



# Monitoring Toolbox V1.2 Quick Start Guide

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



## 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



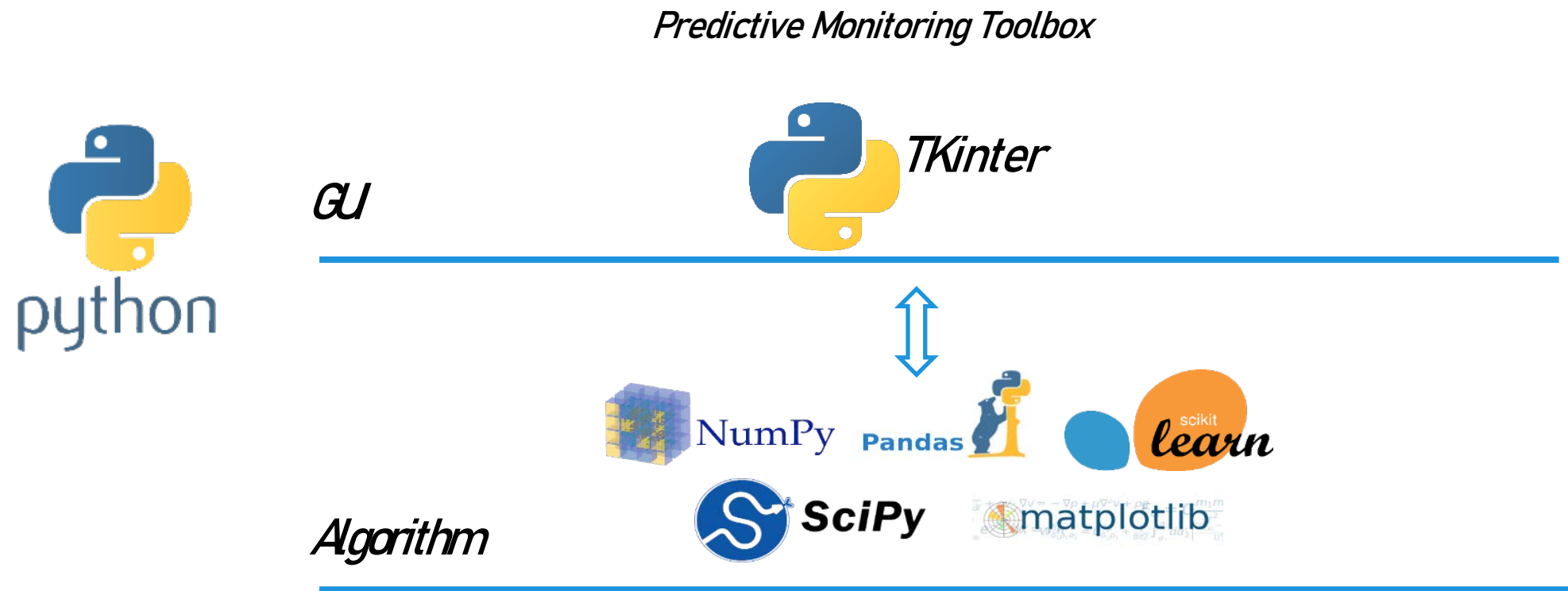
# Overview of Predictive Monitoring Toolbox

# Predictive Monitoring Toolbox



## An overview of the toolbox

The toolbox was developed using Python



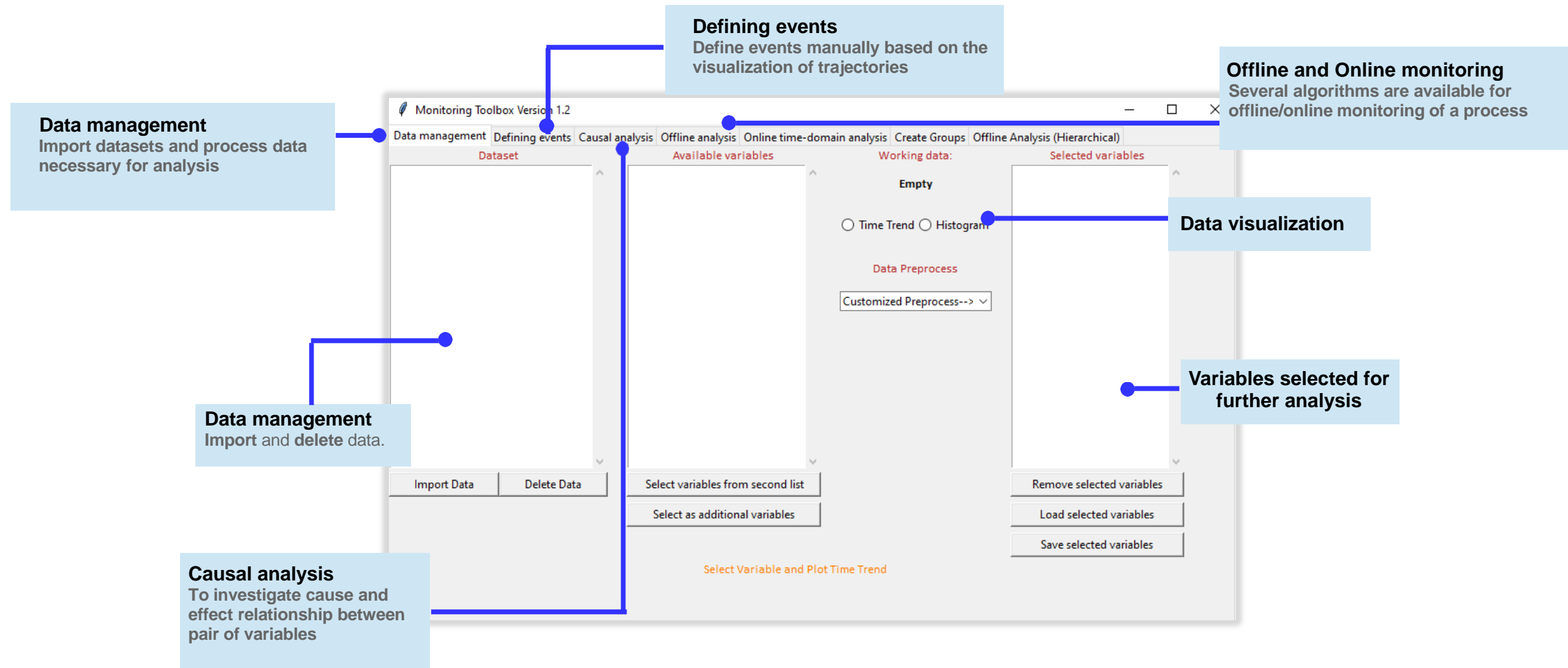
- ❖ *Python has open source license (free to use)*
- ❖ *Python possesses stable libraries*
- ❖ *Python is developer friendly software*

# Predictive Monitoring Toolbox



## An overview of the toolbox

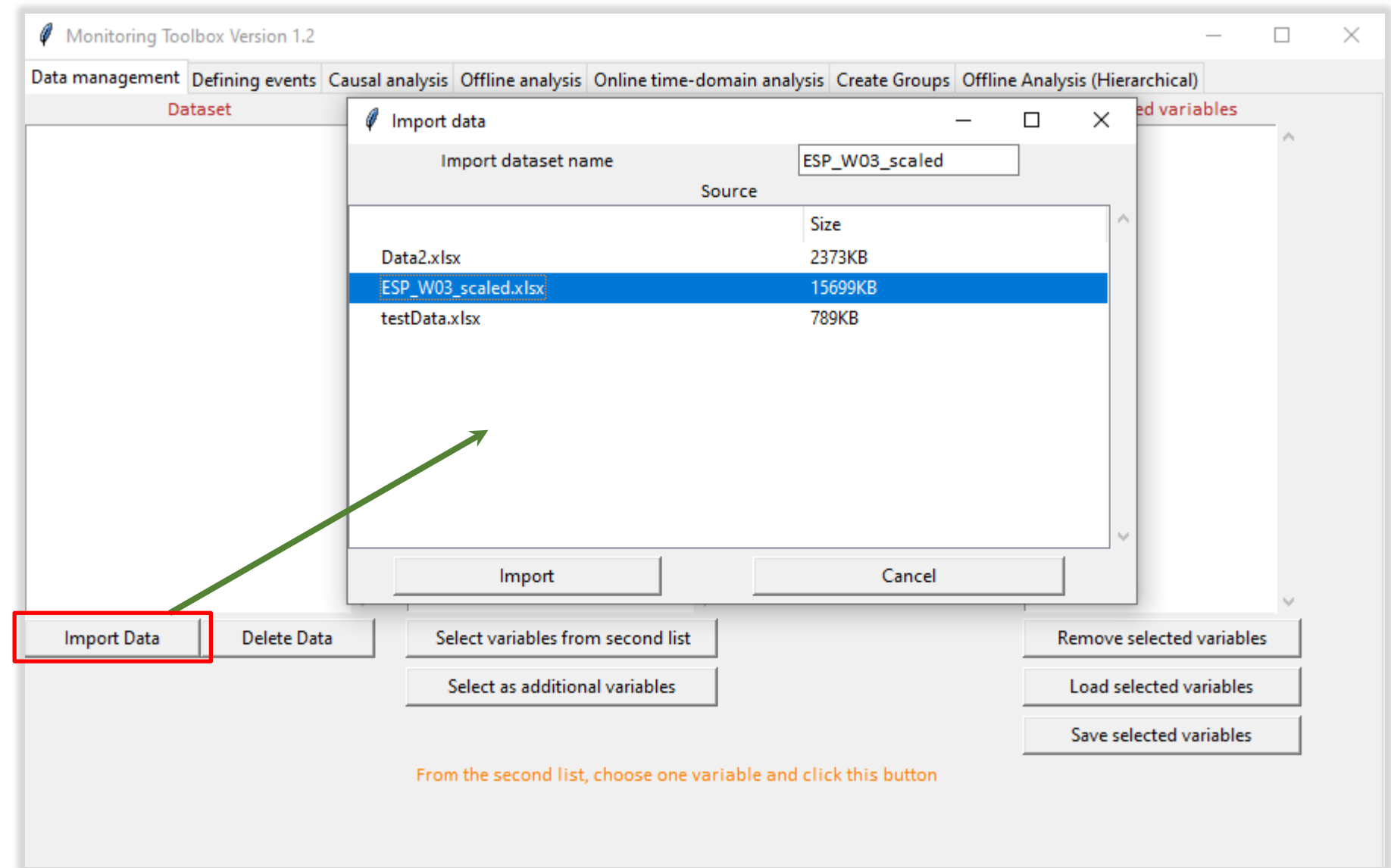
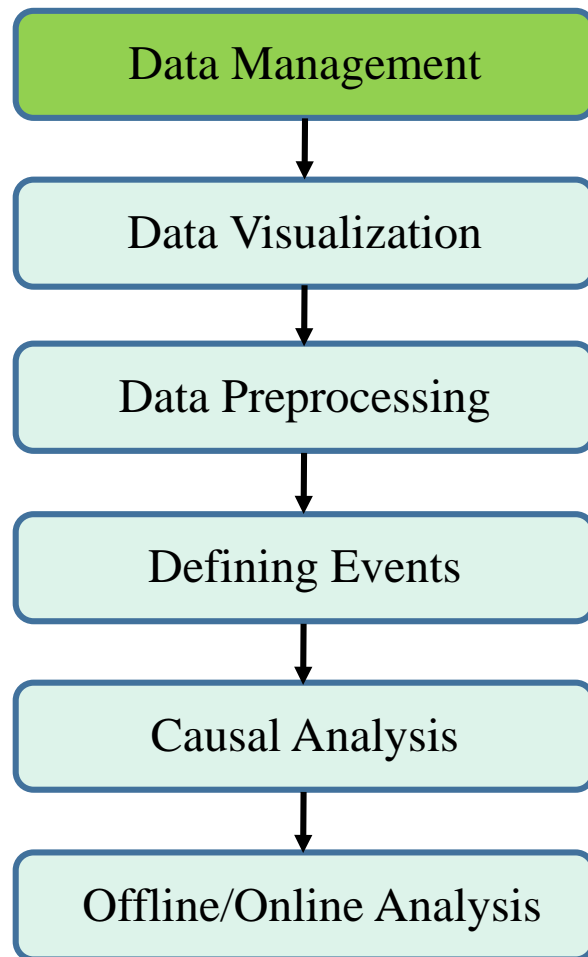
We create this guide to help you minimize the learning curve.



# Predictive Monitoring Toolbox



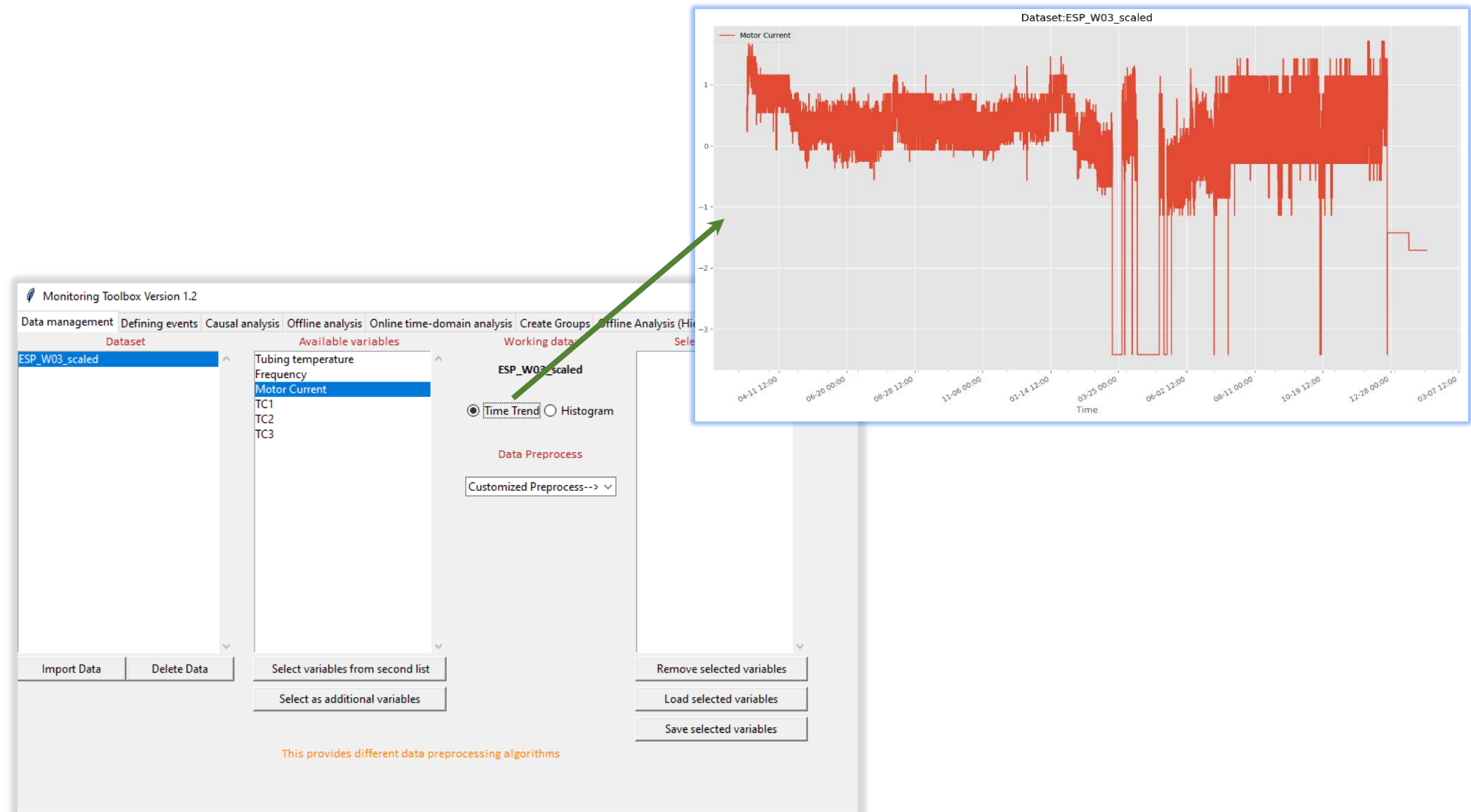
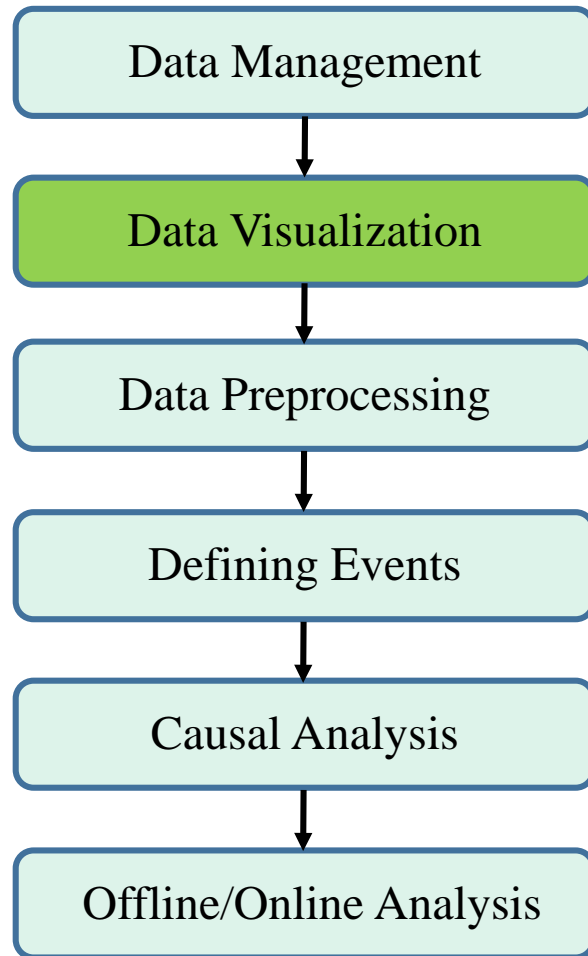
## An overview of the toolbox – Data management



# Predictive Monitoring Toolbox



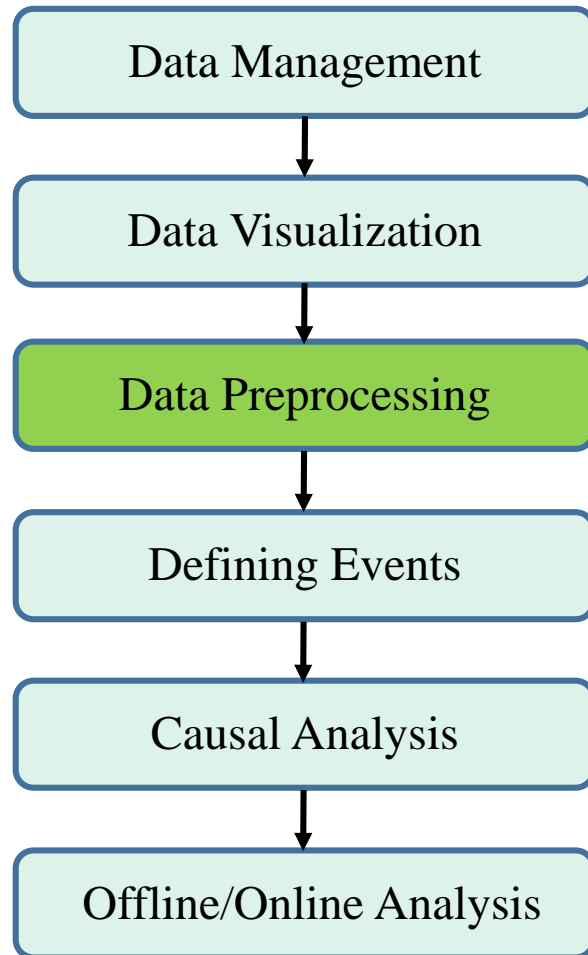
## An overview of the toolbox – Data visualization



# Predictive Monitoring Toolbox



## An overview of the toolbox – Data preprocessing



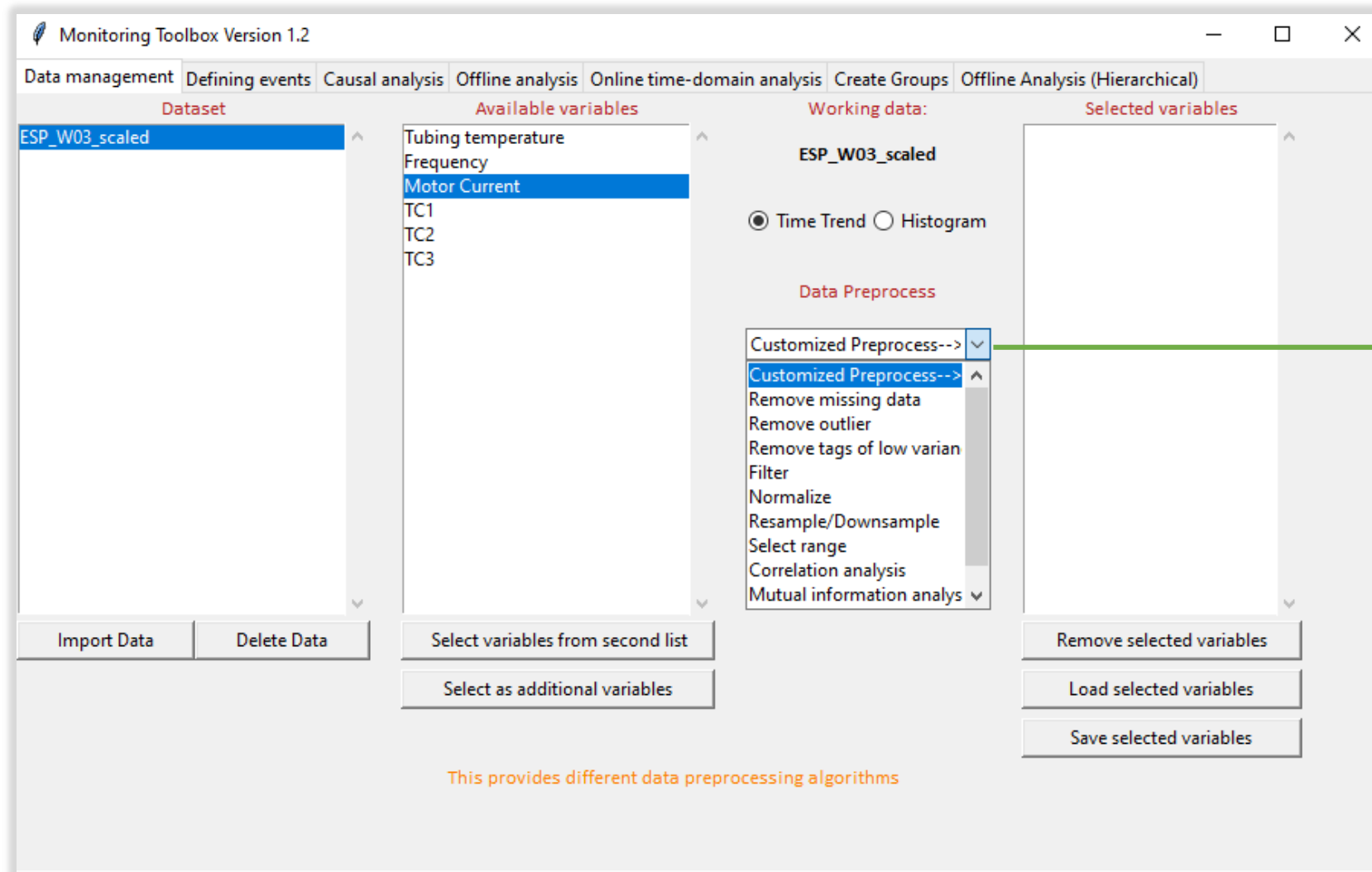
The screenshot shows the "Monitoring Toolbox Version 1.2" software interface. The main window has several tabs: "Data management", "Defining events", "Causal analysis", "Offline analysis", "Online time-domain analysis", "Create Groups", and "Offline Analysis (Hierarchical)". The "Data management" tab is active, showing a "Dataset" list with "ESP\_W03\_scaled" selected. Below the dataset list are "Import Data" and "Delete Data" buttons. To the right is a list of "Available variables" including "Tubing temperature", "Frequency", "Motor Current", "TC1", "TC2", and "TC3". Below this list are "Select variables from second list" and "Select as additional variables" buttons. Further right, the "Working data:" section shows "ESP\_W03\_scaled" and two radio buttons: "Time Trend" (selected) and "Histogram". A red box highlights a "Data Preprocess" section containing a dropdown menu with "Customized Preprocess-->". To the right of this is a "Selected variables" list, which is currently empty. Below the list are "Remove selected variables", "Load selected variables", and "Save selected variables" buttons. At the bottom of the window, a note states: "This provides different data preprocessing algorithms".



# 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

The screenshot shows the 'Monitoring Toolbox Version 1.2' interface. The 'Data management' tab is active, displaying a list of datasets with 'ESP\_W03\_scaled' selected. The 'Available variables' list includes Tubing temperature, Frequency, Motor Current, TC1, TC2, and TC3. The 'Working data' section shows 'ESP\_W03\_scaled' with 'Time Trend' selected. The 'Data Preprocess' dropdown menu is open, listing various algorithms: Customized Preprocess, Remove missing data, Remove outlier, Remove tags of low variance, Filter, Normalize, Resample/Downsample, **Select range** (highlighted), Correlation analysis, and Mutual information analysis. A green arrow points from the 'Select range' option to a 'Select range' dialog box. This dialog box shows a time series plot of 'Tubing temperature' and a histogram of its frequency distribution. The 'Start' time is 2012-03-09 22:15:00 and the 'End' time is 2014-02-03 01:35:00. The plot shows a signal fluctuating around zero with several sharp negative spikes. The histogram shows a distribution centered at zero with a few outliers.

This provides different data preprocessing algorithms

- 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

The screenshot displays the 'Monitoring Toolbox Version 1.2' interface. The main window shows a dataset 'ESP\_W03\_scaled' with available variables: Tubing temperature, Frequency, Motor Current, TC1, TC2, and TC3. The 'Data Preprocess' dropdown menu is open, listing various algorithms: Customized Preprocess-->, Remove missing data, Remove outlier, Remove tags of low variance, Filter, Normalize, Resample/Downsample, Select range, Correlation analysis, Mutual information analysis, and Define auxiliary tags. A green arrow points from 'Define auxiliary tags' to a separate dialog box.

The 'Define auxiliary tags' dialog box shows the following configuration:

- Available variables: Tubing temperature, Frequency, Motor Current, TC1, TC2, TC3
- Selected variables for creating new tag: Frequency(x(1)), Motor Current(x(2)), TC1(x(3))
- Enter the name of new tag: new\_tag
- Formula used to define the new tag:  $x(1)/x(2)+x(3)$
- Buttons: Add the new variable

Missing data

Remove outliers

Low variance

Filter

Normalize

Resample

Range selection

Correlation

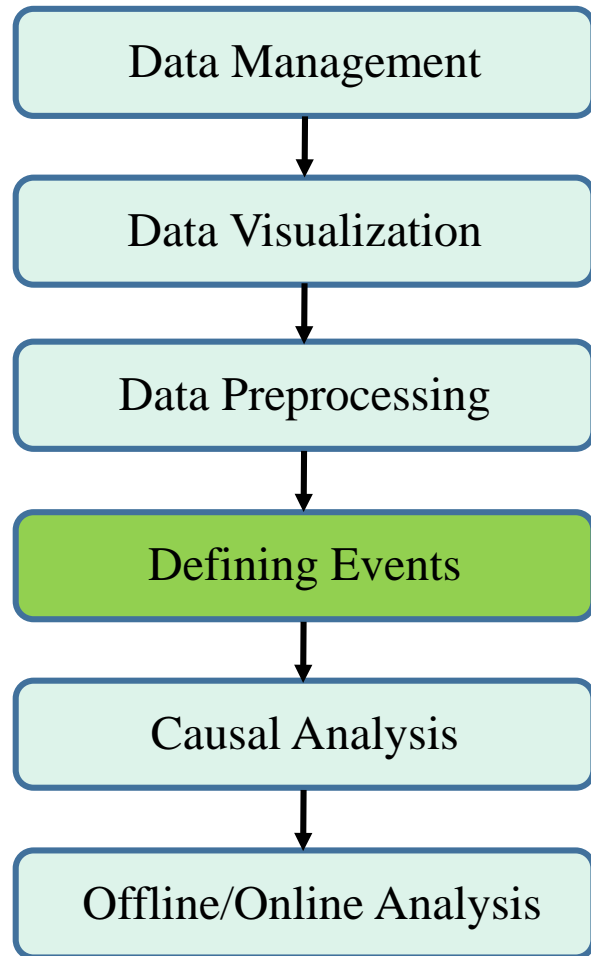
Mutual information

Auxiliary tags

# Predictive Monitoring Toolbox



## An overview of the toolbox – Defining events

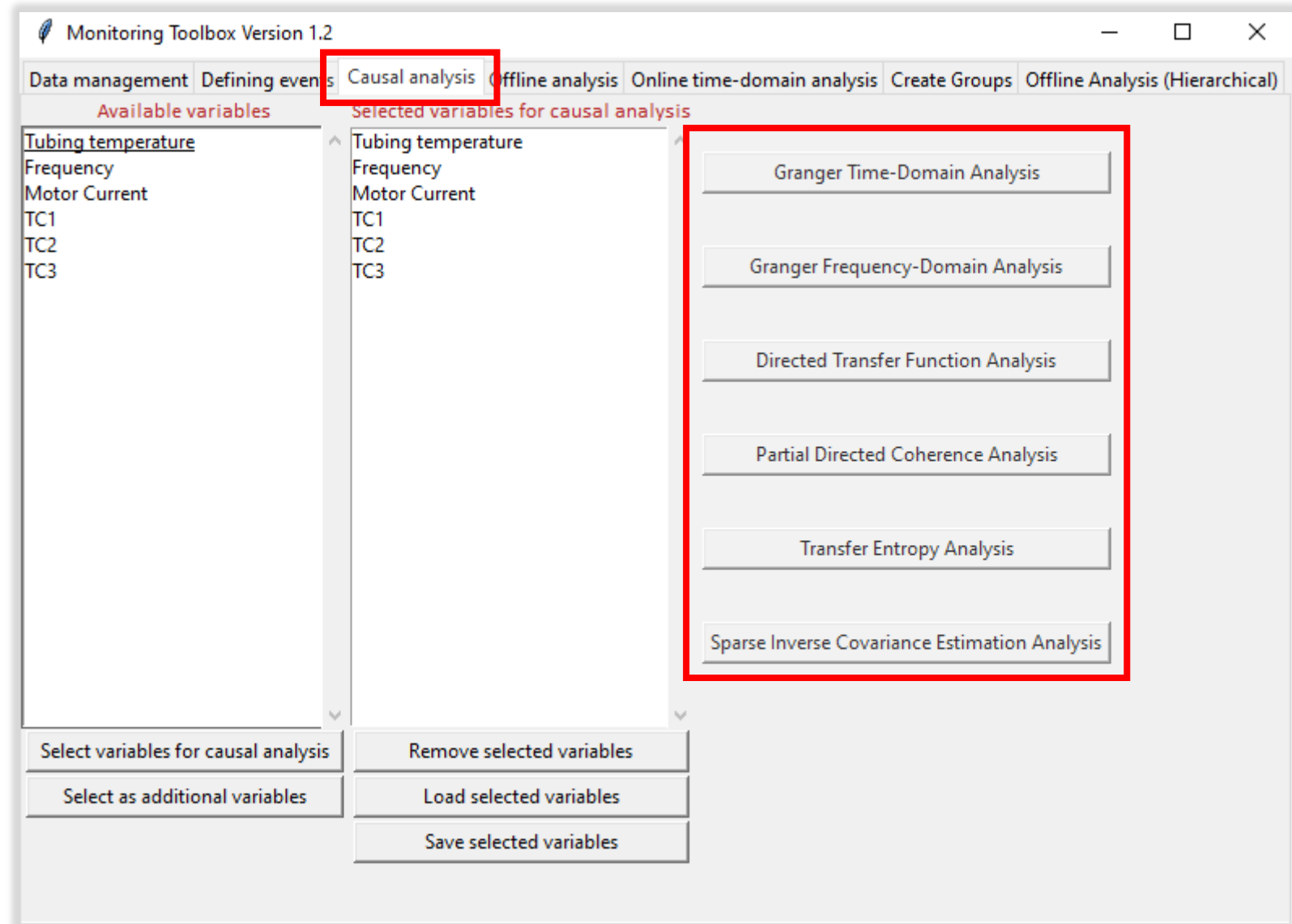
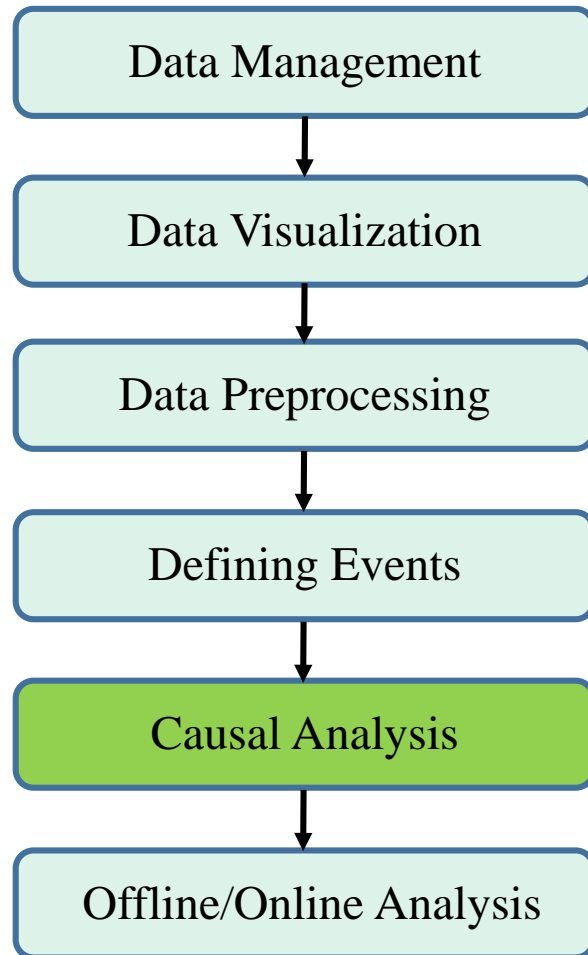


The screenshot displays the 'Monitoring Toolbox Version 1.2' interface. The 'Defining events' tab is active, showing a list of available variables (Tubing temperature, Frequency, Motor Current, TC1, TC2, TC3) and a 'Define Events' button. A red box highlights the 'Defining events' tab. A green arrow points from this box to a 'Select range' dialog box. The dialog box shows the 'Sample' range from 185448 to 200200 and the 'Time Span' from 2013-12-13 20:15:00 to 2014-02-03 01:35:00. Below the dialog box, three plots show the 'Tubing temperature' data for the selected range, with the range highlighted in red, orange, and blue.

# Predictive Monitoring Toolbox



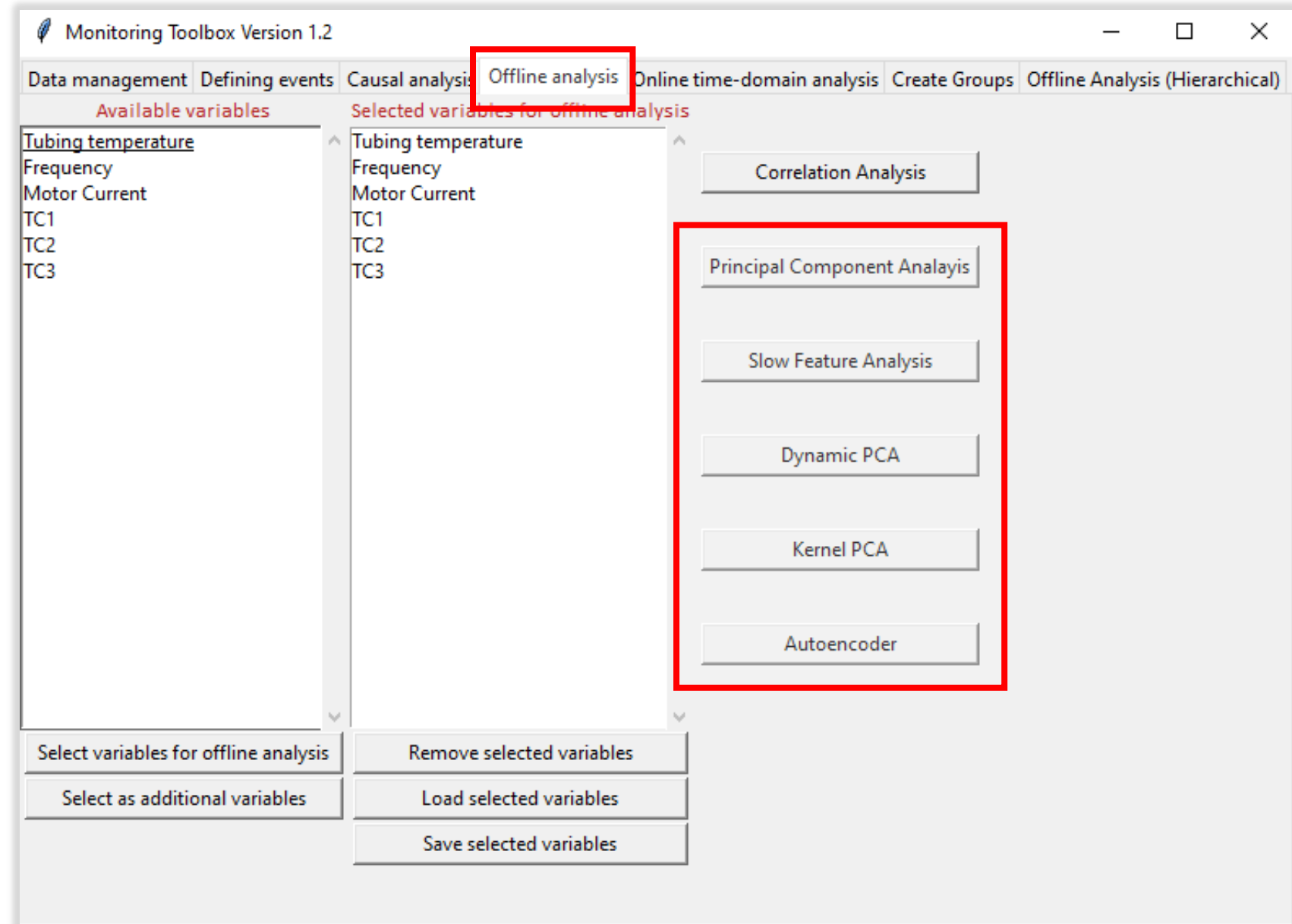
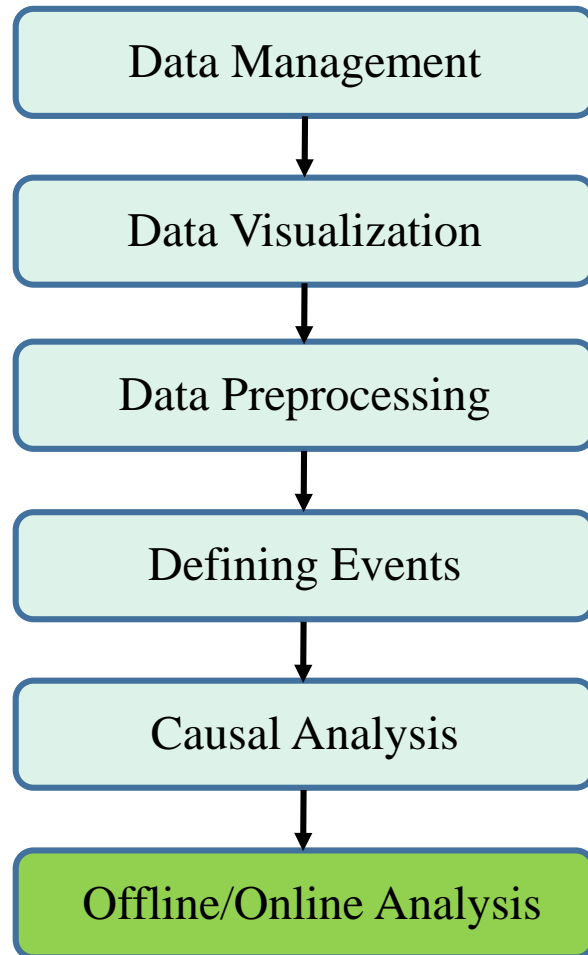
## An overview of the toolbox – Causal analysis



# Predictive Monitoring Toolbox



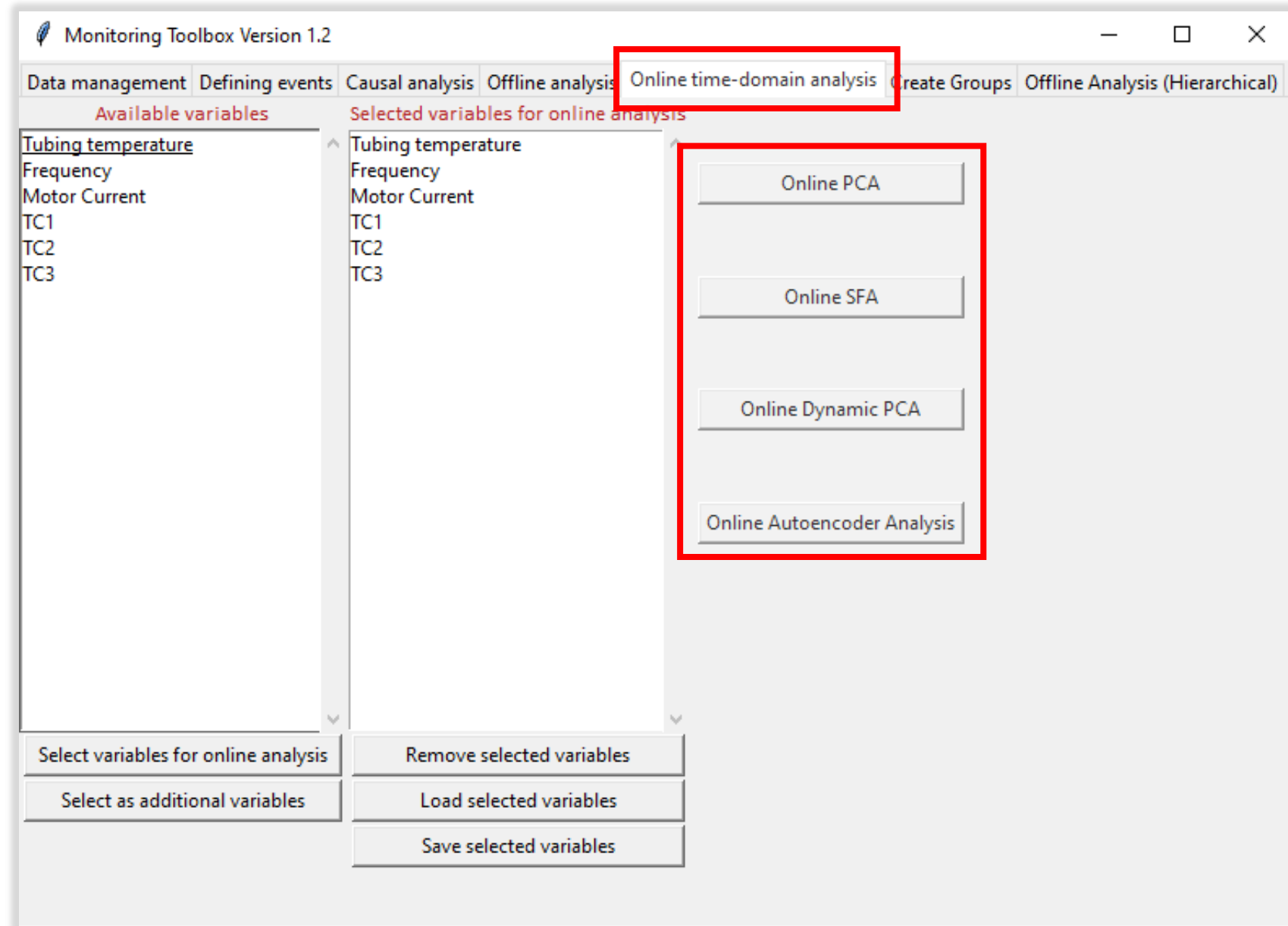
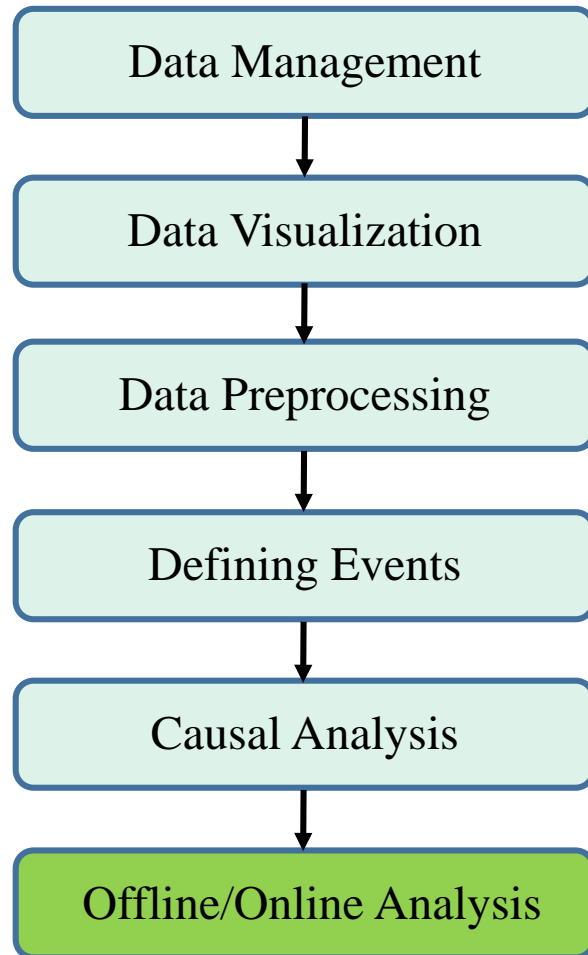
## An overview of the toolbox – Offline analysis



# Predictive Monitoring Toolbox



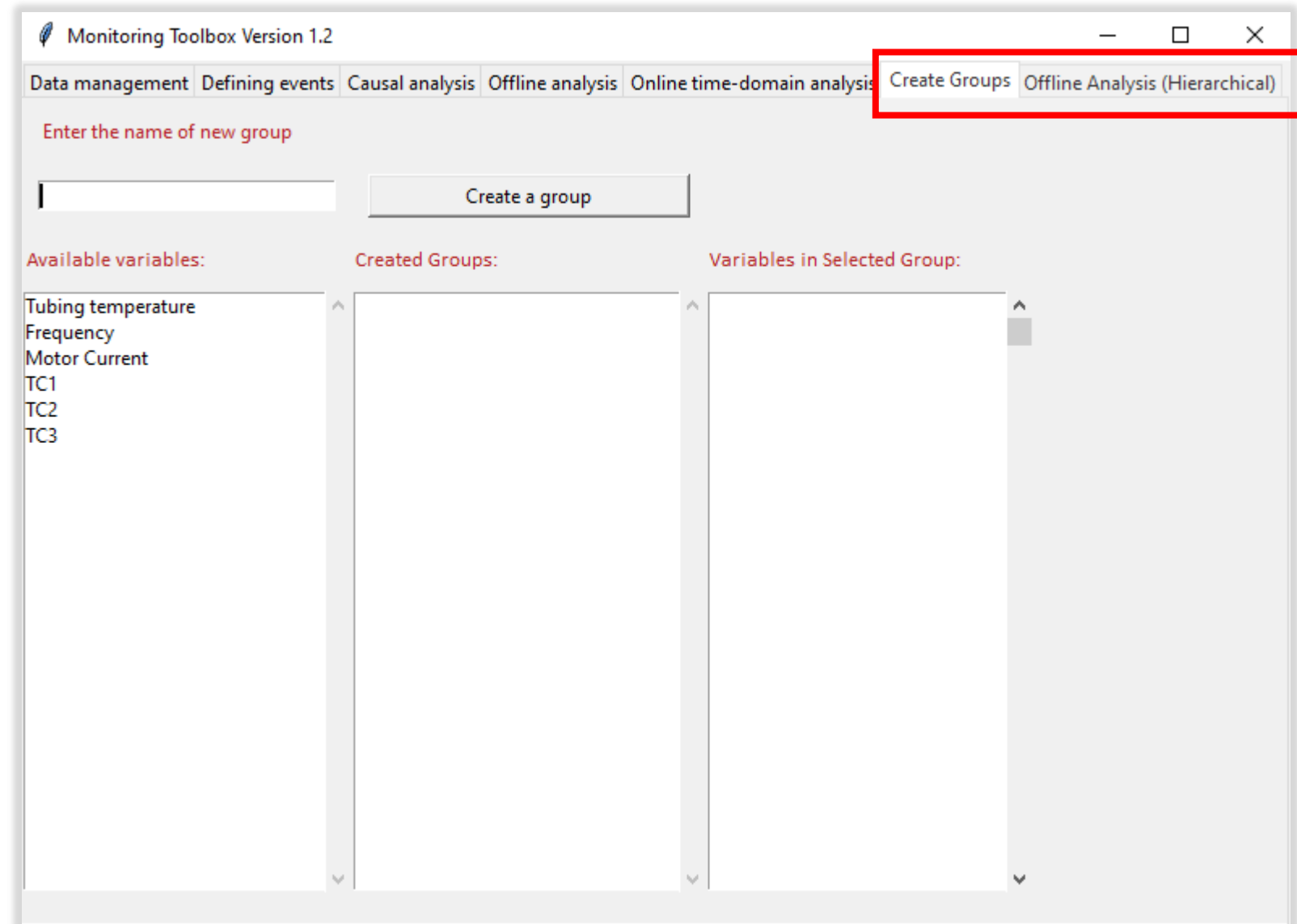
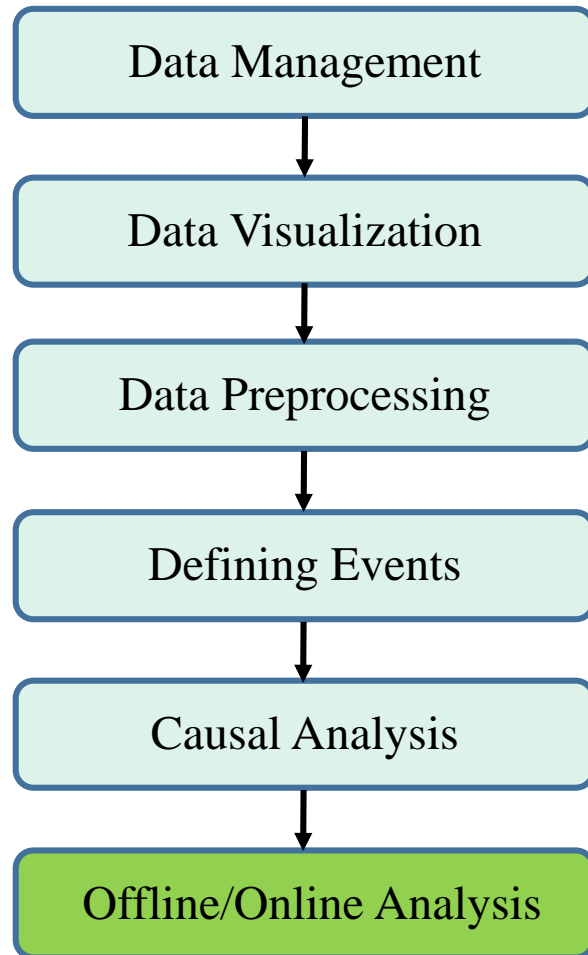
## An overview of the toolbox – Online analysis



# Predictive Monitoring Toolbox



An overview of the toolbox – Hierarchical distributed analysis







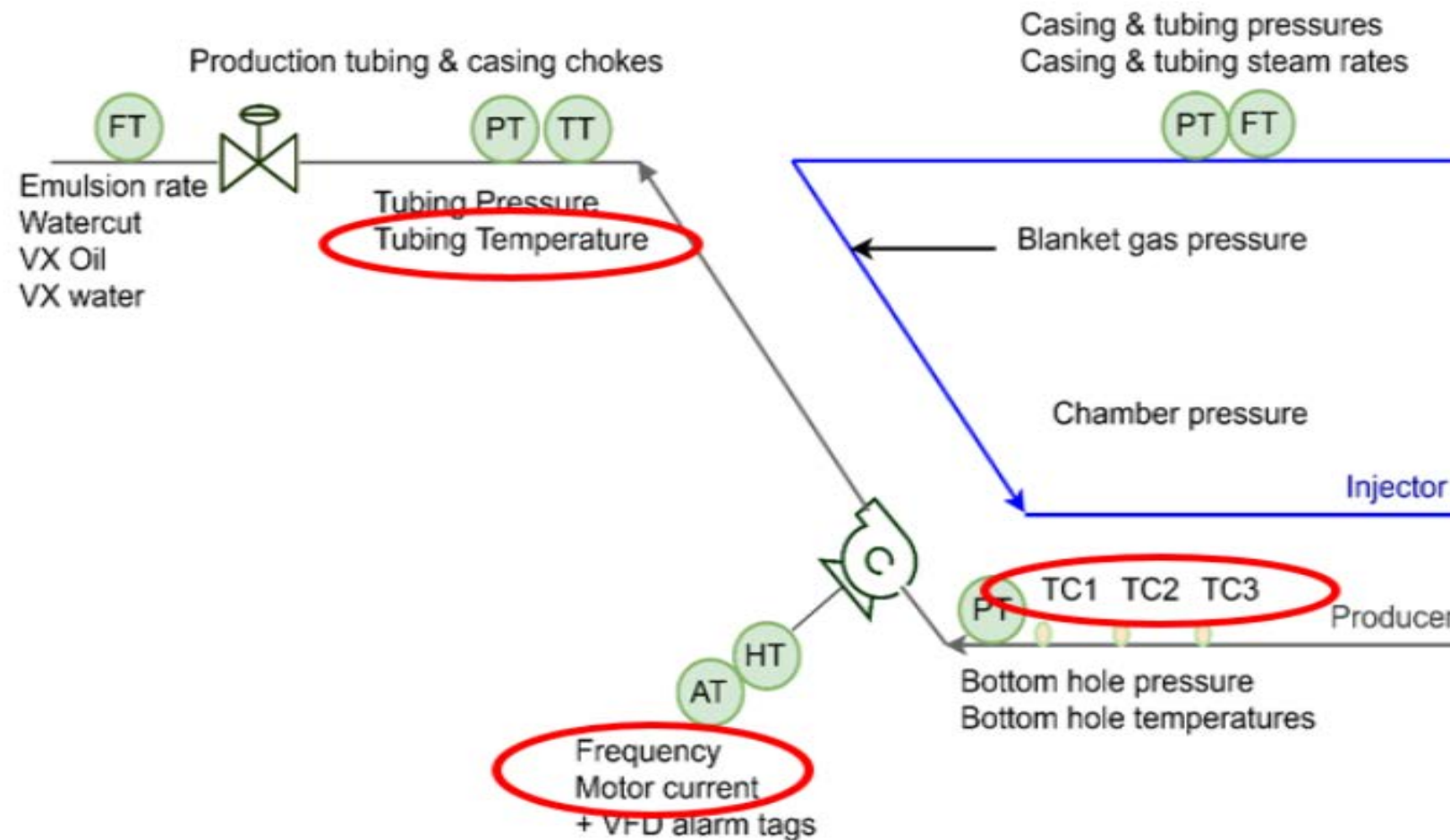
# Start Guide based on a Case Study

# 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

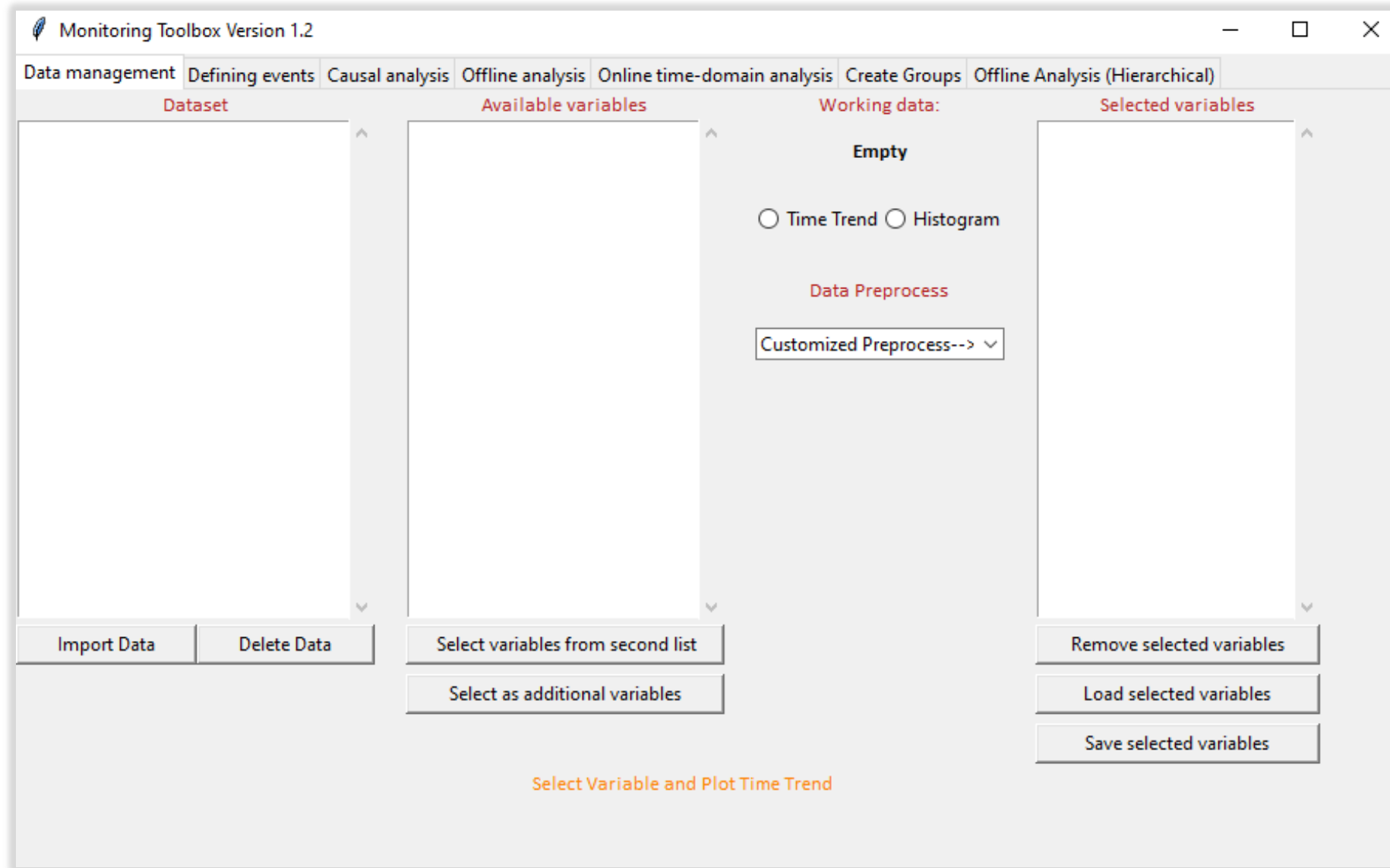
# Predictive Monitoring Toolbox



## To start

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

Click **“Monitoring\_v1.2.exe”** to start the program

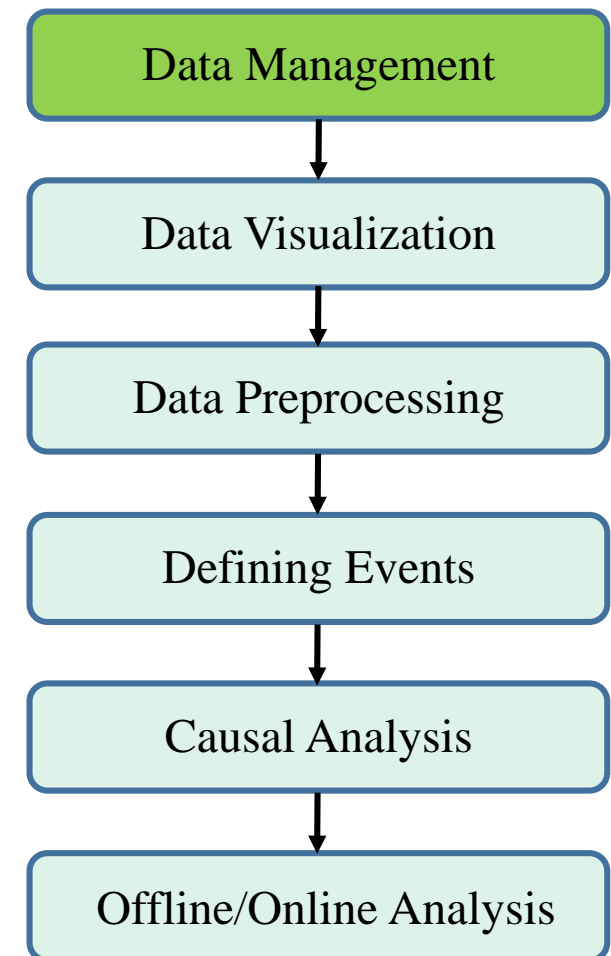
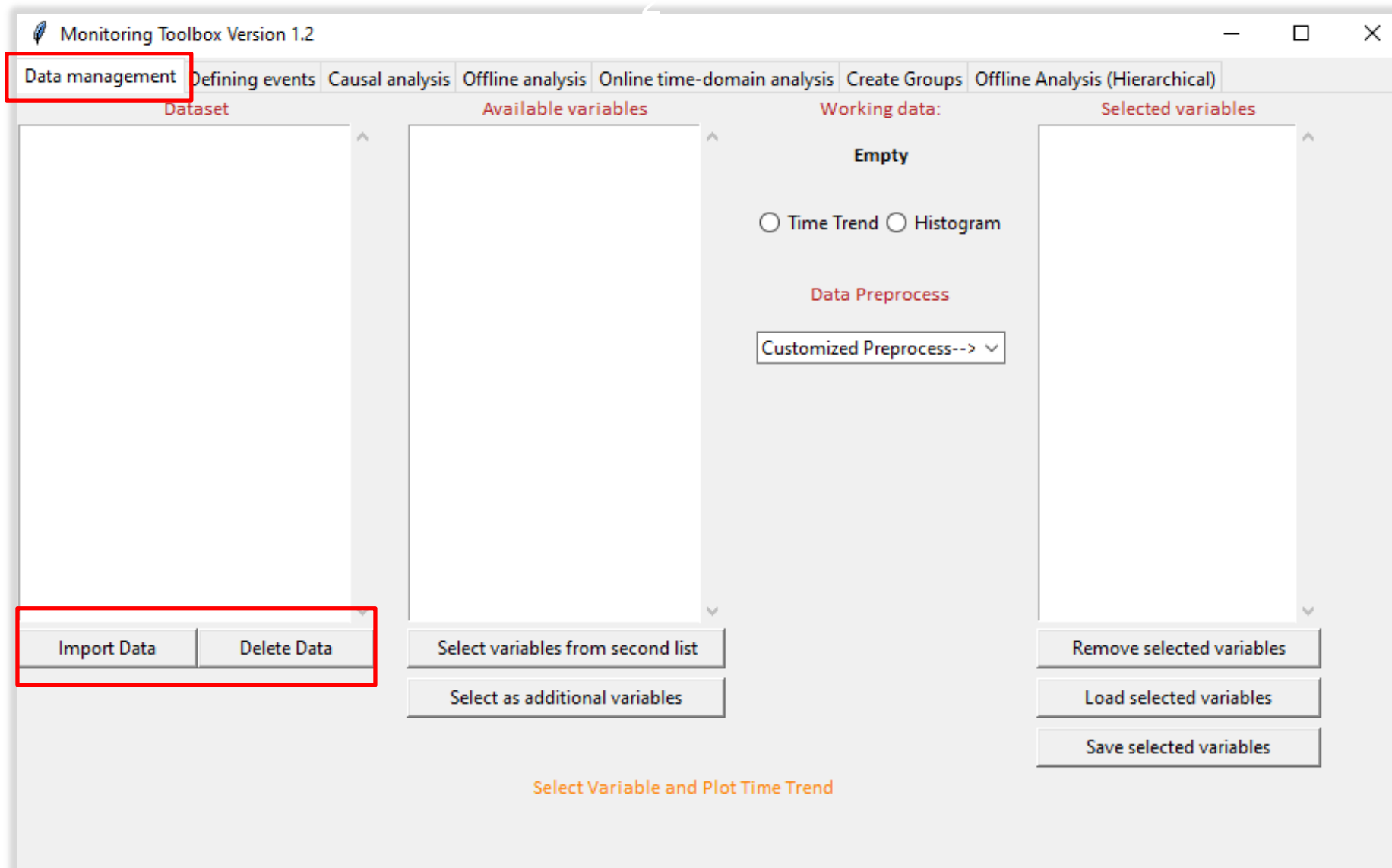


# Predictive Monitoring Toolbox



## 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

2

- Click “Import Data” button
- Select data file “EP\_W03\_scaled.xlsx”
- Click “Import” button

Source	Size
Data2.xlsx	2373KB
ESP_W03_scaled.xlsx	15699KB
flare1_scaled.xlsx	69206KB
testData.xlsx	789KB

1

- Go to “Data management” tab

“ESP\_W03\_Scaled” has been imported

Monitoring Toolbox Version 1.2

Data management Defining events Causal analysis Offline analysis Online time-domain analysis Create Groups Offline Analysis (Hierarchical)

Dataset Available variables Working data: Selected variables

ESP\_W03\_scaled

Empty

Time Trend  Histogram

Data Preprocess

Customized Preprocess-->

Import Data Delete Data

Select variables from second list

Select as additional variables

Remove selected variables

Load selected variables

Save selected variables

Select Data File/Click Import Button

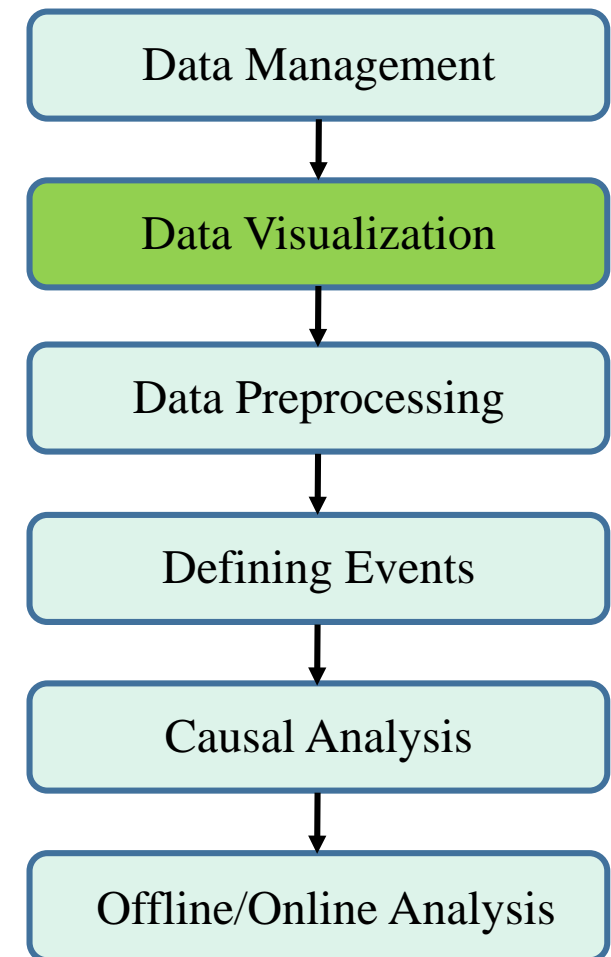
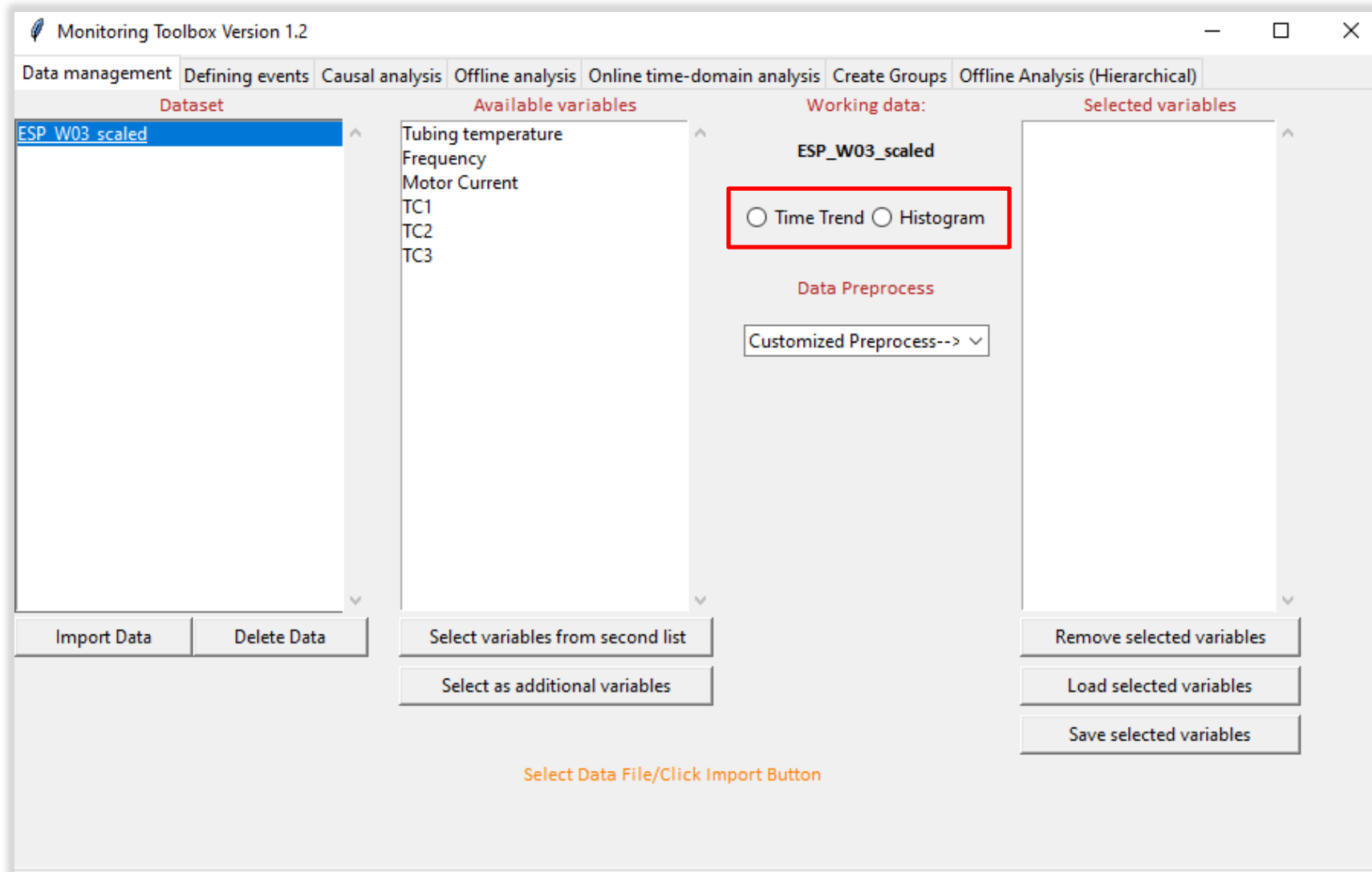
Note: Please save data sets in the “.xlsx” format

# Predictive Monitoring Toolbox



## Data visualization

Under the “**Data management**” tab, the variables can be visualized



# Predictive Monitoring Toolbox



## Data visualization

4

Drag-select to select variables for visualization  
Hold "CTRL" to multi-select if they are not consecutive

3

Click on the "Data management" tab.  
Click on "ESP\_W03\_scaled" from the left workspace (i.e., "Dataset" column) to select it

5

Click "Time Trend" to visualize the selected variables

Select Variable and Plot Histogram

The screenshot displays the Monitoring Toolbox Version 1.2 interface. The 'Data management' tab is active, showing a 'Dataset' column with 'ESP\_W03\_scaled' selected. The 'Available variables' list includes Tubing temperature, Frequency, Motor Current, TC1, TC2, and TC3. The 'Working data' section shows 'ESP\_W03\_scaled' with 'Time Trend' selected. The 'Selected variables' list is empty. The 'Data Preprocess' section has a 'Customized Preprocess-->' button. The 'Time Trend' radio button is highlighted with a red box. A green arrow points from the 'Time Trend' button to a time-series plot of 'Motor Current' (red) and 'TC1' (blue) over time. The plot shows significant fluctuations and a sharp drop in TC1 around 03-25 00:00.

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



# Predictive Monitoring Toolbox



## Data visualization

The screenshot displays the 'Monitoring Toolbox Version 1.2' interface. The 'Dataset' list on the left contains 'ESP\_W03\_scaled'. The 'Available variables' list includes 'Tubing temperature', 'Frequency', 'Motor Current', 'TC1', 'TC2', and 'TC3'. The 'Working data' section shows 'ESP\_W03\_scaled' with the 'Histogram' option selected. The 'Data Preprocess' dropdown is set to 'Customized Preprocess-->'. The 'Selected variables' list is empty. The histogram plot on the right, titled 'Dataset:ESP\_W03\_scaled', shows the frequency distribution for 'Motor Current' (red bars) and 'TC1' (blue bars). The x-axis ranges from -5 to 2, and the y-axis (Frequency) ranges from 0 to 100,000. A green arrow points from the 'Histogram' button to the plot, and a blue arrow points from a callout box to the button.

**6** Click "Histogram" to generate histogram for selected variables

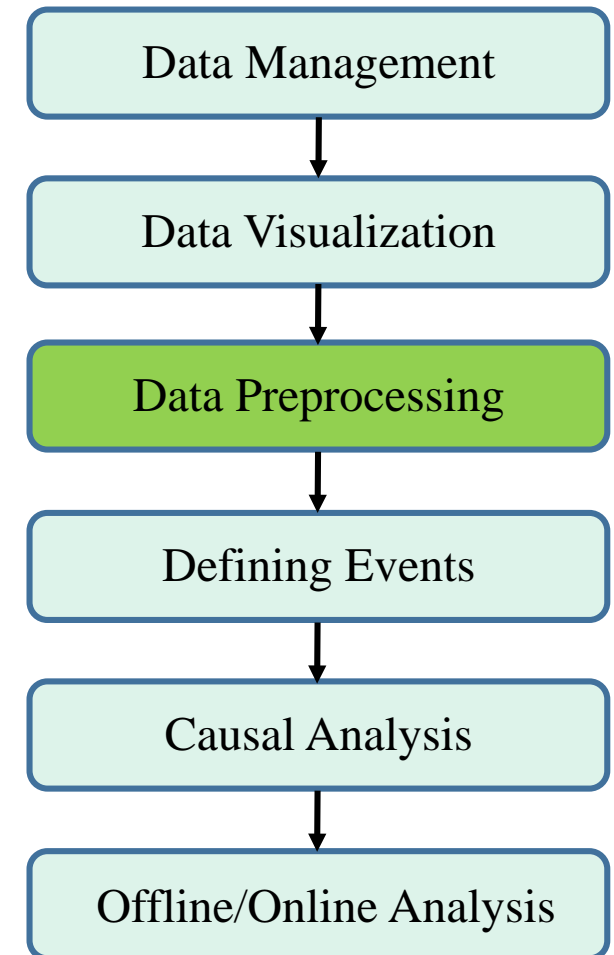
# Predictive Monitoring Toolbox



## Data preprocessing

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

The screenshot shows the "Monitoring Toolbox Version 1.2" interface. The "Data management" tab is active. The "Dataset" list contains "ESP\_W03\_scaled". The "Available variables" list includes "Tubing temperature", "Frequency", "Motor Current", "TC1", "TC2", and "TC3". The "Working data:" section shows "ESP\_W03\_scaled" with radio buttons for "Time Trend" and "Histogram". A red box highlights the "Data Preprocess" section, which contains a dropdown menu currently set to "Customized Preprocess--&gt;". The "Selected variables" list is empty. At the bottom, there are buttons for "Import Data", "Delete Data", "Select variables from second list", "Select as additional variables", "Remove selected variables", "Load selected variables", and "Save selected variables". A note at the bottom states: "This provides different data preprocessing algorithms".



# 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”

The screenshot displays the Monitoring Toolbox Version 1.2 interface. The 'Data management' tab is active. In the 'Dataset' column, 'ESP\_W03\_scaled' is selected and highlighted with a red box. In the 'Available variables' column, 'ESP\_W03\_scaled\_RM' is highlighted with a green box. The 'Data Preprocess' dropdown menu is open, with 'Remove missing data' selected and highlighted with a red box. A green arrow points from the 'Remove missing data' option to the 'ESP\_W03\_scaled\_RM' variable in the 'Dataset' column. Below the main interface, a text box states: 'This provides different data preprocessing algorithms'.

# 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

The screenshot shows the 'Monitoring Toolbox Version 1.2' interface. The 'Data management' tab is active. The 'Dataset' list contains 'ESP\_W03\_scaled' and 'ESP\_W03\_scaled\_RM', with the latter highlighted in a red box. The 'Available variables' list includes 'Tubing temperature', 'Frequency', 'Motor Current', 'TC1', 'TC2', and 'TC3'. The 'Working data:' section shows 'ESP\_W03\_scaled\_RM' selected, with 'Time Trend' and 'Histogram' options. The 'Data Preprocess:' section has a dropdown menu open, showing 'Customized Preprocessing-->', 'Remove missing data', 'Remove outlier', 'Remove tags of low variance', 'Filter', 'Normalize', 'Resample/Downsample', 'Select range', 'Correlation analysis' (highlighted in a red box), and 'Mutual information analysis'. At the bottom, there are buttons for 'Import Data', 'Delete Data', 'Select variables from second list', 'Select as additional variables', 'Remove selected variables', 'Load selected variables', and 'Save selected variables'. A note at the bottom states: 'This provides different data preprocessing algorithms'.

# Predictive Monitoring Toolbox



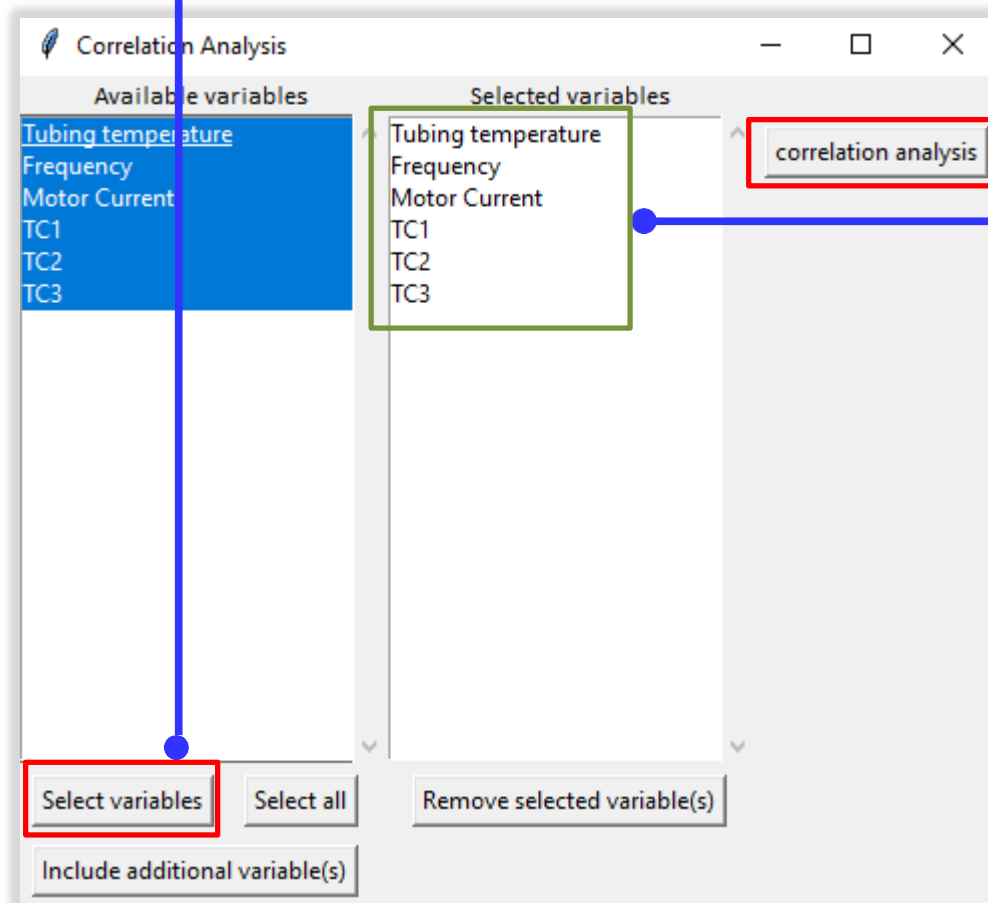
## Data preprocessing – Correlation analysis

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

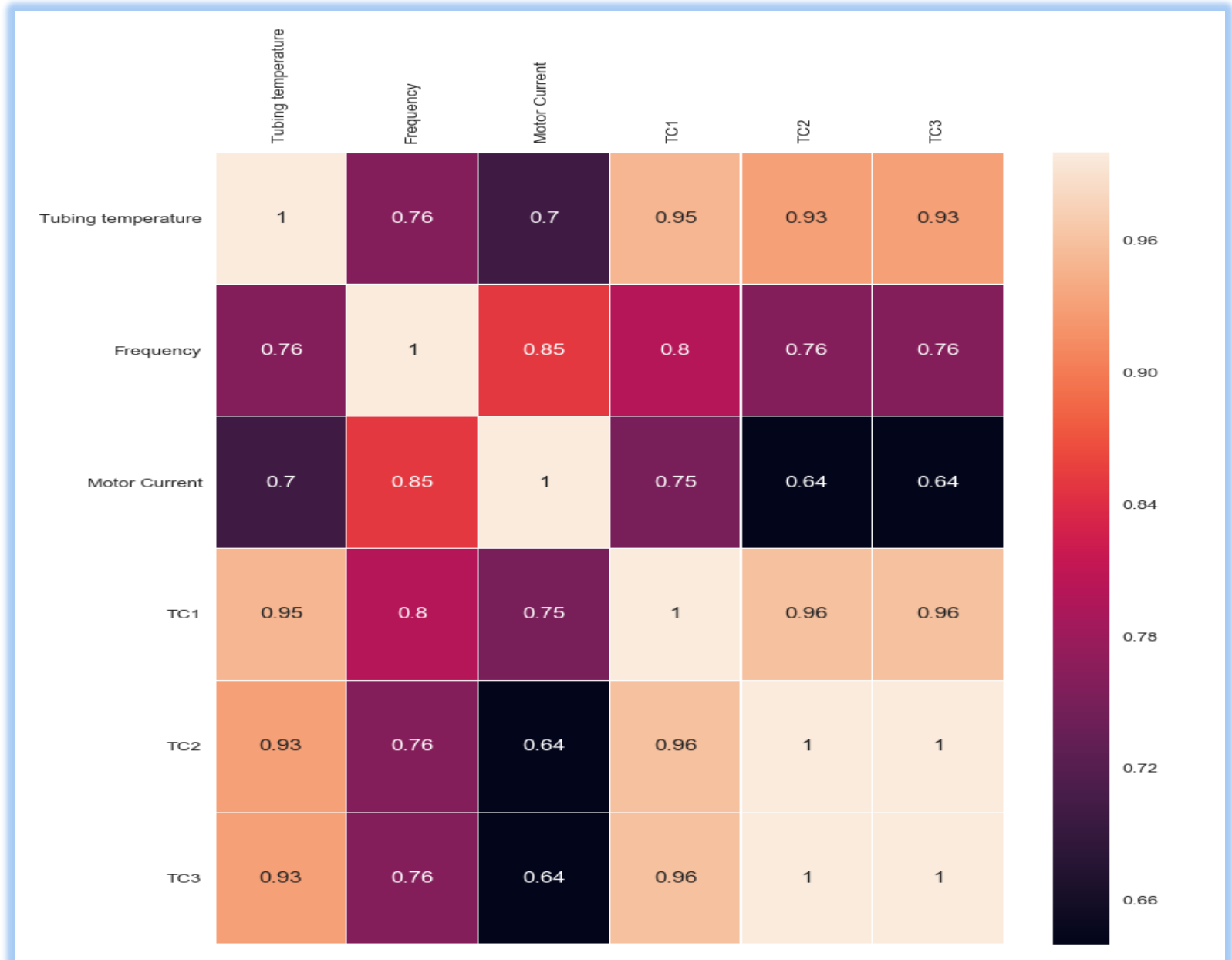


## Correlation analysis – Result

### 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

The screenshot shows the Predictive Monitoring Toolbox software interface. The 'Data management' tab is active. The 'Dataset' list contains 'ESP\_W03\_scaled' and 'ESP\_W03\_scaled\_RM', with the latter highlighted in blue. The 'Available variables' list includes 'Tubing temperature', 'Frequency', 'Motor Current', 'TC1', 'TC2', and 'TC3'. The 'Working data' section shows 'ESP\_W03\_scaled\_RM' selected. The 'Data Preprocess' section has a dropdown menu open, showing options like 'Remove missing data', 'Remove outlier', 'Filter', 'Normalize', 'Resample/Down sample', 'Select range', 'Correlation analysis', 'Mutual information analysis', and 'Define auxiliary tags', with the last option highlighted in blue. The 'Selected variables' list is currently empty. At the bottom, there are buttons for 'Import Data', 'Delete Data', 'Select variables from second list', 'Select as additional variables', 'Remove selected variables', 'Load selected variables', and 'Save selected variables'.

This provides different data preprocessing algorithms



# 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”**

Define auxiliary tags

Available variables

- Tubing temperature
- Frequency
- Motor Current
- TC1
- TC2
- TC3

Selected variables for creating new tag

- Tubing temperature(x(1))
- Frequency(x(2))
- Motor Current(x(3))
- TC1(x(4))
- TC2(x(5))
- TC3(x(6))

Enter the name of new tag:

Formula used to define the new tag:

Add the new variable

Select variables    Select all    Remove selected variable(s)

Include additional variable(s)

# 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

The screenshot shows a dialog box titled "Define auxiliary tags" with three main sections:

- Available variables:** A list containing "Tubing temperature", "Frequency", "Motor Current", "TC1", "TC2", and "TC3".
- Selected variables for creating new tag:** A list containing "Tubing temperature(x(1))", "Frequency(x(2))", "Motor Current(x(3))", "TC1(x(4))", "TC2(x(5))", and "TC3(x(6))".
- Enter the name of new tag:** A text input field containing "new\_variable\_1", which is highlighted with a red rectangular box.

Below the input field is a section labeled "Formula used to define the new tag:" with an empty text area. At the bottom of the dialog are four buttons: "Select variables", "Select all", "Remove selected variable(s)", and "Include additional variable(s)".

# 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

Define auxiliary tags

Available variables	Selected variables for creating new tag	Enter the name of new tag:
Tubing temperature Frequency Motor Current TC1 TC2 TC3	Tubing temperature(x(1)) Frequency(x(2)) Motor Current(x(3)) TC1(x(4)) TC2(x(5)) TC3(x(6))	new_variable_1

Formula used to define the new tag:

$x(4)/x(5)$

Add the new variable

Select variables    Select all    Remove selected variable(s)

Include additional variable(s)

# Predictive Monitoring Toolbox



## Data preprocessing – Defining auxiliary tags

6 Click the “Add the new variable” button to create this new variable

Define auxiliary tags

Available variables	Selected variables for creating
Tubing temperature	Tubing temperature(x(1))
Frequency	Frequency(x(2))
Motor Current	Motor Current(x(3))
TC1	TC1(x(4))
TC2	TC2(x(5))
TC3	TC3(x(6))

Select variables    Select all    Remove selected variable(s)

Include additional variable(s)

Add the new variable

Monitoring Toolbox Version 1.2

Data management    Defining events    Causal analysis    Offline analysis    Online time-domain analysis    Create Groups    Offline Analysis (Hierarchical)

Dataset: ESP\_W03\_scaled

Available variables: Tubing temperature, Frequency, Motor Current, TC1, TC2, TC3, new\_variable\_1

Working data: ESP\_W03\_scaled

Time Trend     Histogram

Data Preprocess: Define auxiliary tags

Remove selected variables

Load selected variables

Save selected variables

This provides different data preprocessing algorithms

# 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

The screenshot shows the 'Monitoring Toolbox Version 1.2' interface. The 'Available variables' list is highlighted with a red box and contains: Tubing temperature, Frequency, Motor Current, TC1, TC2, TC3, and new\_variable\_1. Below this list, the 'Select variables from second list' button is also highlighted with a red box. A blue line connects this button to the 'Available variables' list. A green arrow points from the 'Select variables from second list' button to the 'Selected variables' list in the next screenshot.

From the second list, choose one variable and click this button

The screenshot shows the 'Monitoring Toolbox Version 1.2' interface. The 'Selected variables' list is highlighted with a red box and contains: Tubing temperature, Frequency, Motor Current, TC1, TC2, and TC3. Below this list, the 'Save selected variables' button is also highlighted with a red box. A green arrow points from the 'Select variables from second list' button in the previous screenshot to the 'Selected variables' list.

From the second list, choose one variable and click this button

2

### Save selected variables

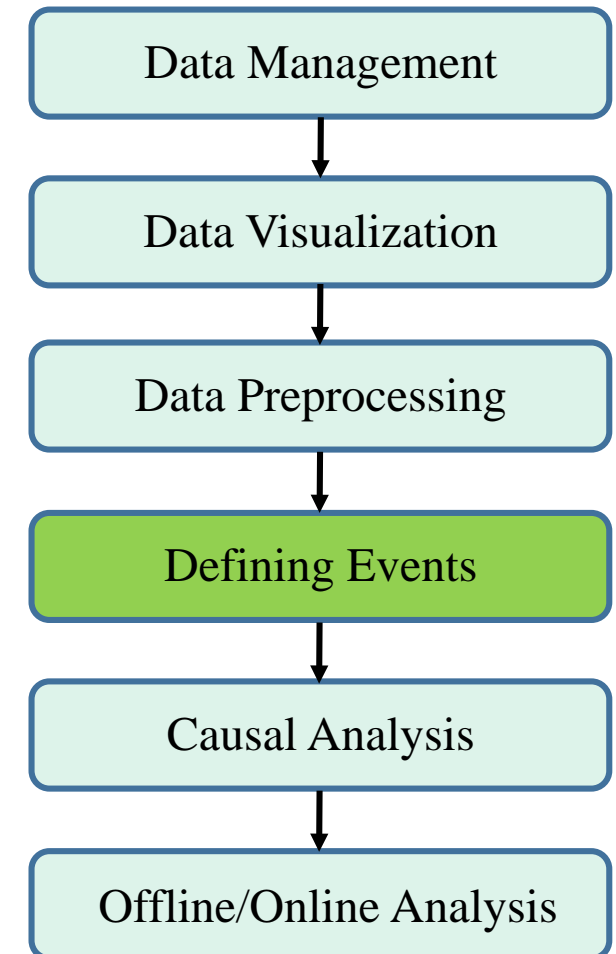
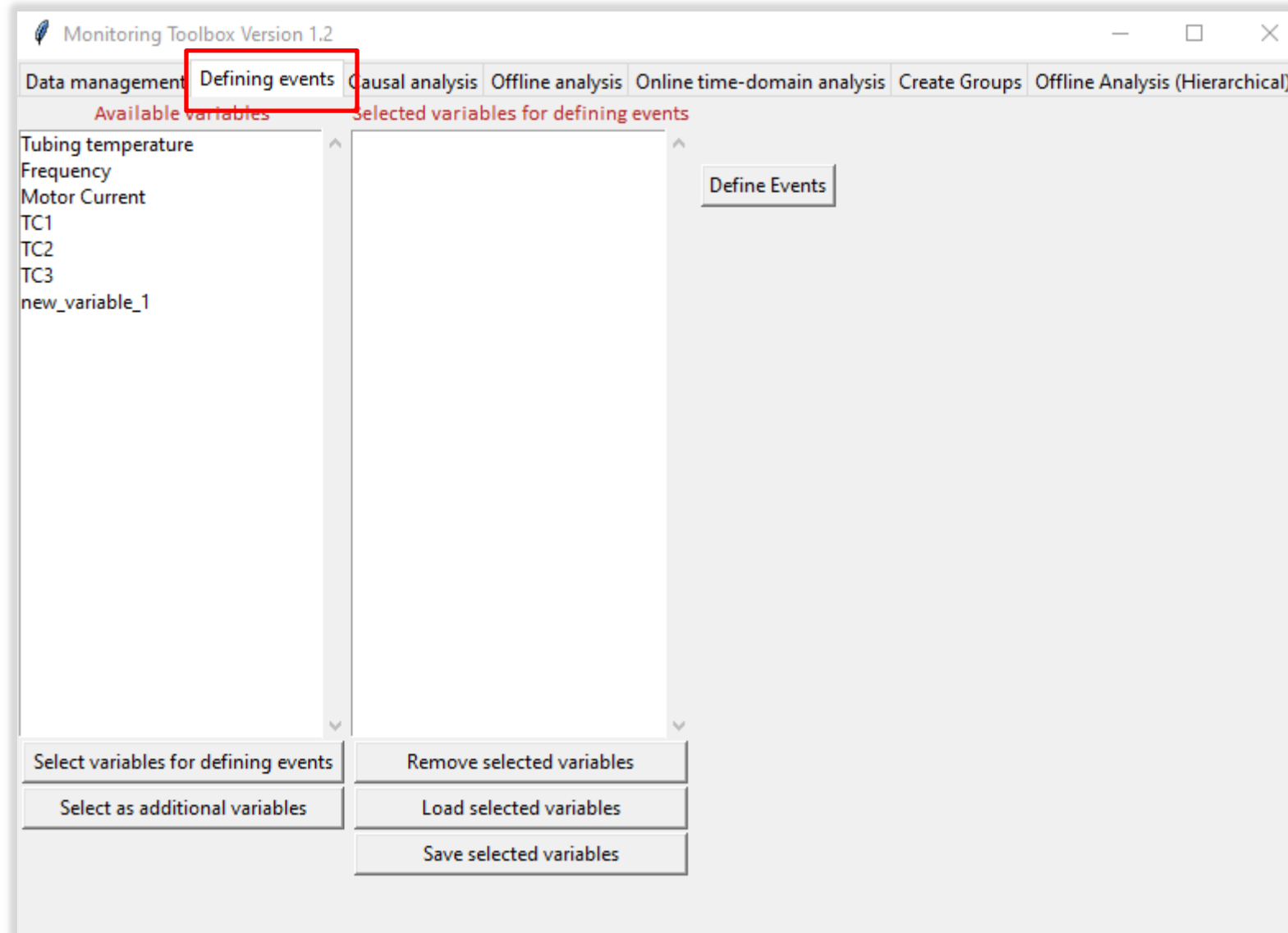
- The variables selected are shown in the third column
- Click “Save selected variables” button to save this list of variables for further analysis

# 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

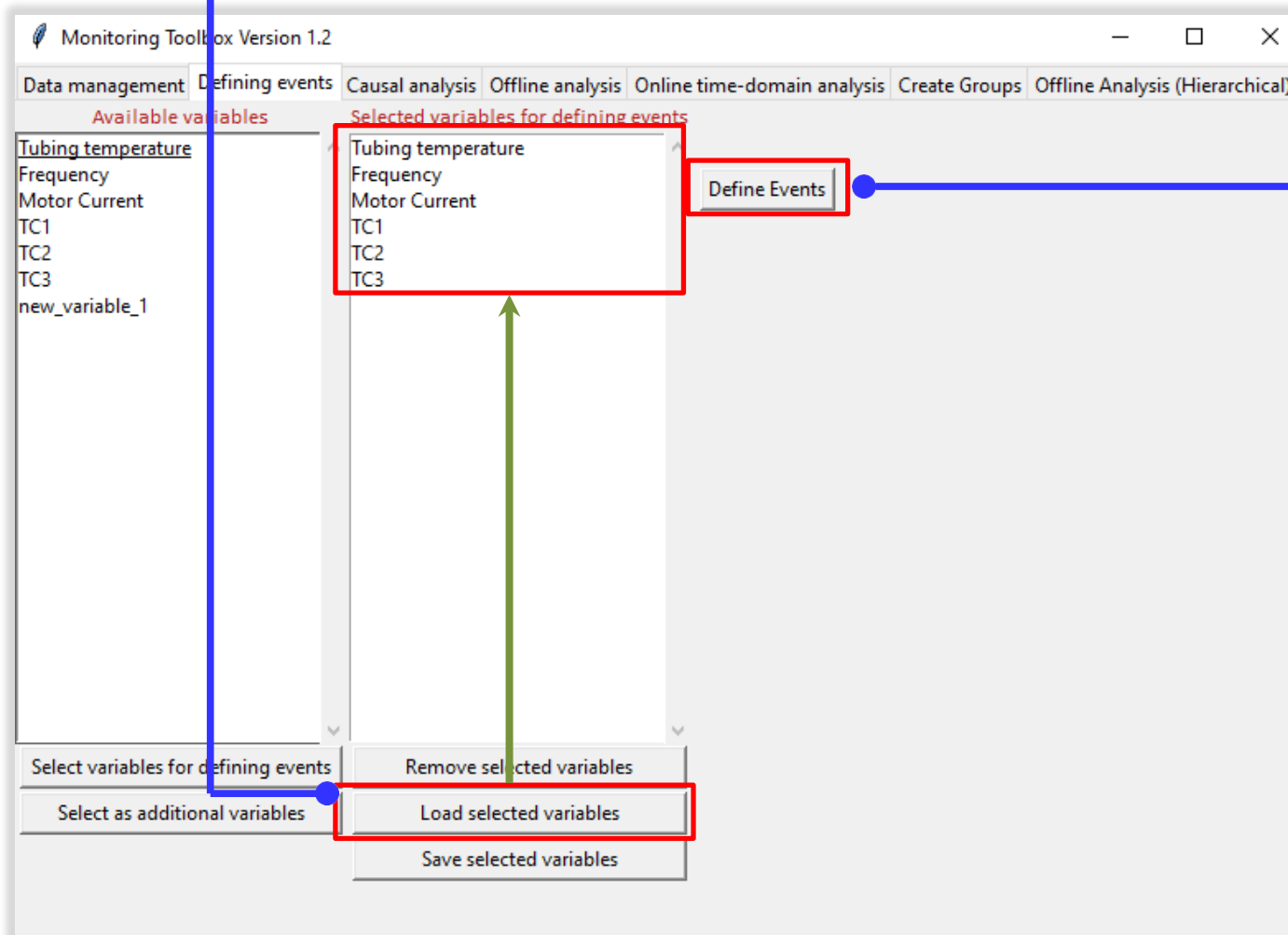
Under the “**Defining events**” tab, different types of events that took place can be defined

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

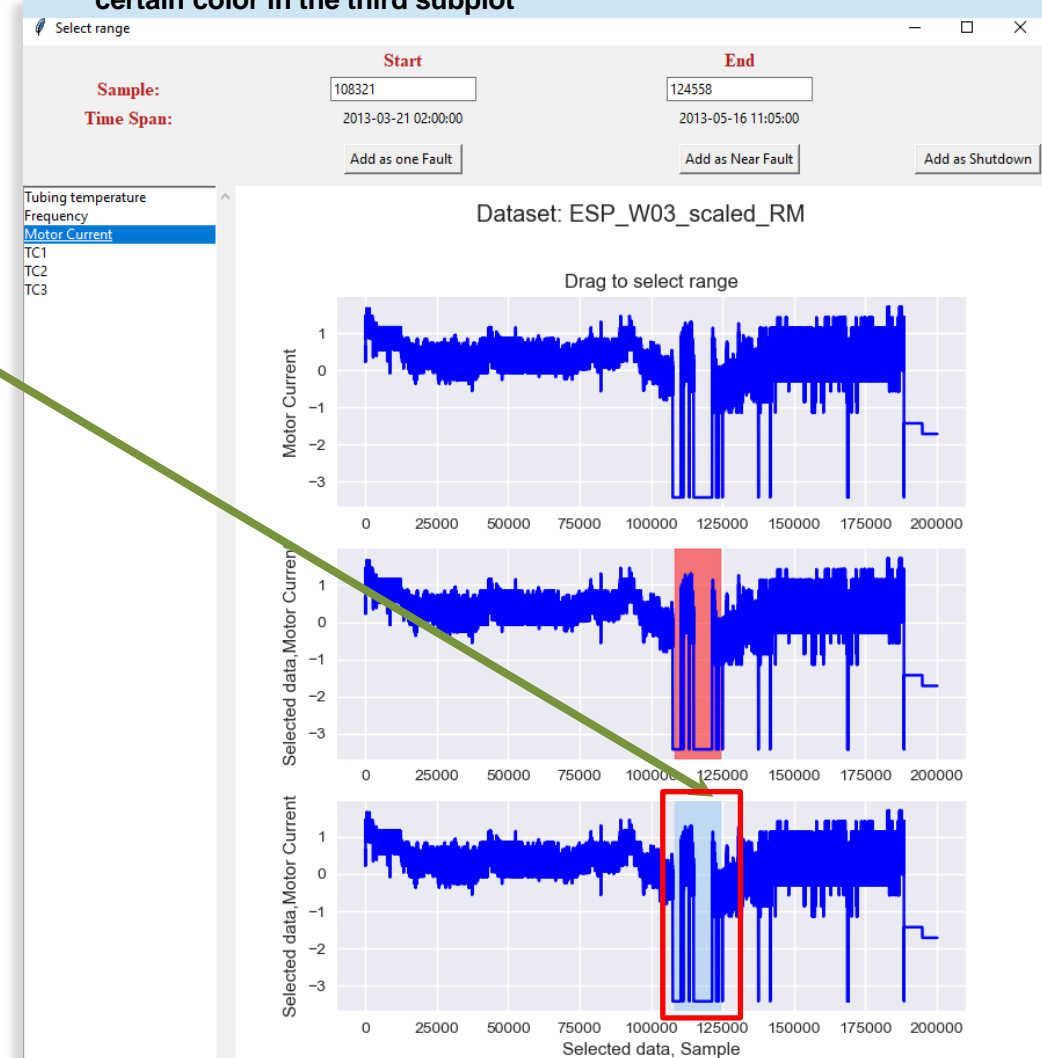
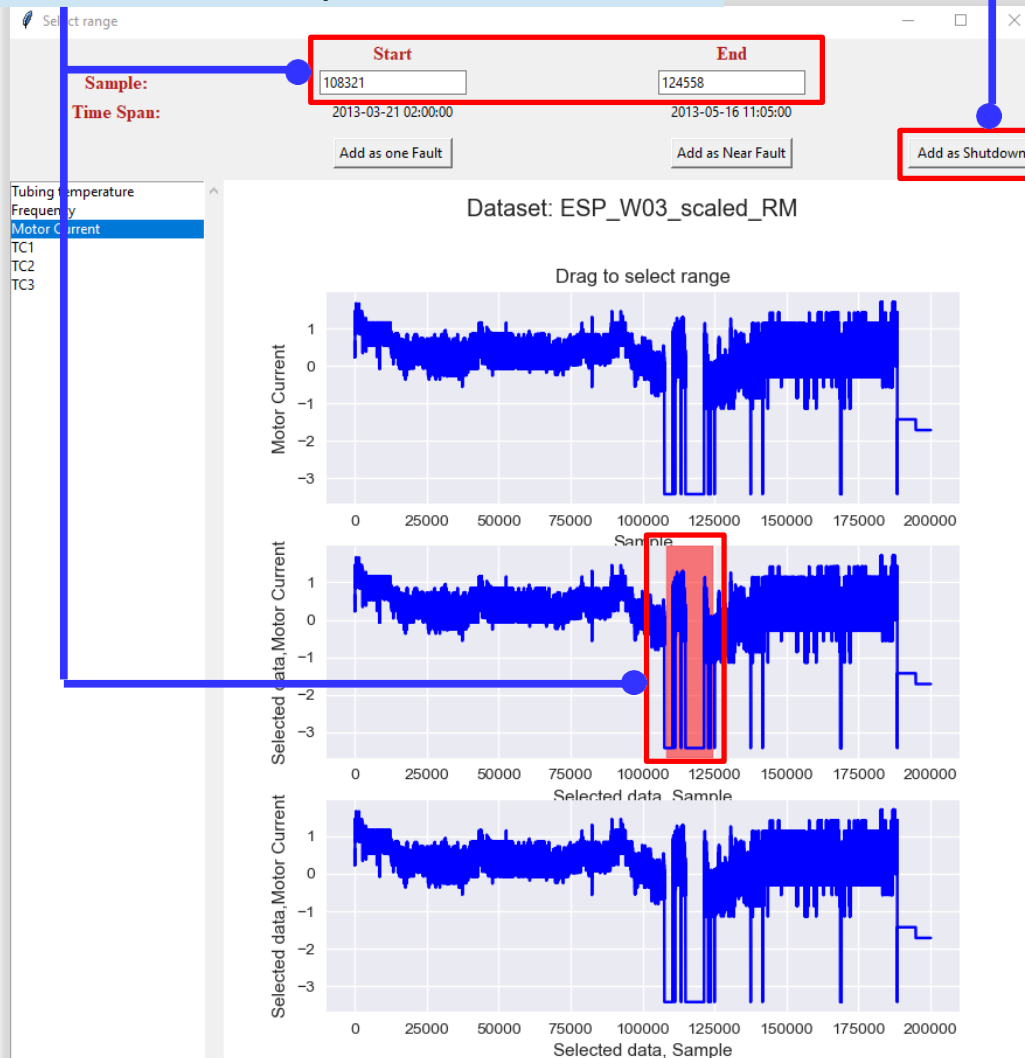


## Defining Events

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”, “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

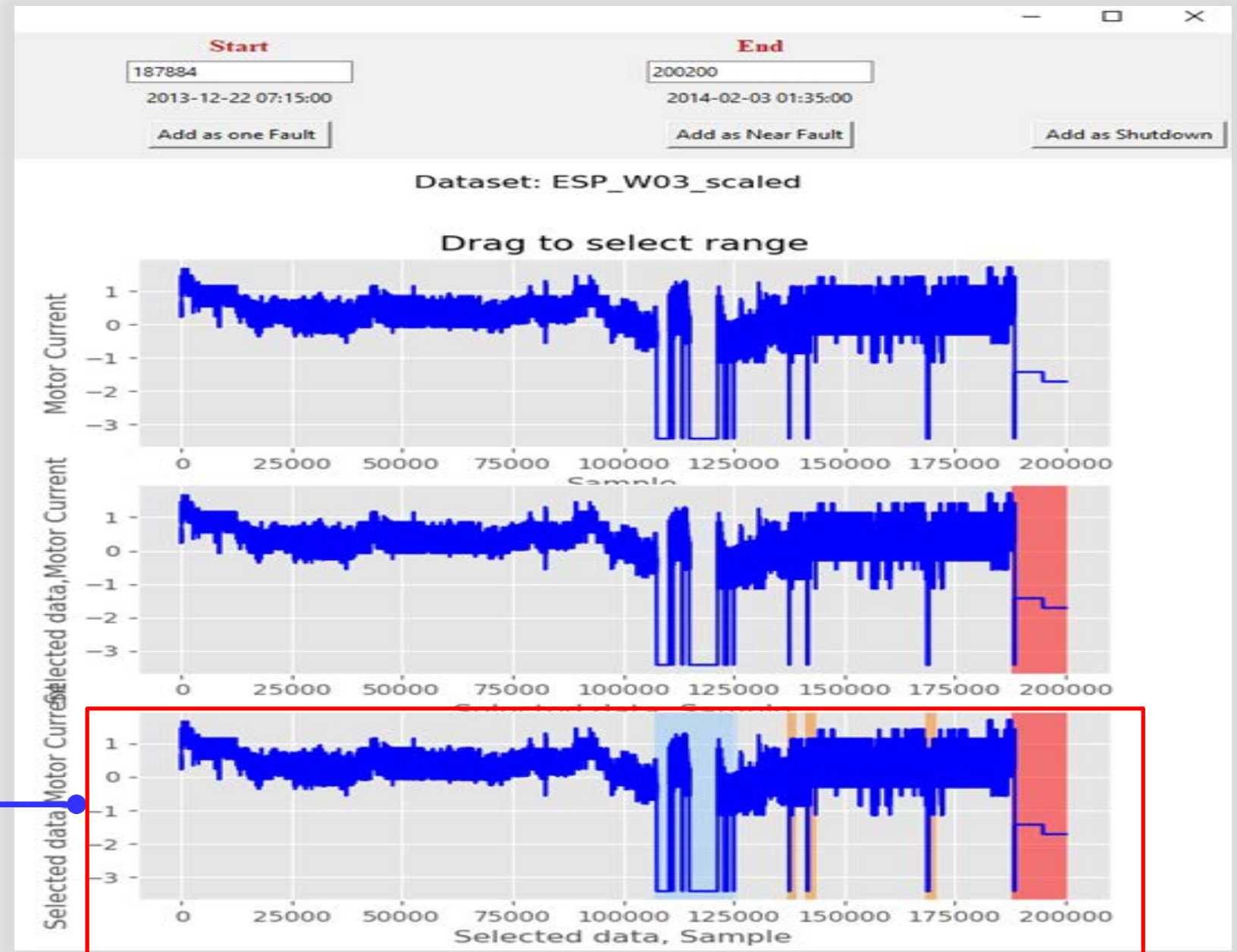


## Defining Events

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

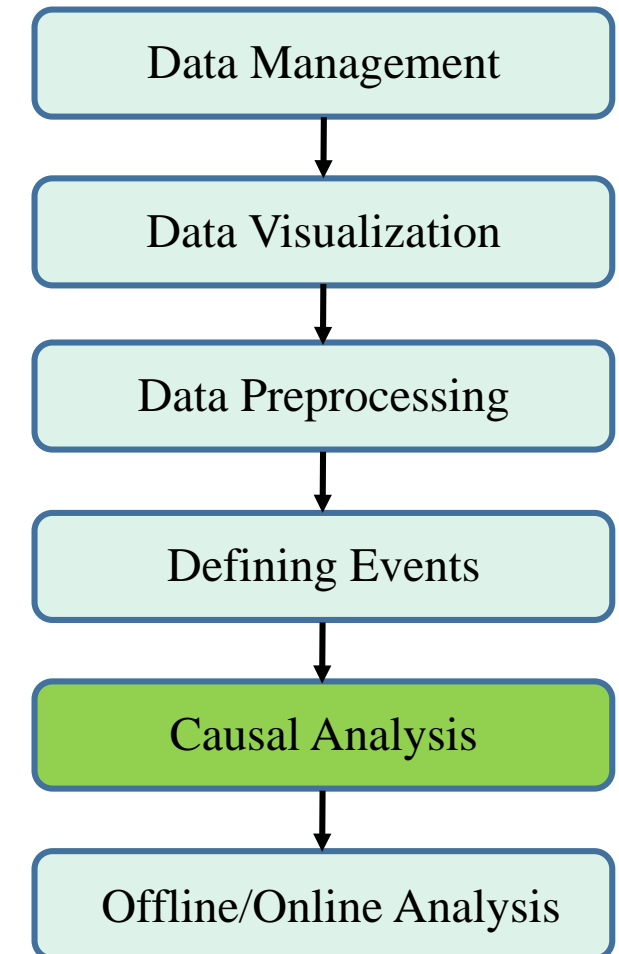
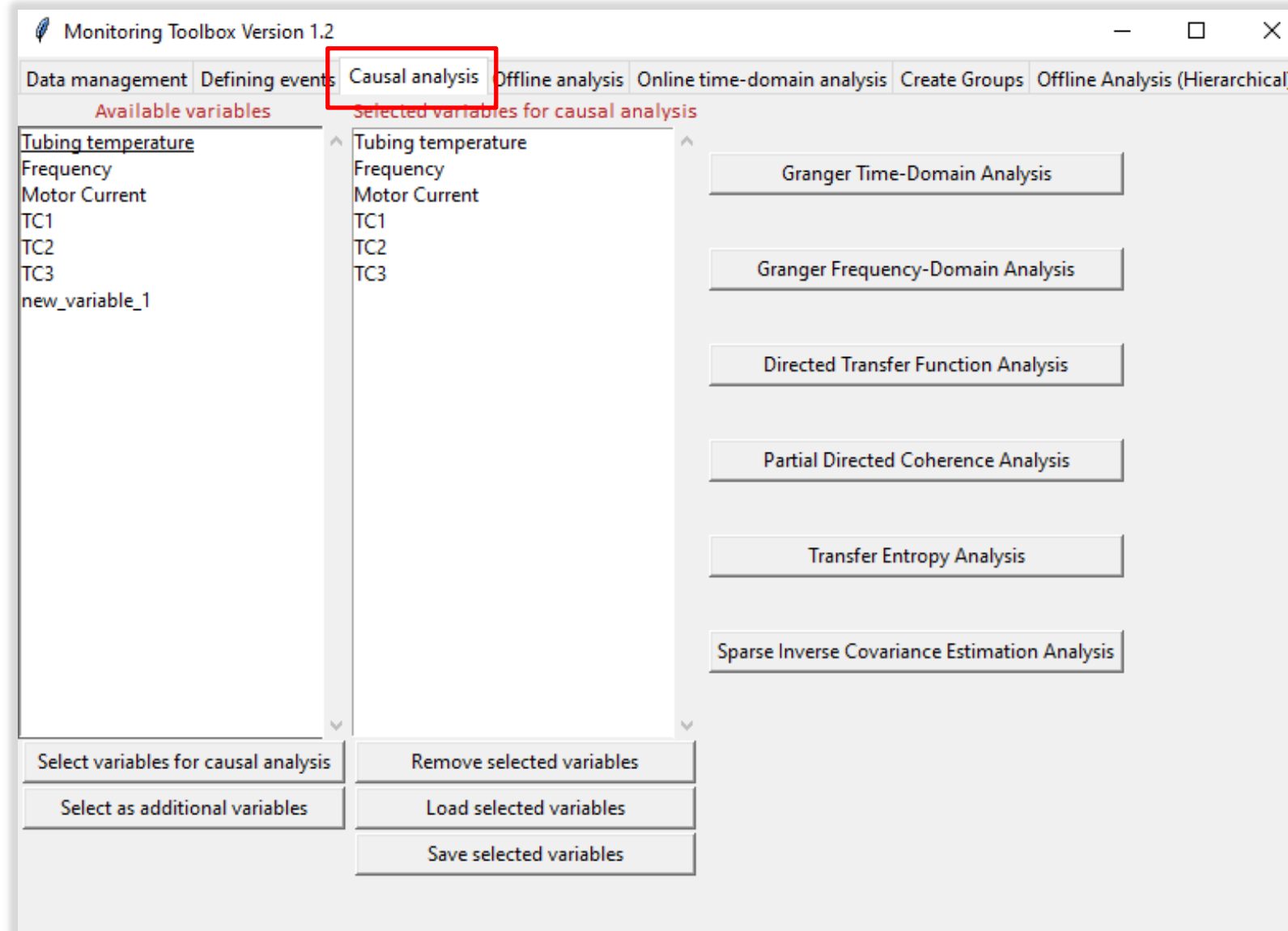


# Predictive Monitoring Toolbox



## 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

The screenshot displays the "Monitoring Toolbox Version 1.2" software interface. The "Causal analysis" tab is selected and highlighted with a red box. A blue callout box with the number "1" and the text "Go to 'Causal analysis' tab" points to the selected tab. The interface is divided into three main sections:

- Available variables:** A list of variables including Tubing temperature, Frequency, Motor Current, TC1, TC2, TC3, and new\_variable\_1.
- Selected variables for causal analysis:** A list of variables that have been selected for analysis, including Tubing temperature, Frequency, Motor Current, TC1, TC2, and TC3.
- Analysis Algorithms:** A vertical stack of six buttons for different causal analysis methods: Granger Time-Domain Analysis, Granger Frequency-Domain Analysis, Directed Transfer Function Analysis, Partial Directed Coherence Analysis, Transfer Entropy Analysis, and Sparse Inverse Covariance Estimation Analysis.

At the bottom of the interface, there are several control buttons: "Select variables for causal analysis", "Remove selected variables", "Select as additional variables", "Load selected variables", and "Save selected 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

Monitoring Toolbox Version 1.2

Data management | Defining events | **Causal analysis** | Offline analysis | Online time-domain analysis | Create Groups | Offline Analysis (Hierarchical)

Available variables: Tubing temperature, Frequency, Motor Current, TC1, TC2, TC3, new\_variable\_1

Selected variables for causal analysis: Tubing temperature, Frequency, Motor Current, TC1, TC2, TC3

Buttons: Granger Time-Domain Analysis, **Granger Frequency-Domain Analysis**, Directed Transfer Function Analysis, Partial Directed Coherence Analysis, Transfer Entropy Analysis, Sparse Inverse Covariance Estimation Analysis

Bottom buttons: Select variables for causal analysis, Remove selected variables, Select as additional variables, Load selected variables, Save selected variables

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

Granger frequency-domain analysis

Variables used in analysis

Frequency: 256

Max order: 20

Variables used in analysis: Tubing temperature, Frequency, Motor Current, TC1, TC2, TC3

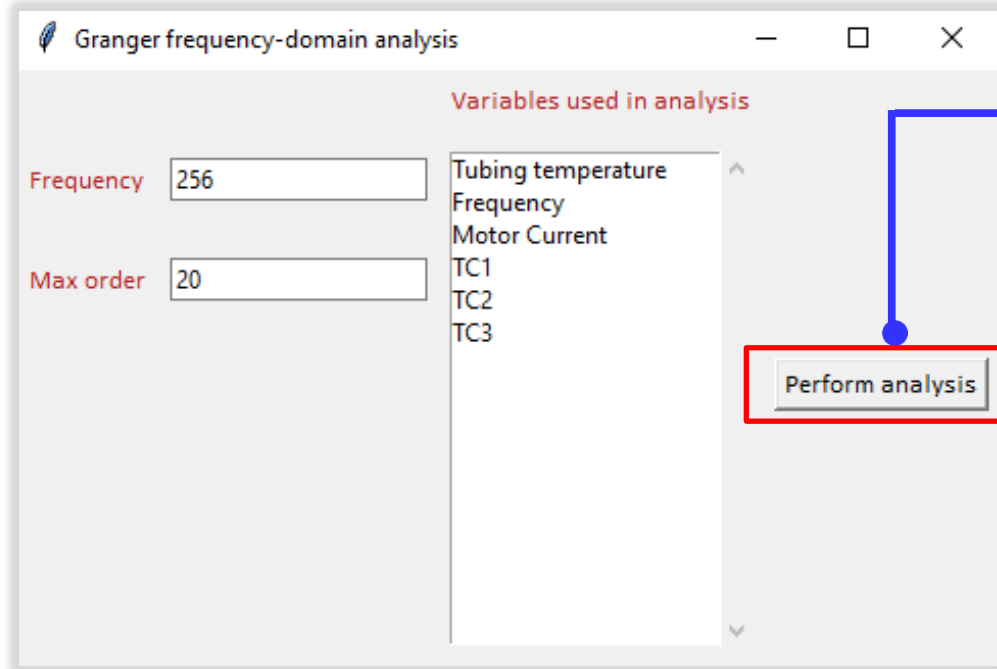
Perform analysis

# Predictive Monitoring Toolbox

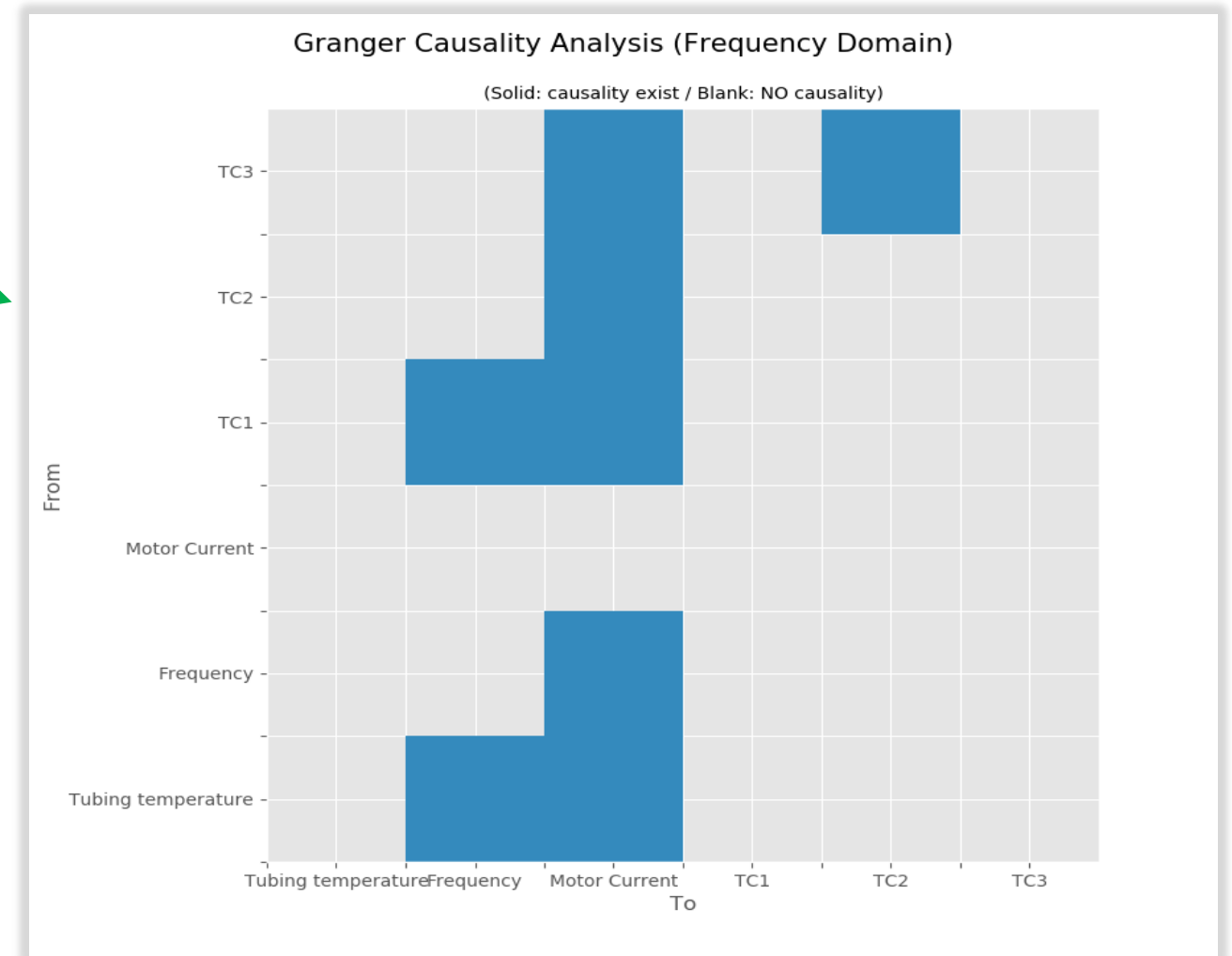


## Causal analysis

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



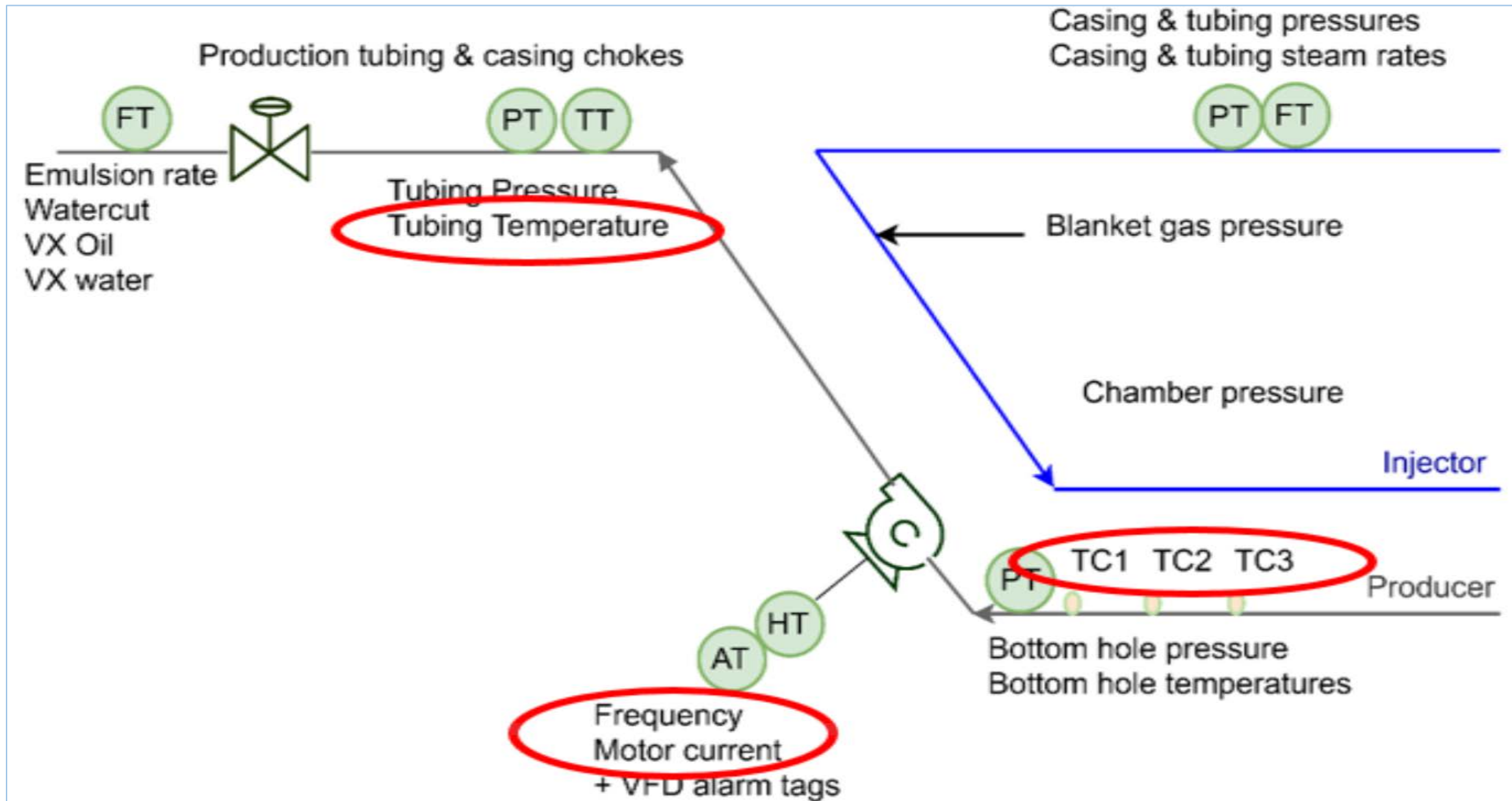
3 Click the “Granger Frequency-Domain Analysis” button, the results will be shown



# Predictive Monitoring Toolbox



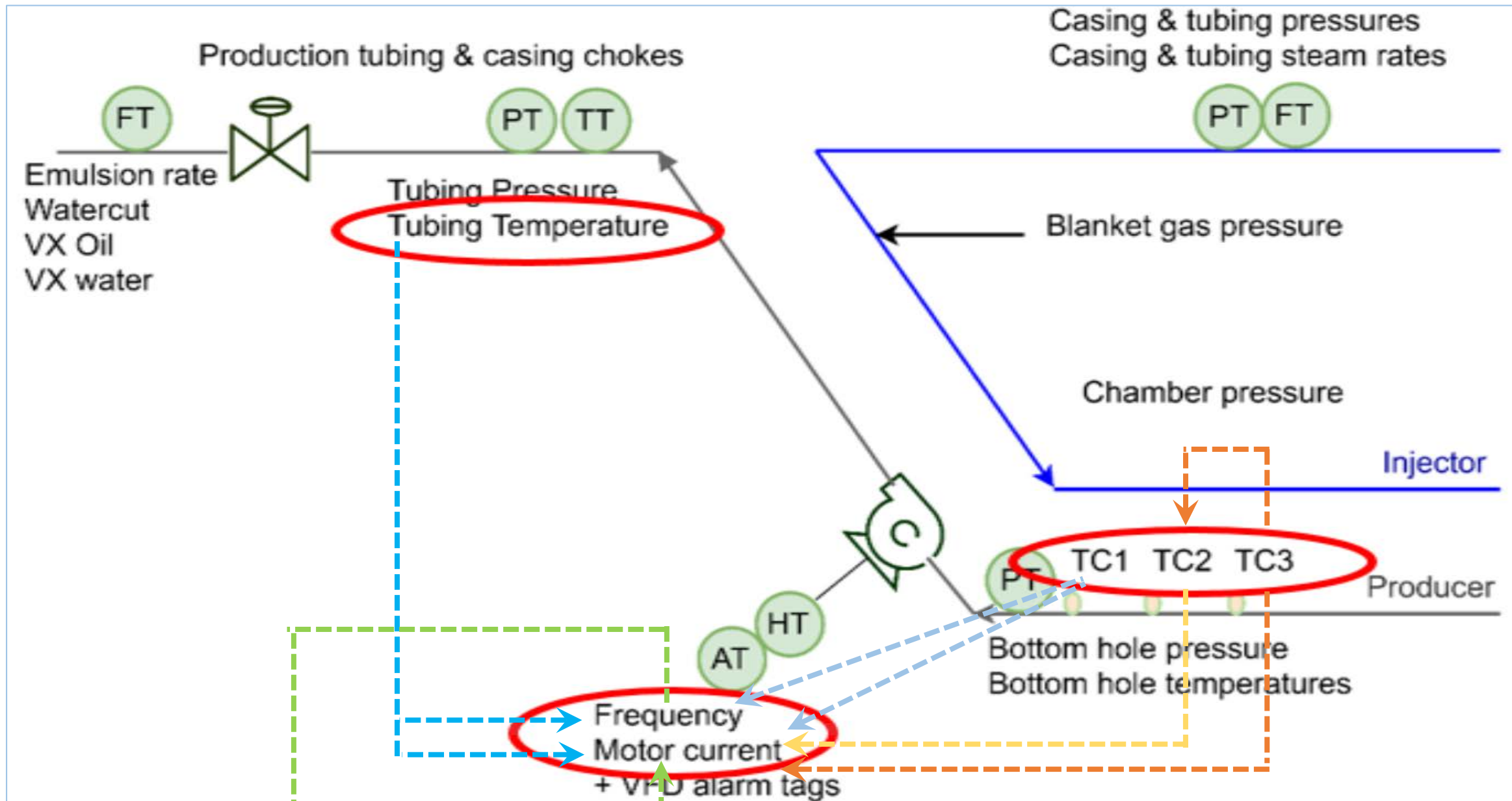
Causal analysis – interpreting the results: **Each dashed line indicates a cause-effect relationship**



# Predictive Monitoring Toolbox



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

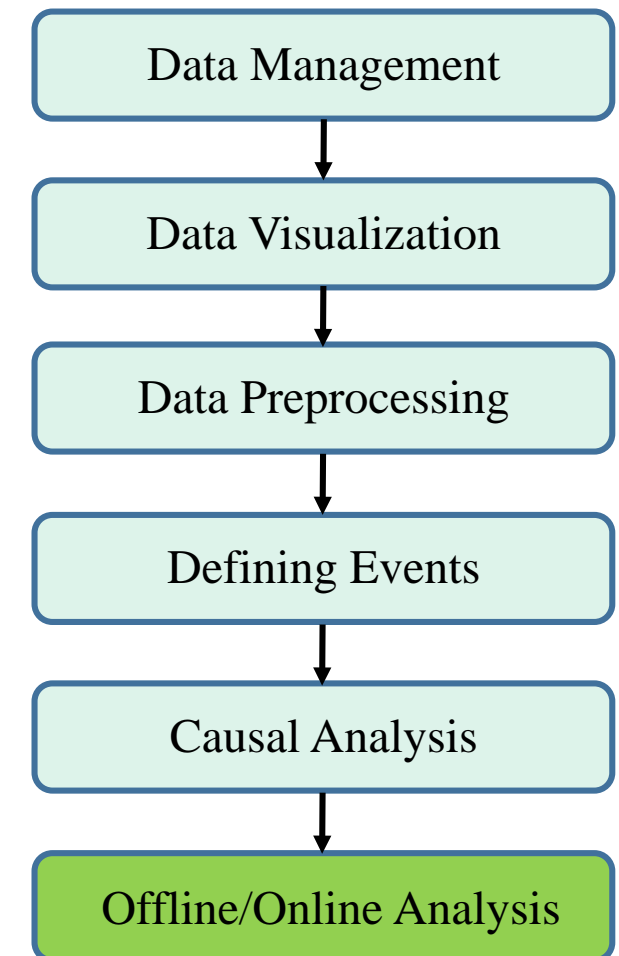
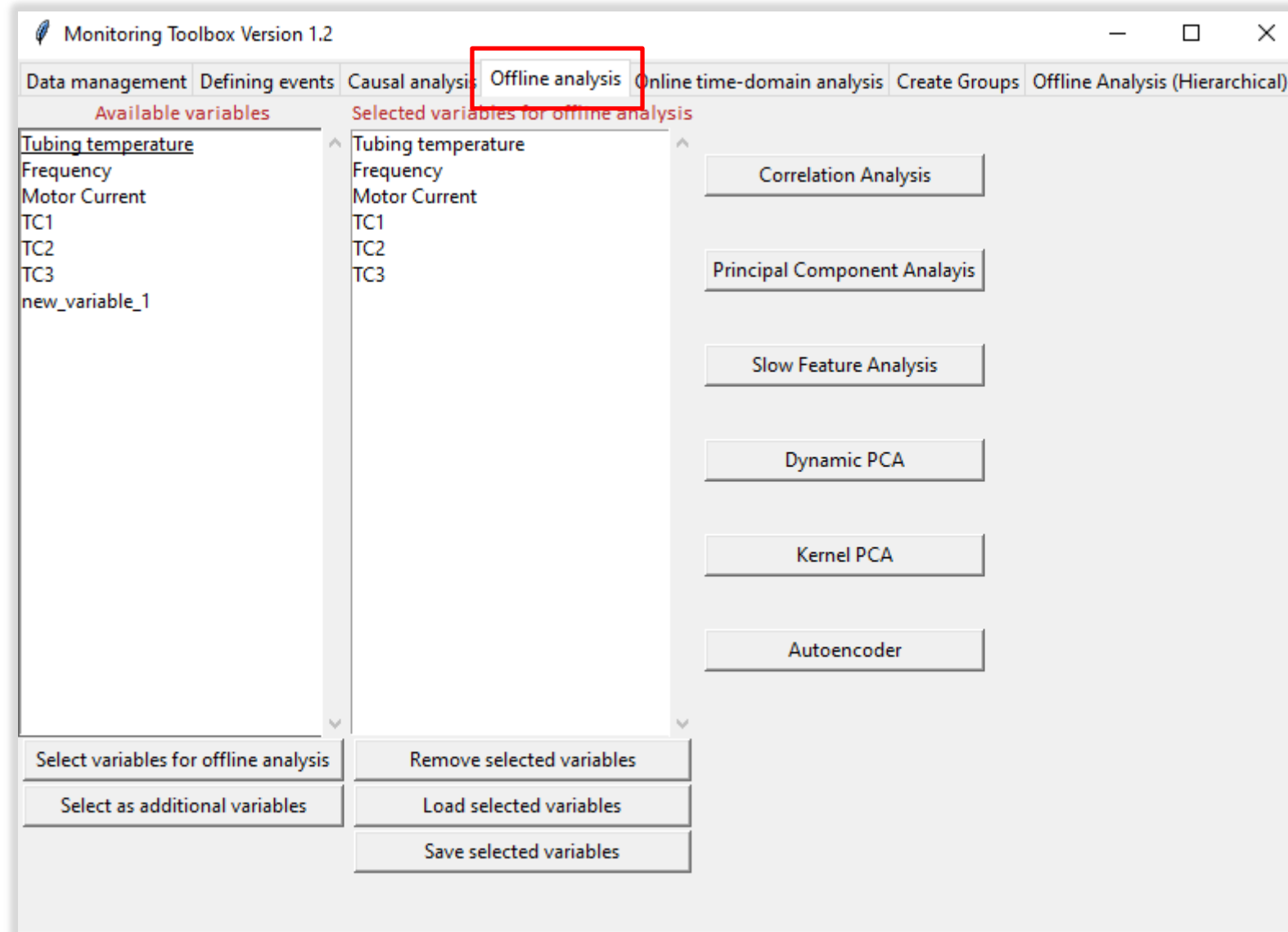


# Predictive Monitoring Toolbox



## 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

The screenshot displays the 'Monitoring Toolbox Version 1.2' interface. The 'Offline analysis' tab is active, showing a list of 'Available variables' (Tubing temperature, Frequency, Motor Current, TC1, TC2, TC3, new\_variable\_1) and 'Selected variables for offline analysis' (Tubing temperature, Frequency, Motor Current, TC1, TC2, TC3). A red box highlights the 'Principal Component Analysis' button. A blue callout box with the number '1' points to this button, containing the text: 'Click "Principal component analysis" button' and 'A new window pops up, and principal component analysis can be conducted'. A green arrow points from the 'Principal Component Analysis' button to a separate window titled 'Principal Component Analysis'. This window contains the following settings:

- Number of Components: 3
- Control Limit Confidence (%): 95
- Variables considered in PCA: Tubing temperature, Frequency, Motor Current, TC1, TC2, TC3
- Dimension of visualization: 1
- PC indices: 1, 2, 3
- Buttons: Scree Plot, T Squared Score, Q-Statistic, Visualize Scores

# 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.

Principal Component Analysis

Variables considered in PCA

Number of Components: 3

Control Limit Confidence (%): 95

Scree Plot

T Squared Score

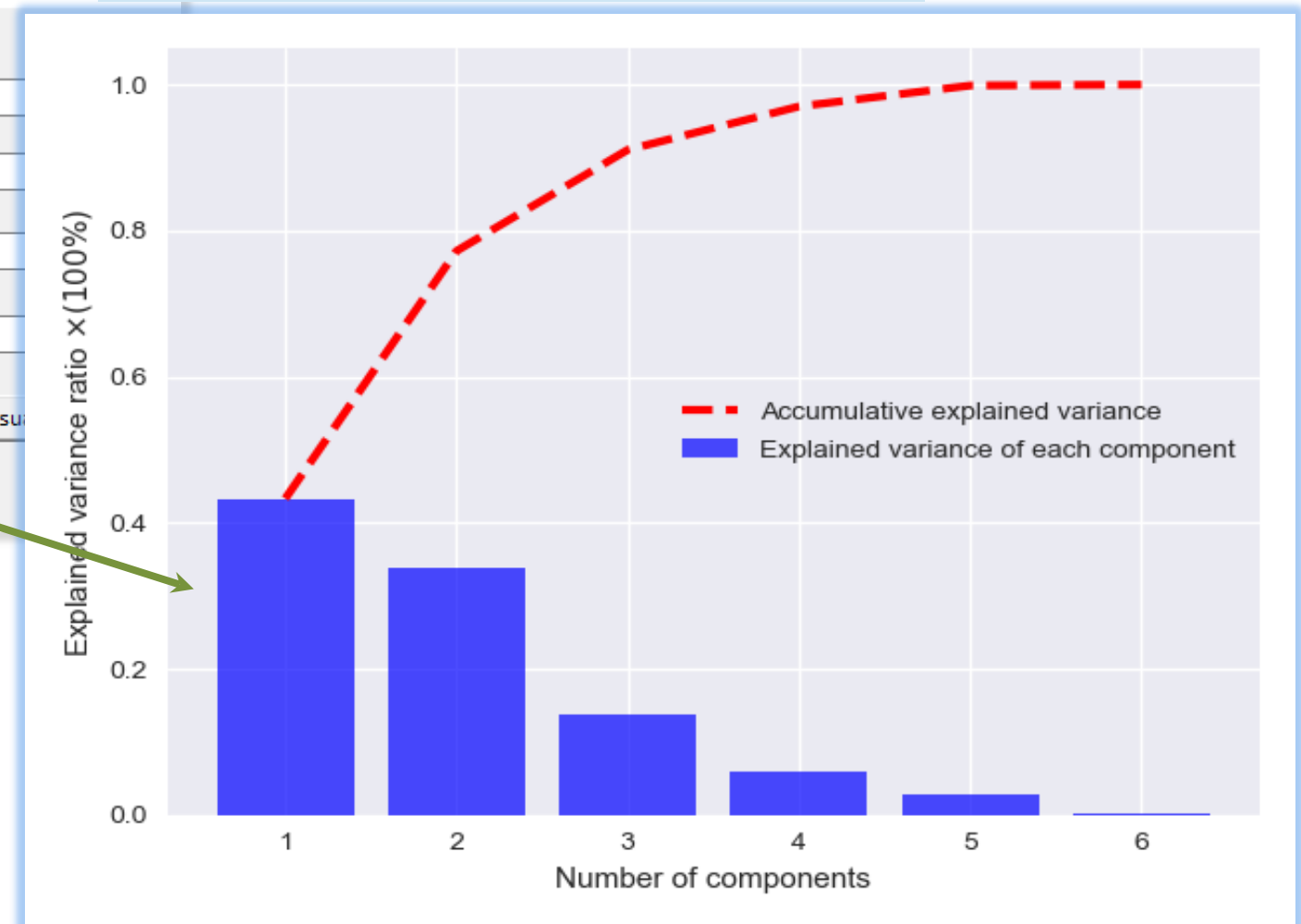
Q-Statistic

Variables considered in PCA:

- Tubing temperature
- Frequency
- Motor Current
- TC1
- TC2
- TC3

Dimension of visualization: 1

PC indices: 1, 2, 3



# 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.

The software interface shows the following settings:

- Number of Components: 4
- Control Limit Confidence (%): 95
- Dimension of visualization: 1
- PC indices: 1, 2, 3

Variables considered in PCA:

- Tubing temperature
- Frequency
- Motor Current
- TC1
- TC2
- TC3

Buttons: Scree Plot, T Squared Score, Q-Statistic

The plot shows the T<sup>2</sup> Score based on PCA (Y-axis, 0.00 to 0.12) versus Selected data, Sample (X-axis, 0 to 200,000). The plot displays a blue line representing the T-squared score over time. A horizontal dashed red line indicates the control limit. Vertical lines indicate detected events: a blue shaded region for 'shutdown' (around sample 110,000), orange vertical lines for 'near fault', and a red shaded region for 'fault' (around sample 190,000).

# 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

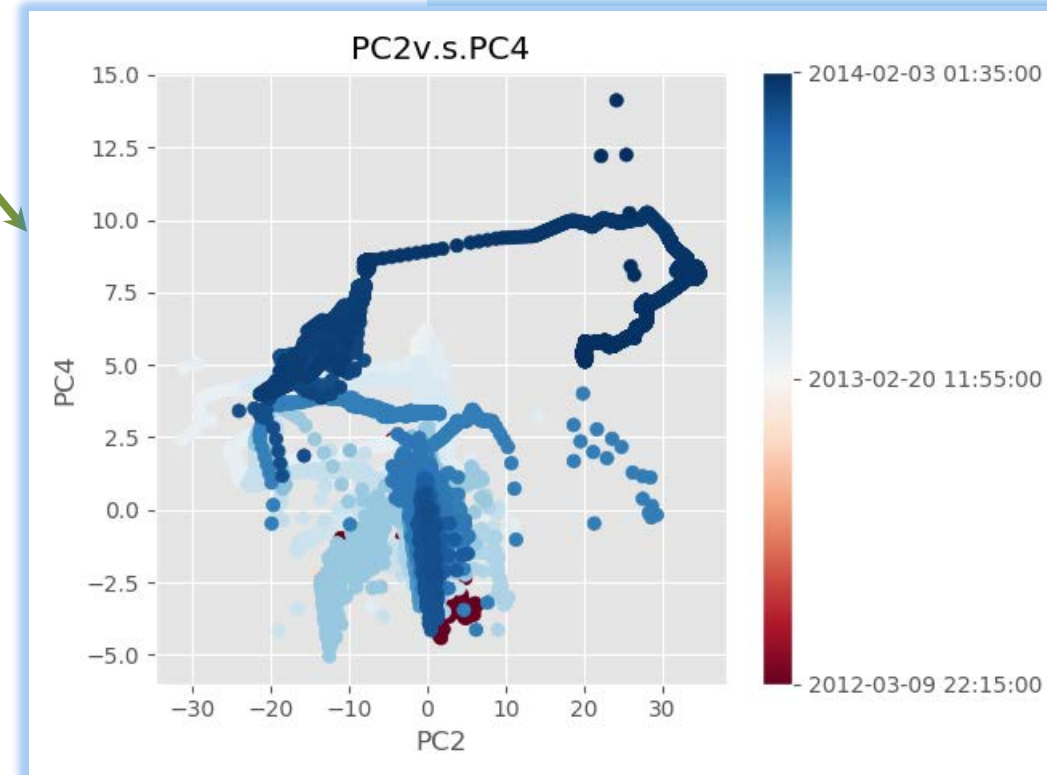


## Offline monitoring using PCA – Visualizing principal components of PCA (2D)

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)

Principal Component Analysis

Number of Components: 4

Control Limit Confidence (%): 95

Variables considered in PCA:

- Tubing temperature
- Frequency
- Motor Current
- TC1
- TC2
- TC3

Dimension of visualization: 3

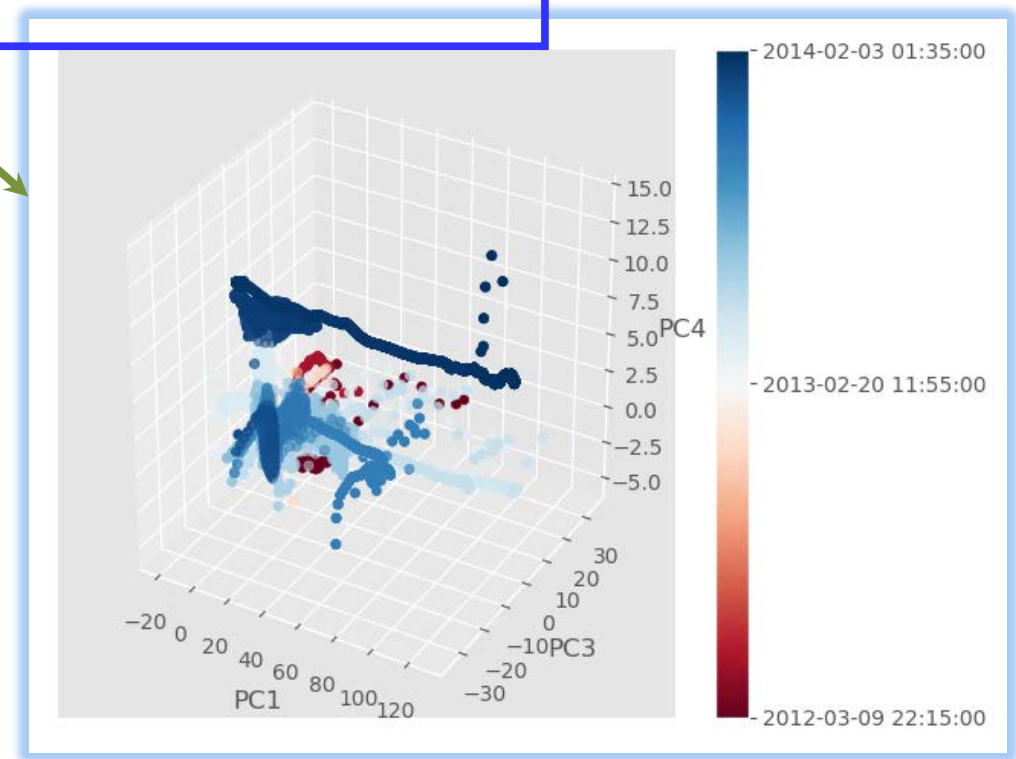
PC indices: 1, 3, 4

Buttons: Scree Plot, T Squared Score, Q-Statistic, Visualize Scores

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

The screenshot displays the 'Monitoring Toolbox Version 1.2' interface. The 'Offline analysis' tab is active, showing a list of available variables on the left and selected variables for offline analysis in the middle. The 'Slow Feature Analysis' button is highlighted with a red box. A blue callout box with the number '1' points to this button, containing the text: 'Click "Slow Feature Analysis" button' and 'A new window pops up, and principal component analysis can be conducted'. A green arrow points from the 'Slow Feature Analysis' button to a separate window titled 'Slow Feature Analysis'. This window contains the following settings:

- Number of Features: 5
- Control Limit Confidence (%): 95
- Variables considered in SFA: Tubing temperature, Frequency, Motor Current, TC1, TC2, TC3
- Dimension of visualization: 1
- Feature indices: 1, 2, 3

Buttons for 'Slowness Assessment', 'T Squared Score', and 'Visualize Scores' are also visible in the 'Slow Feature Analysis' window.

# 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.

Slow Feature Analysis

Variables considered in SFA

Number of Features: 5

Control Limit Confidence (%): 95

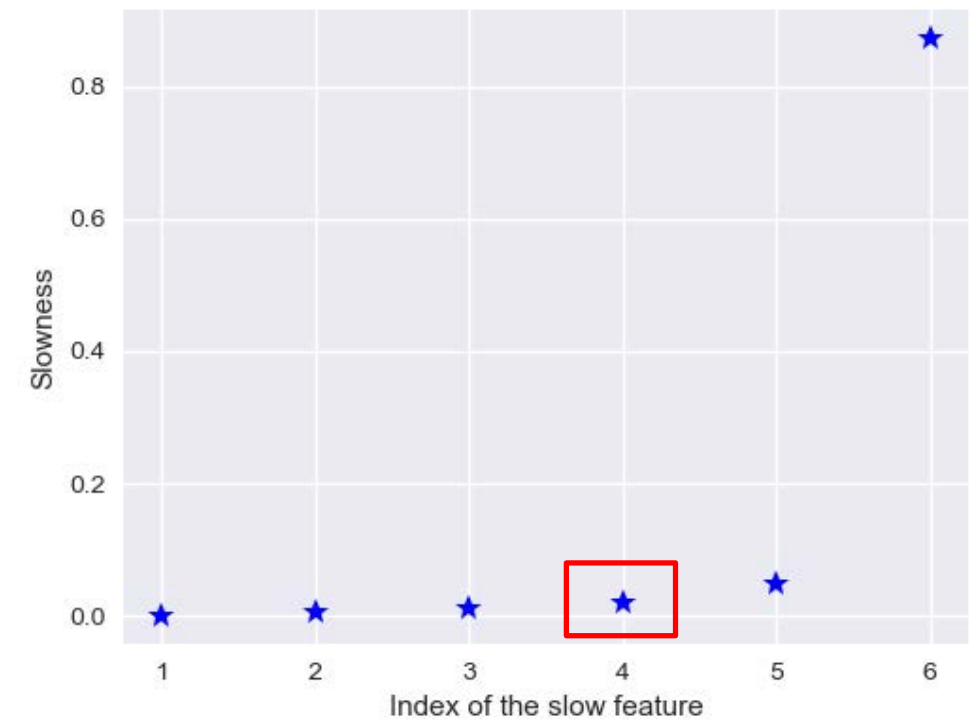
T Squared Score

Dimension of visualization: 1

Feature indices: 1, 2, 3

Visualize Scores

Variables in list: Tubing temperature, Frequency, Motor Current, TC1, TC2, TC3





# Predictive Monitoring Toolbox



## Offline monitoring – Slow feature analysis (SFA)

The screenshot shows the 'Slow Feature Analysis' window. On the left, there are input fields for 'Number of Features' (set to 4), 'Control Limit Confidence (%)' (set to 99), and 'T Squared Score'. A 'Slowness Assessment' button is located below the 'Number of Features' field. In the center, a list of 'Variables considered in SFA' includes Tubing temperature, Frequency, Motor Current, TC1, TC2, and TC3. On the right, there are input fields for 'Dimension of visualization' (set to 1), 'Feature indices' (set to 1), and another field (set to 2). A 'Visualize Scores' button is at the bottom right. Two callout boxes with blue circles containing the numbers 3 and 4 provide instructions: callout 3 points to the 'Number of Features' field with the text 'Enter "4" as the number of components.', and callout 4 points to the 'Control Limit Confidence (%)' field with the text 'Enter "99" so that the control limit confidence is made 99%.'

# Predictive Monitoring Toolbox



## Offline monitoring – Slow feature analysis (SFA)

Slow Feature Analysis

Variables considered in SFA

Number of Features: 4

Control Limit Confidence (%): 99

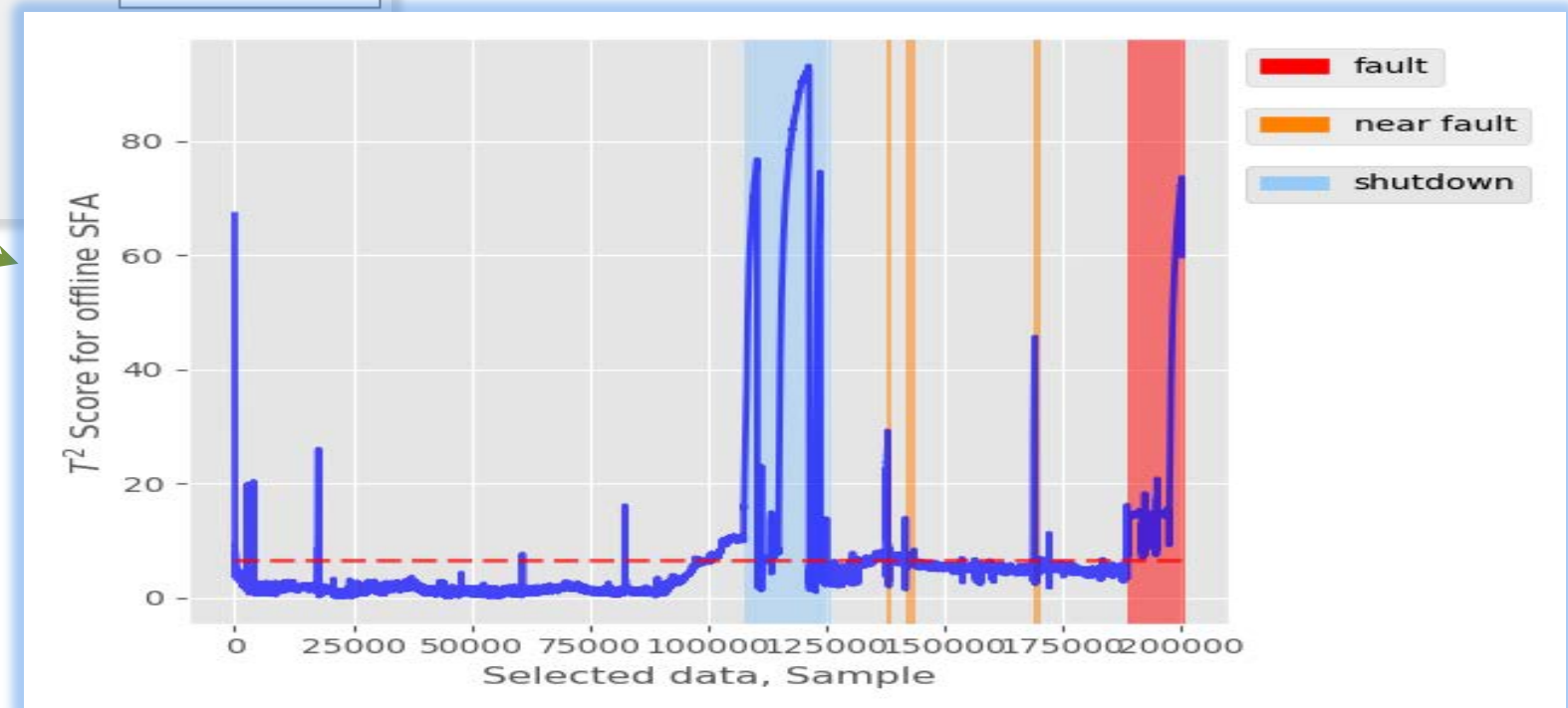
Dimension of visualization: 1

Feature indices: 1, 2, 3

Variables: Tubing temperature, Frequency, Motor Current, TC1, TC2, TC3

Buttons: Slowness Assessment, T Squared Score

- 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.

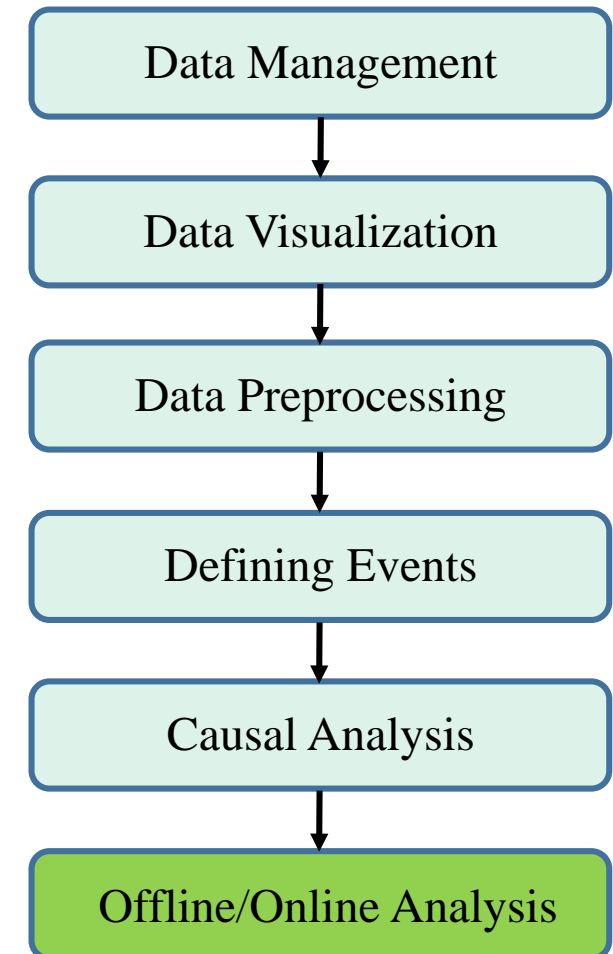
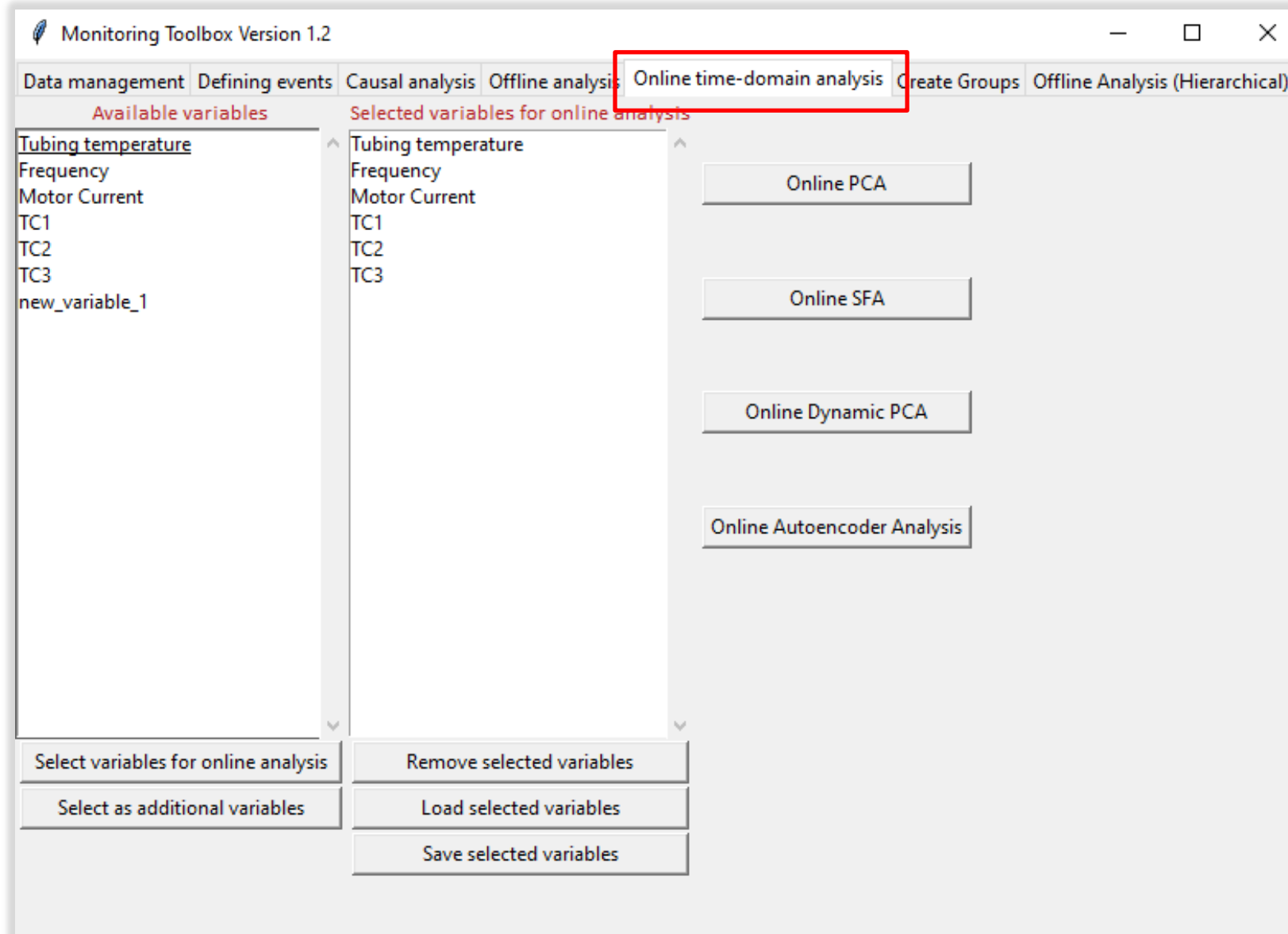


# Predictive Monitoring Toolbox



## 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

1

Click “Online PCA” button

- A new window pops up, and principal component analysis can be conducted

The screenshot displays the Monitoring Toolbox Version 1.2 interface. The main window has a menu bar with options: Data management, Defining events, Causal analysis, Offline analysis, Online time-domain analysis, Create Groups, and Offline Analysis (Hie). Below the menu bar, there are two columns of variables: 'Available variables' and 'Selected variables for online analysis'. The 'Available variables' list includes Tubing temperature, Frequency, Motor Current, TC1, TC2, TC3, and new\_variable\_1. The 'Selected variables for online analysis' list includes Tubing temperature, Frequency, Motor Current, TC1, TC2, and TC3. A red box highlights the 'Online PCA' button in the 'Selected variables for online analysis' column. A blue arrow points from this button to a callout box containing the instruction 'Click “Online PCA” button' and a bullet point: '• A new window pops up, and principal component analysis can be conducted'. A green arrow points from the 'Online PCA' button to a secondary window titled 'Online Principal Component Analysis'. This window contains several input fields: 'Number of Components' (3), 'Control Limit Confidence (%)' (95), and 'Update Window (Samples)' (2000). It also has a list of 'Variables considered in online PCA' (Tubing temperature, Frequency, Motor Current, TC1, TC2, TC3) and a 'Dimension of visualization' field (1). Below these fields are buttons for 'Online PCA analysis', 'T Squared Score', and 'Q-Statistic'. On the right side of the window, there are 'PC indices' fields (1, 2, 3) and a 'Visualize Scores' button.

# Predictive Monitoring Toolbox



## Online monitoring – Principal component analysis (PCA)

2

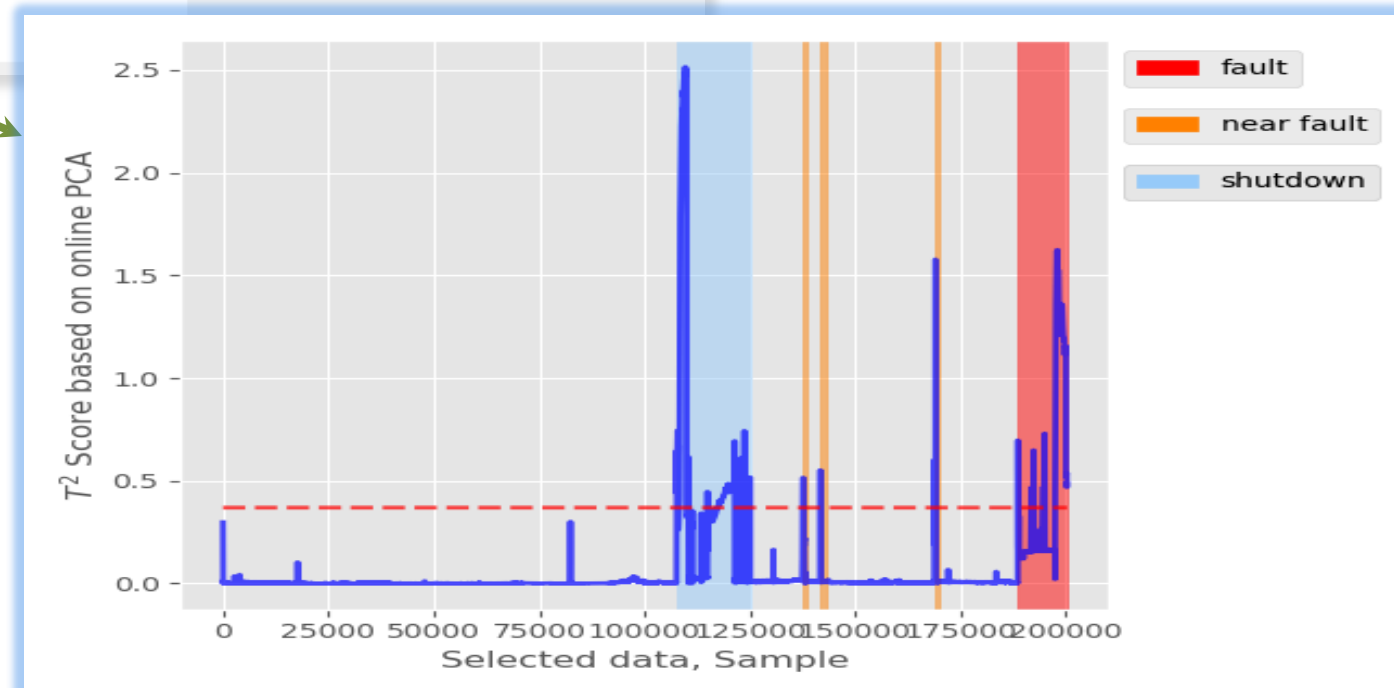
Click “Online PCA analysis” button

- 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.

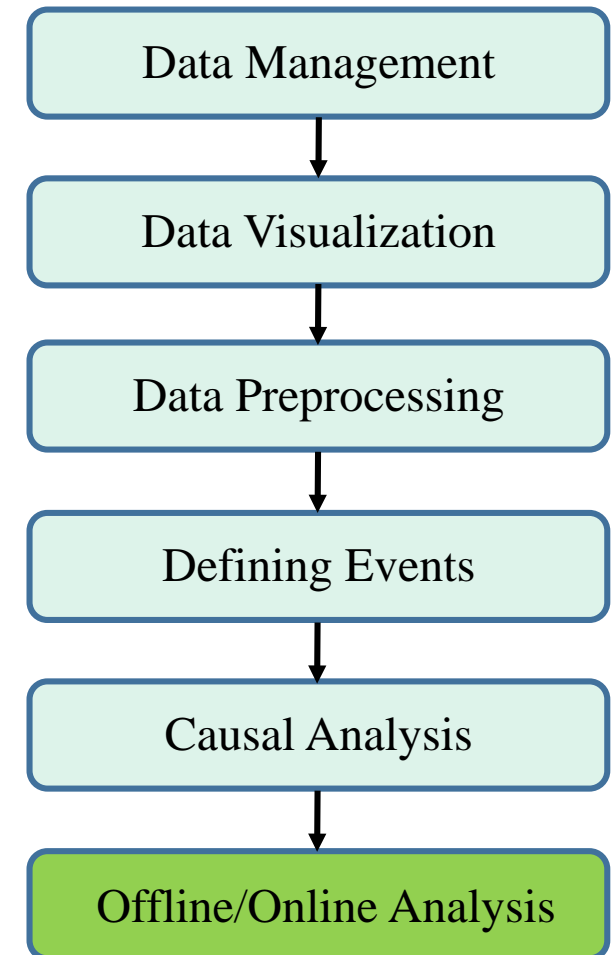
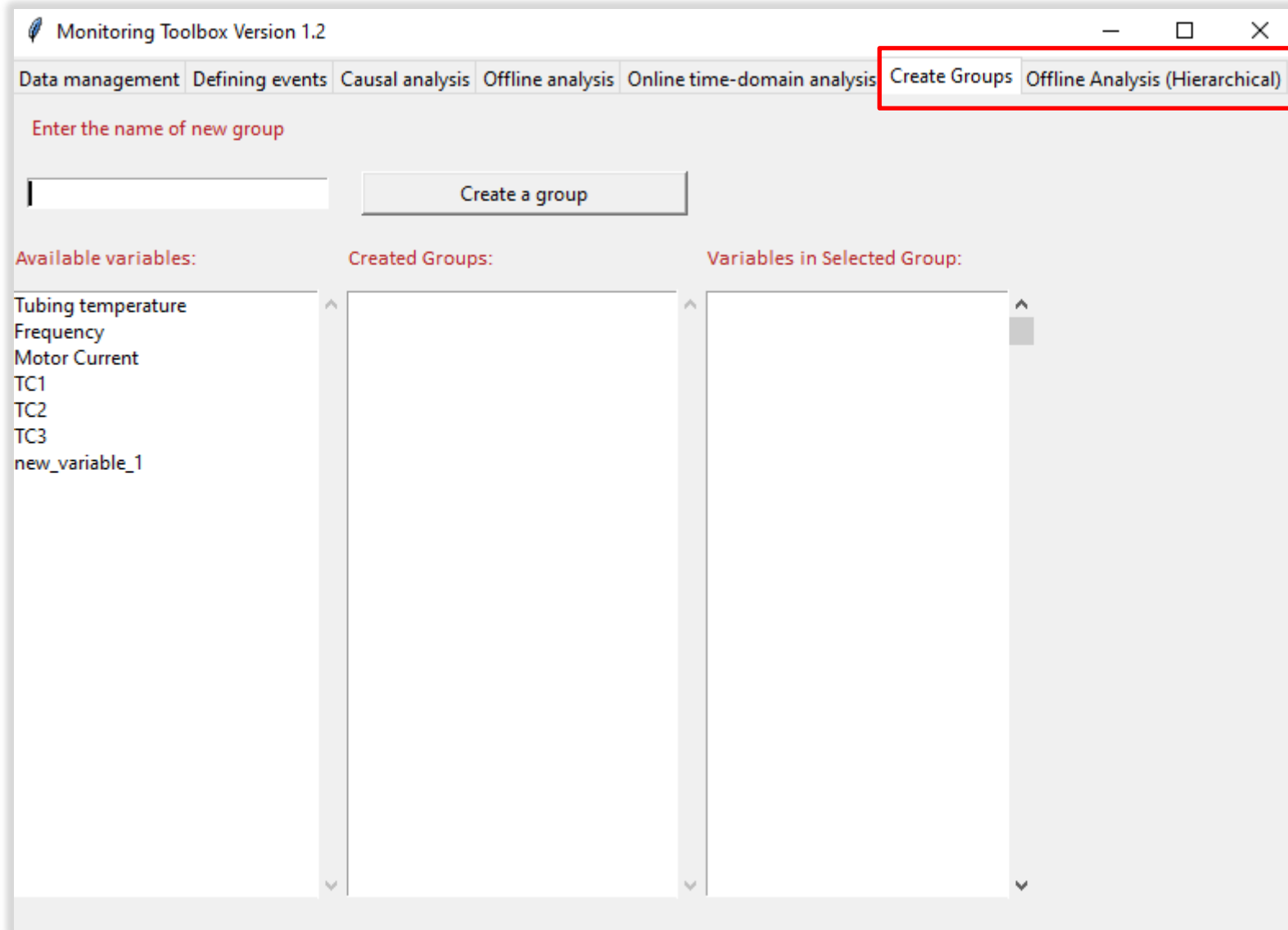


# Predictive Monitoring Toolbox



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

Monitoring Toolbox Version 1.2

Data management | Defining events | Causal analysis | Offline analysis | Online time-domain analysis | **Create Groups** | Offline Analysis (Hierarchical)

Enter the name of new group

G2

Create a group

Available variables:

Tubing temperature  
Frequency  
Motor Current  
**TC1**  
**TC2**  
**TC3**  
new\_variable\_1

Created Groups:

G1  
G2

Variables in Selected Group:

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

The screenshot shows the 'Monitoring Toolbox Version 1.2' application window. The 'Create Groups' tab is active. At the top, there is a text input field containing 'G2' and a 'Create a group' button. Below this, the interface is divided into three columns: 'Available variables:', 'Created Groups:', and 'Variables in Selected Group:'. The 'Available variables' list includes 'Tubing temperature', 'Frequency', 'Motor Current', 'TC1', 'TC2', 'TC3', and 'new\_variable\_1'. The 'Created Groups' list shows 'G1' and 'G2', with 'G2' highlighted in blue. The 'Variables in Selected Group' list shows 'TC1', 'TC2', and 'TC3'. A blue line connects the 'G2' group in the 'Created Groups' list to the 'Variables in Selected Group' list, indicating that the variables in that list are assigned to group G2.

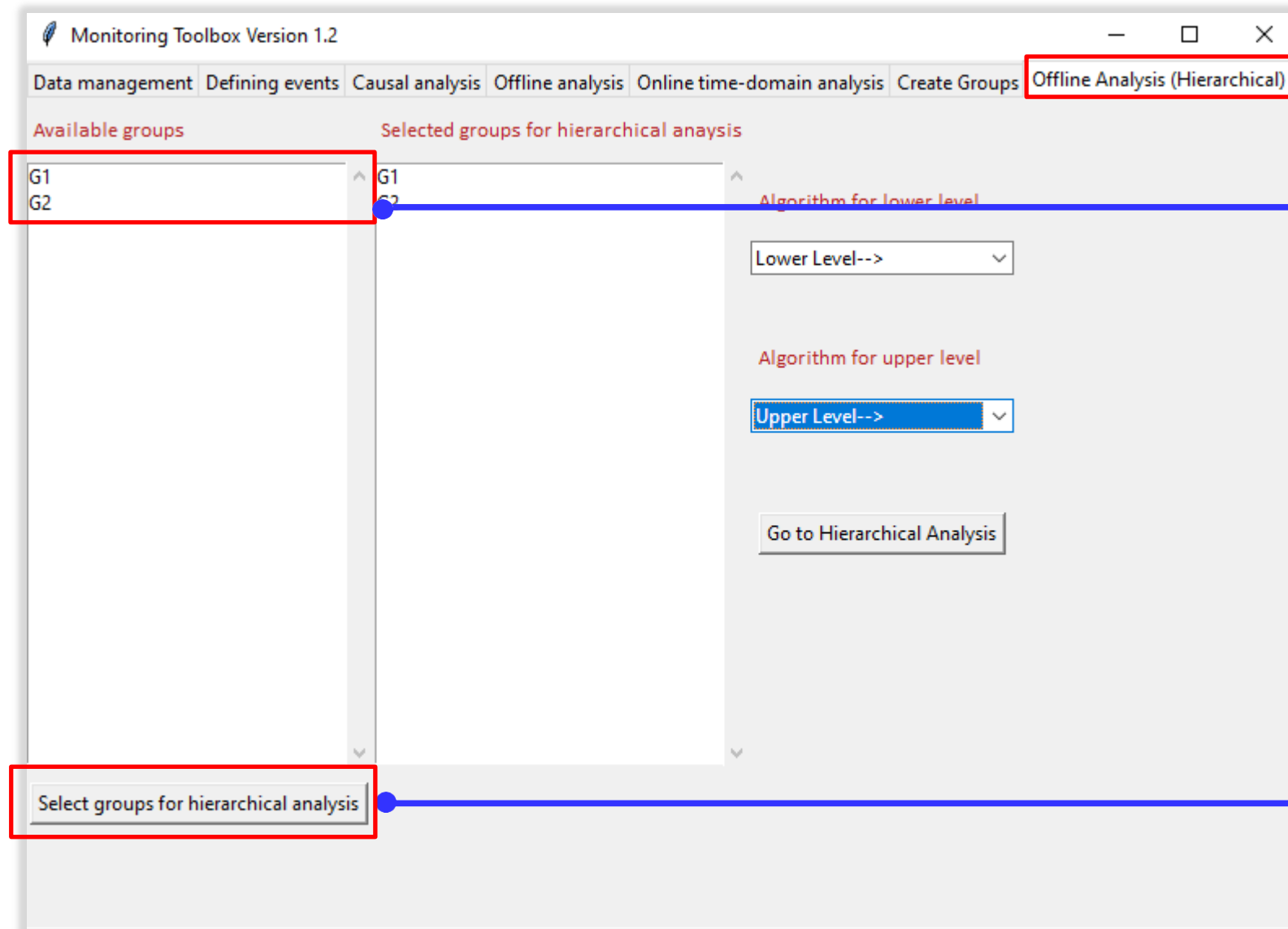
6

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

The screenshot shows the 'Monitoring Toolbox Version 1.2' interface. The 'Offline Analysis (Hierarchical)' tab is active. On the left, under 'Available groups', G1 and G2 are listed. On the right, under 'Selected groups for hierarchical analysis', G1 and G2 are also listed. Three red boxes highlight the configuration steps: 1. The 'Algorithm for lower level' dropdown menu is set to 'PCA'. 2. The 'Algorithm for upper level' dropdown menu is set to 'SFA'. 3. The 'Go to Hierarchical Analysis' button is highlighted. A blue line connects these three boxes to the numbered callouts on the right.

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

Hierarchical Offline Analysis

	Groups considered	Variables in each group	
Number of Components in a local agent	2	G1 G2	Dimension of visualization
Number of Components in Upper Layer	2	TC1 TC2 TC3	PC indices
Lag Samples (used for DPCA only)	5		
<b>Hierarchical Analysis</b>			Visualize Scores
<b>Show Statistics</b>			

- 4 Click "Hierarchical Analysis" button to conduct analysis
- 5 Click "Show Statistics" button to visualize results

