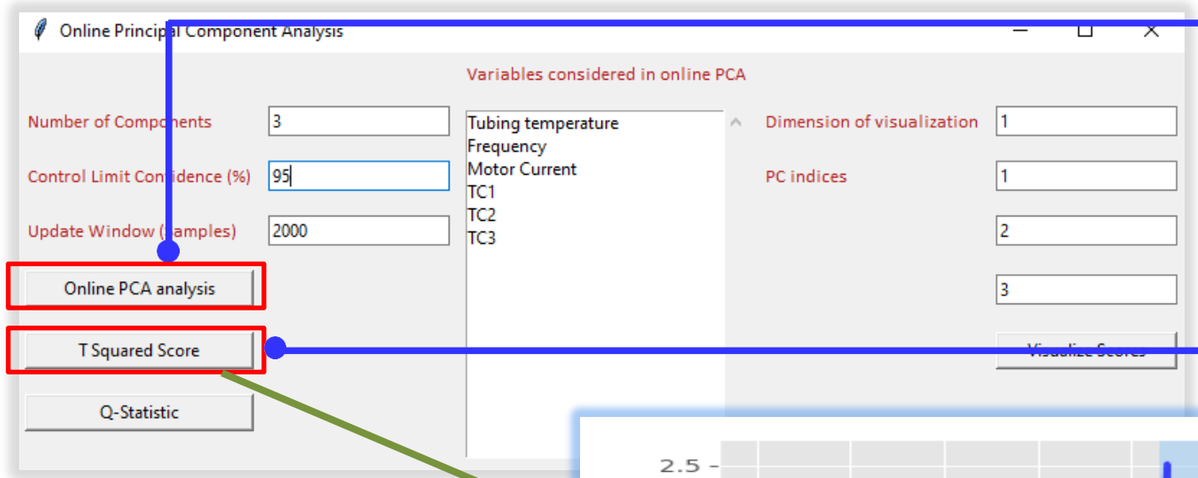
Predictive Monitoring Toolbox

Online monitoring – Principal component analysis



Predictive Monitoring Toolbox

Online monitoring – Principal component analysis (PCA)



2

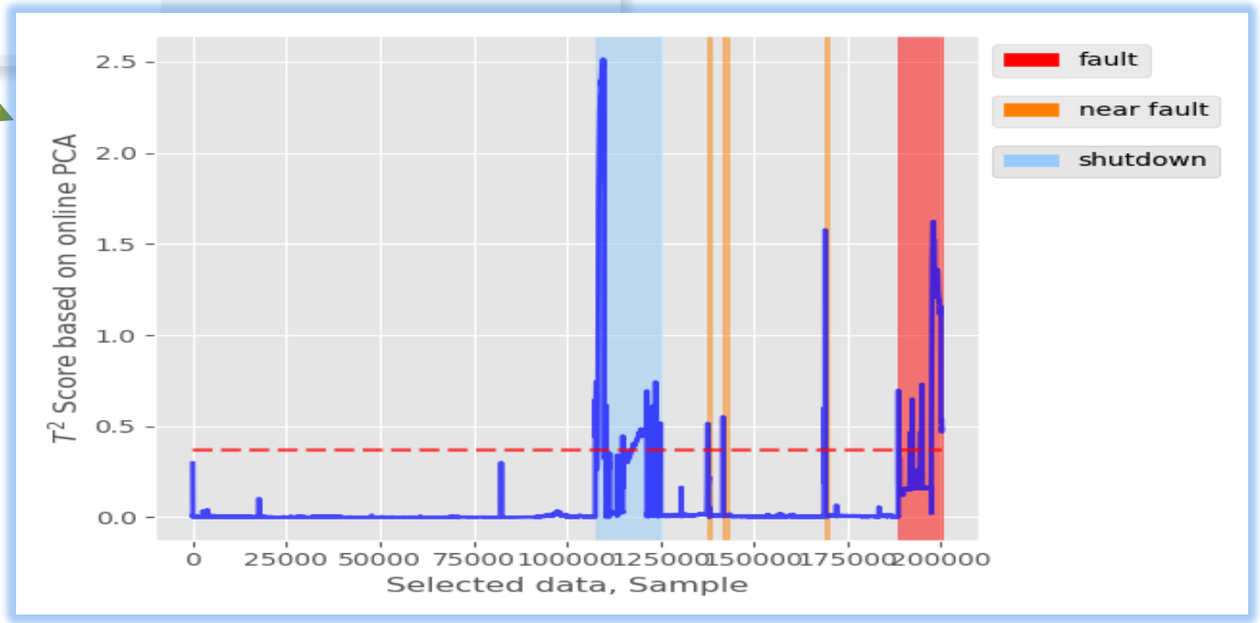
Click "**Online PCA analysis**" button

- PCA is conducted, and the model is updated after every 2000 samples (the size of the update window).

3

Click "**T Squared Score**" button

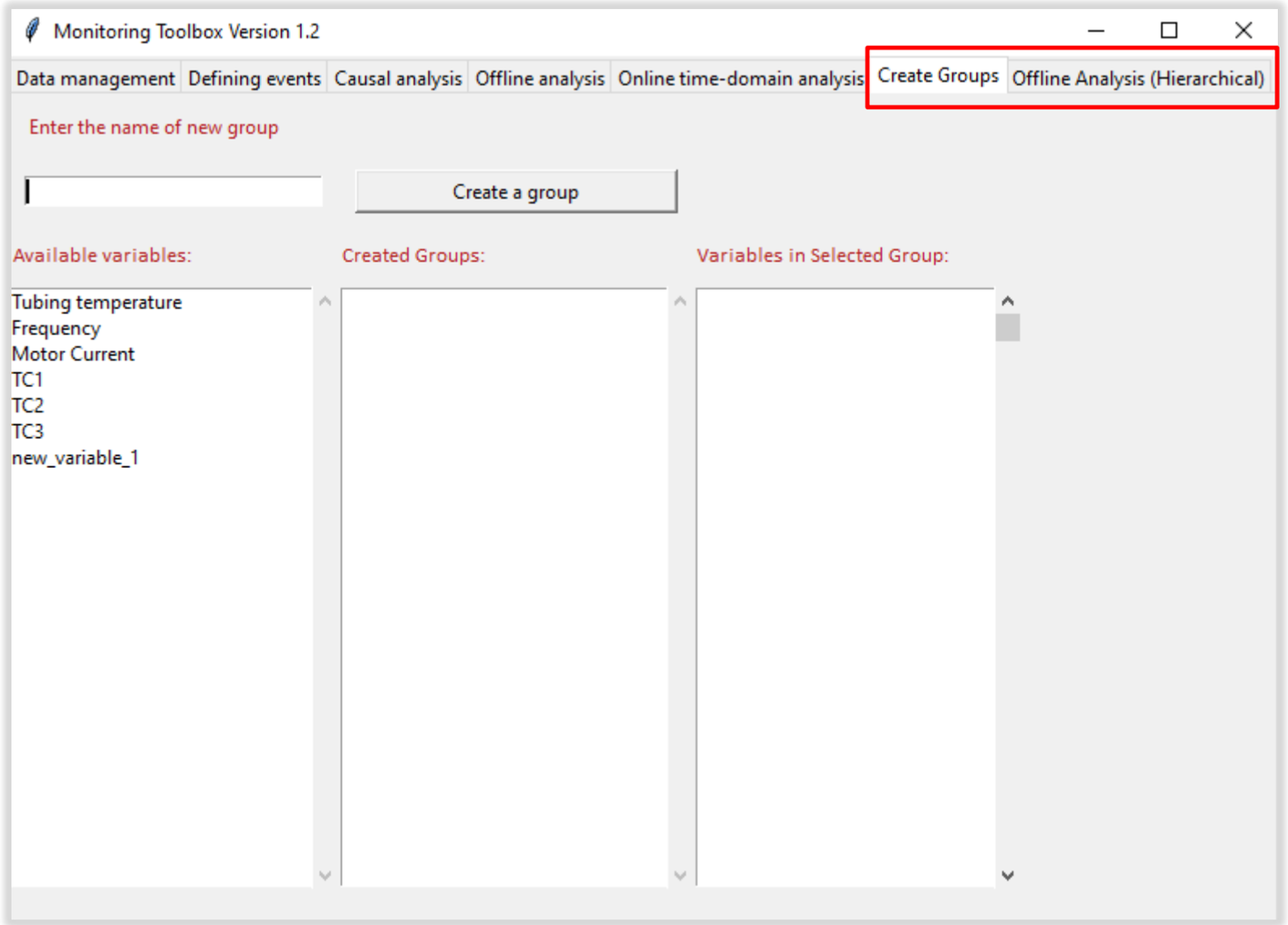
- This generates the T-squared score based on the online PCA analysis.

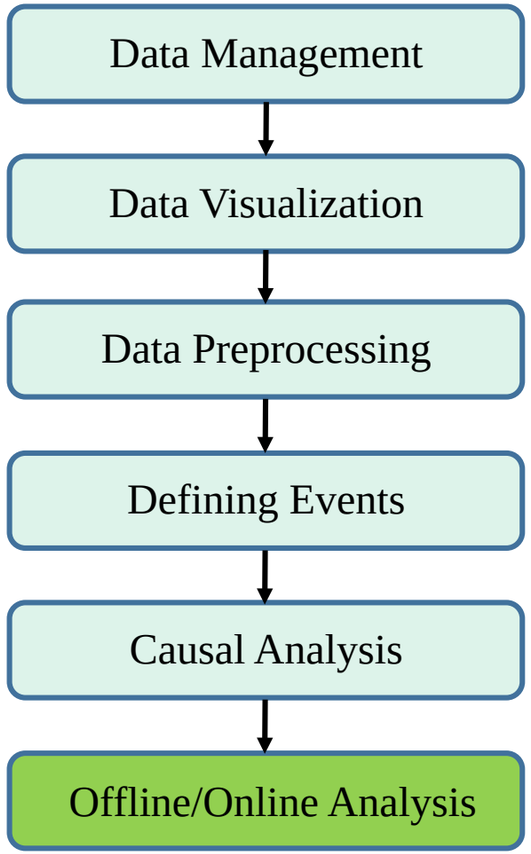


Predictive Monitoring Toolbox

Hierarchical distributed monitoring (this is favorable when the number of variables is large)

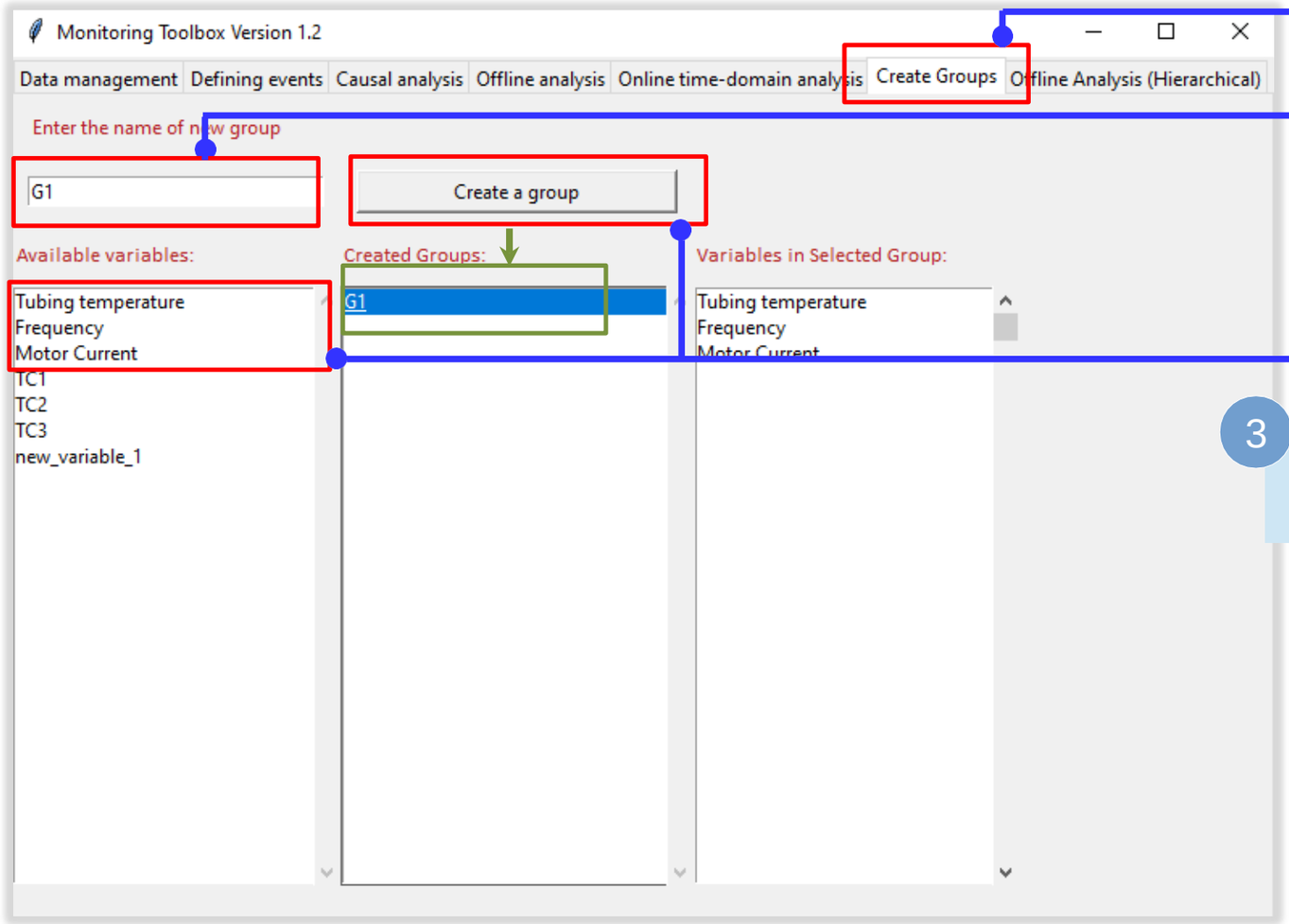
Both "**Create Groups**" and "Offline Analysis (Hierarchical)" tabs are used.





Predictive Monitoring Toolbox

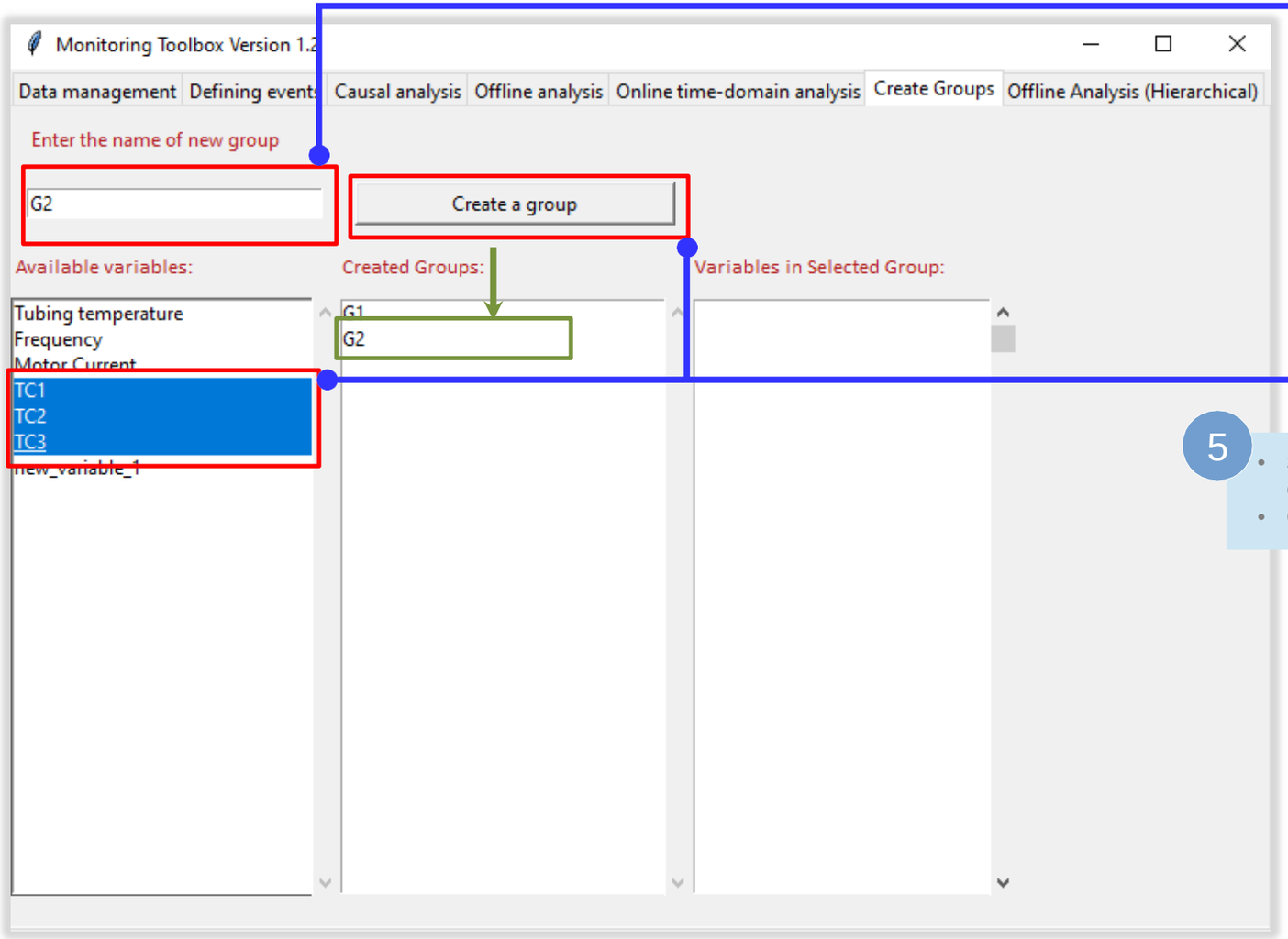
Hierarchical distributed monitoring – Construct groups of variables



1. Go to "**Create Groups**" Tab to assign variables to each of the groups for low-level analysis
2. Enter "**G1**" as the name of this group
3.
   - Select "**Tubing temperature**", "**Frequency**", "**Motor Current**" from the first list of the current window
   - Click "**Create a group**" button

Predictive Monitoring Toolbox

Hierarchical distributed monitoring – Construct groups of variables

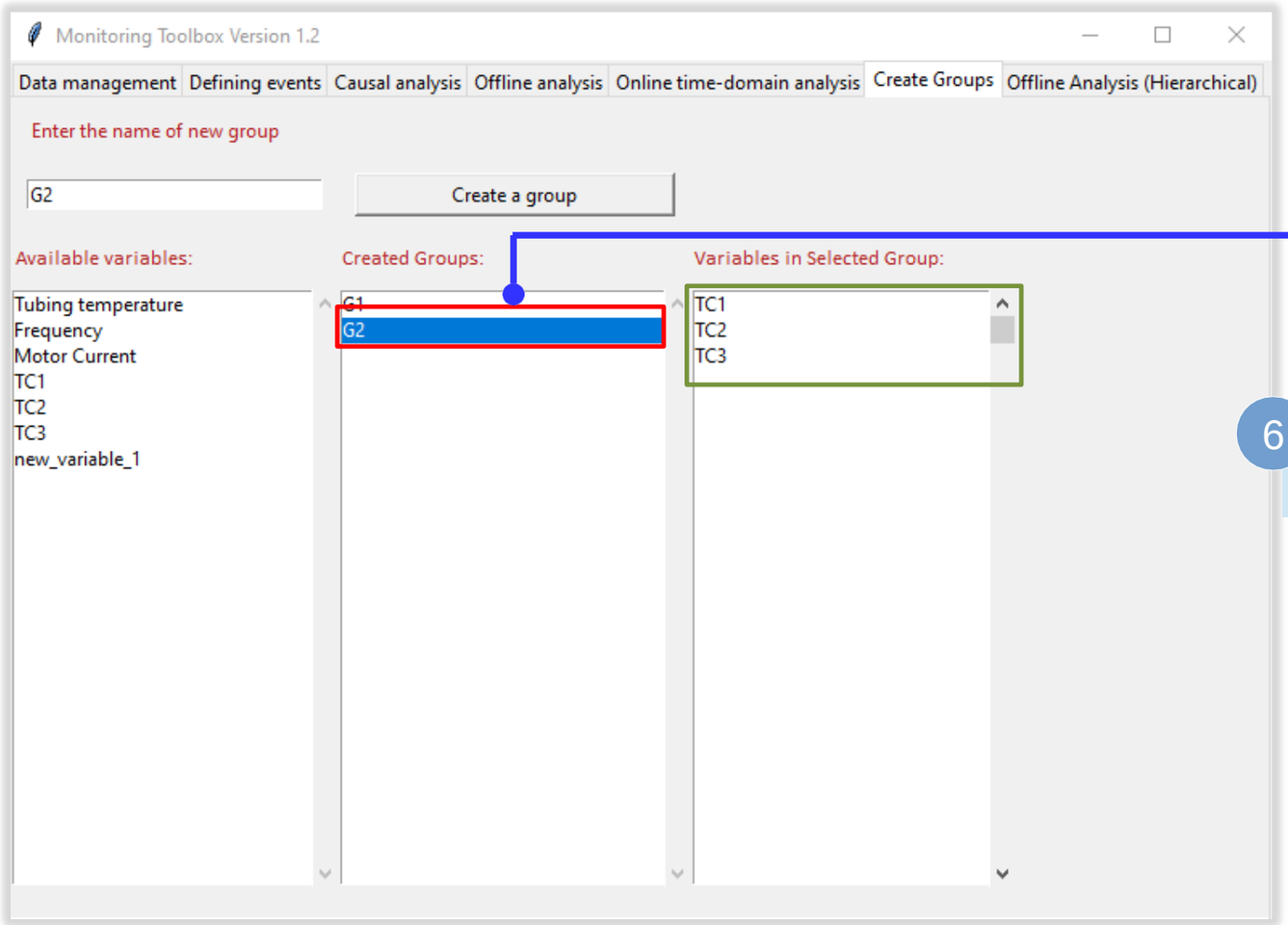


4 Enter "**G2**" as the name of this group

5
- Select "**TC1**", "**TC2**", "**TC3**" from the first list of the current window
- Click "**Create a group**" button

Predictive Monitoring Toolbox

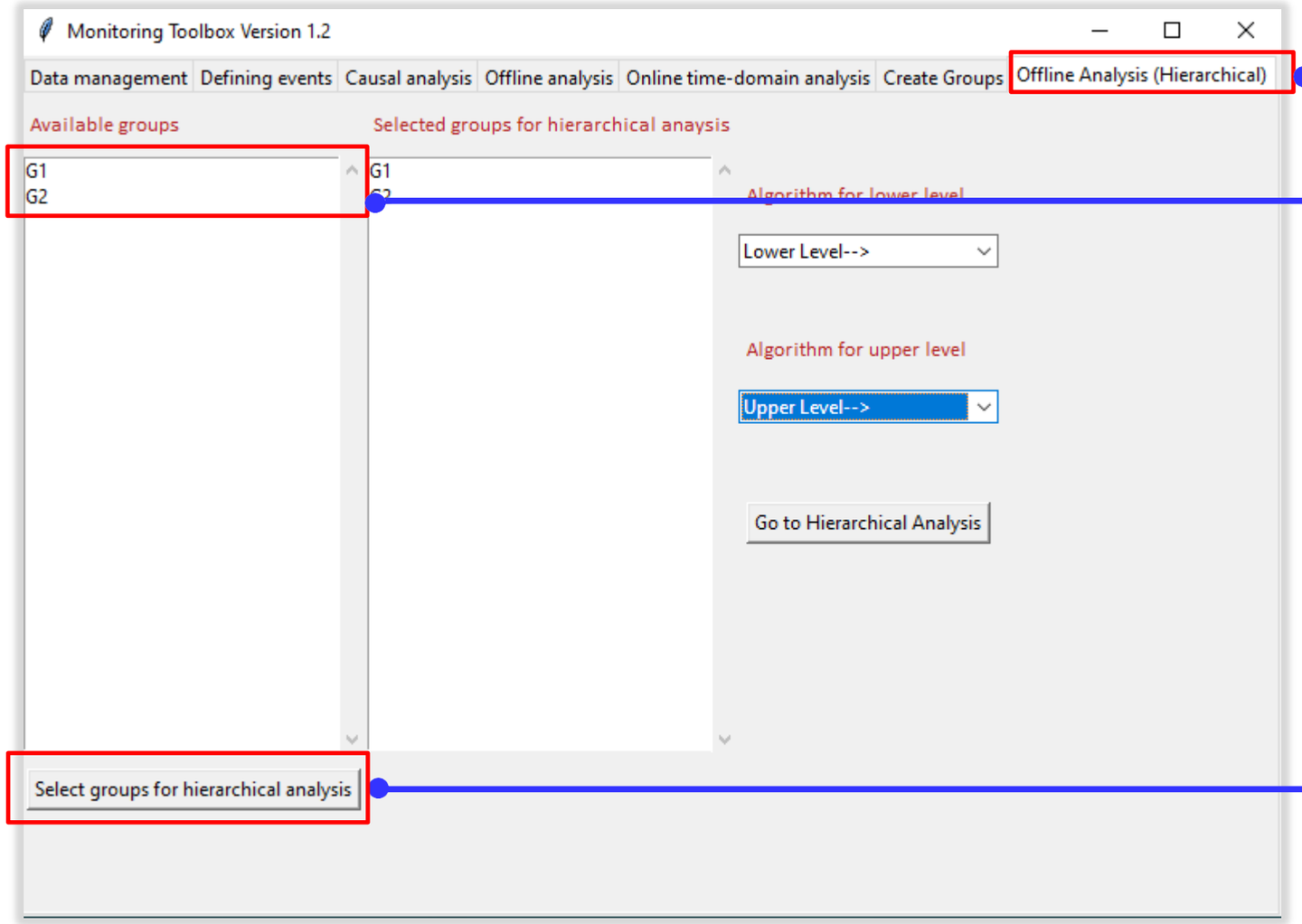
Hierarchical distributed monitoring – Construct groups of variables



6 Click on “G1” or “G2” to see the variables assigned to each group

Predictive Monitoring Toolbox

Hierarchical distributed monitoring – Monitoring analysis



1

Go to "**Offline Analysis (Hierarchical)**" tab for hierarchical distributed monitoring

2

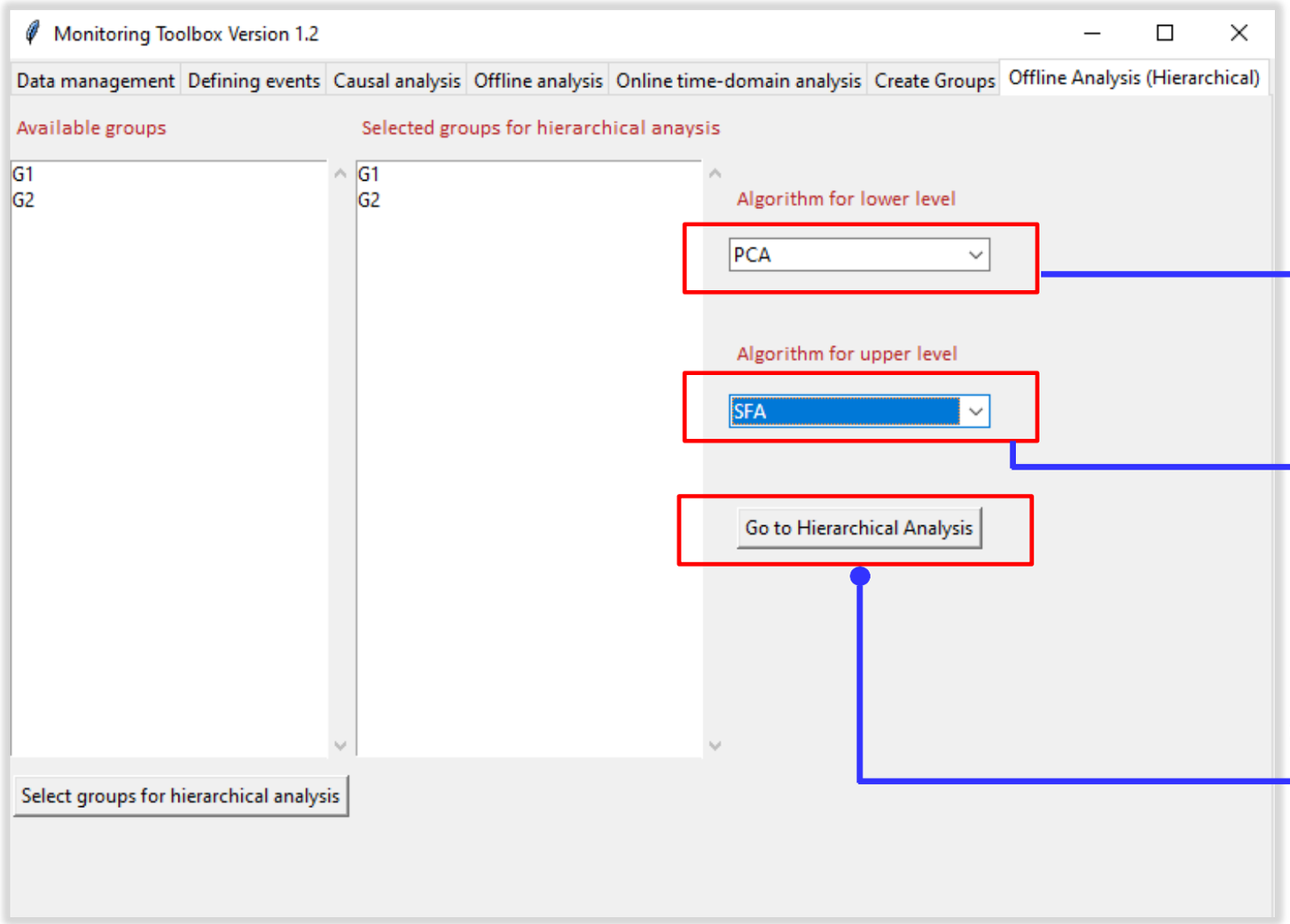
- Select "**G1**" and "**G2**" from the left list

3

- Click "**Select groups for hierarchical analysis**" so that these two groups of variables will be taken into account

Predictive Monitoring Toolbox

Hierarchical distributed monitoring – Monitoring analysis



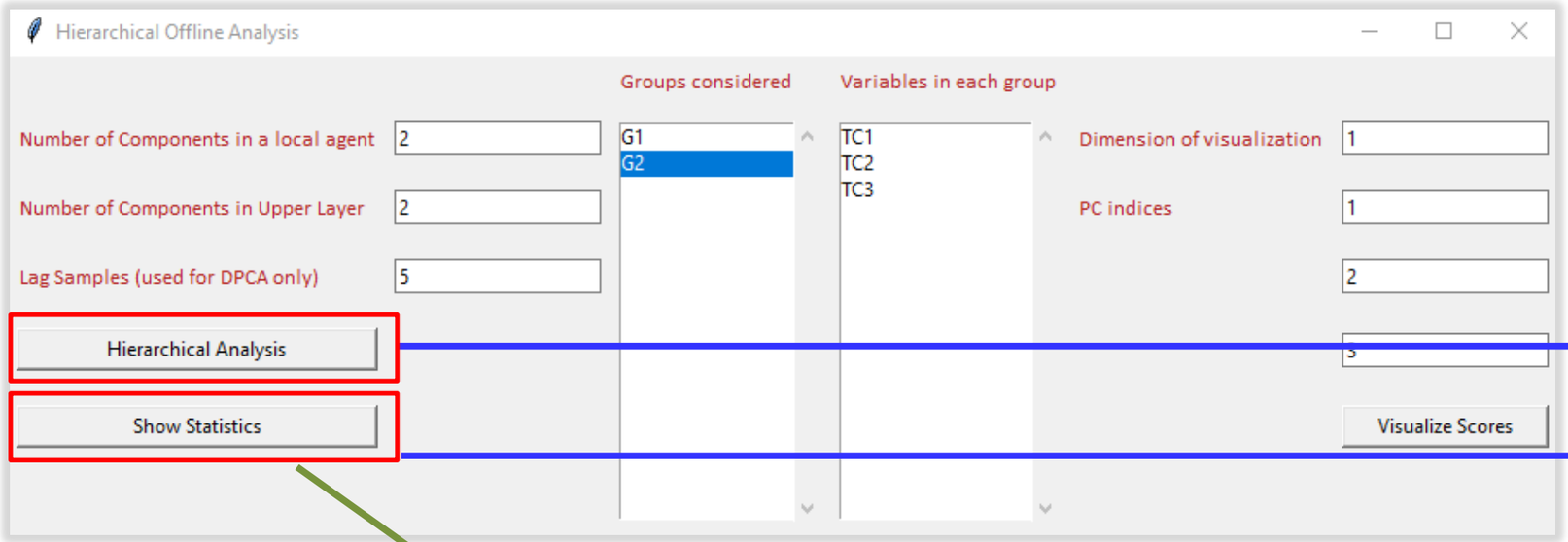
1 Choose "**PCA**" as the monitoring algorithm for each local group

2 Choose "**SFA**" as the monitoring algorithm for the upper level

3 Click "**Go to Hierarchical Analysis**" button to conduct two-layer hierarchical monitoring

Predictive Monitoring Toolbox

Hierarchical distributed monitoring – Monitoring analysis



4 Click “Hierarchical Analysis” button to conduct analysis

5 Click “Show Statistics” button to visualize results