SIEMENS

–

Feasibility Study
Equipment Predictive Analytics (EPA)
Technical Specification

SIEMENS

TABLE OF CONTENTS

SIEMENS

TABLE OF FIGURES

SIEMENS

1. Goal of feasibility study

Siemens Equipment Predictive Analytics (EPA) provides decision support for the maintenance of mission critical process equipment in process industries.

Often customers are not sure if the available data and its quality are suitable for an EPA implementation. Therefore, Siemens offers a feasibility study where the customer data is being assessed with EPA and analyzed whether the solution is suitable for the customer's scope or not. If the provided data doesn't fulfill the app's requirements, the result of the study could be a recommendation for what needs to be done so the solution can be implemented.

2. Project Scope

The customer has to provide the data which we request in the chapter 4 Prerequisites. This data will be analyzed by Siemens data scientists.

As described in the chapter 6 Introduction to EPA it is necessary to perform several steps in order to be able to predict any anomaly. The first step is to screen the complete dataset of the customer to check if all necessary data is provided.
Concerning the time and effort only a few data is being analyzed and interpreted for the PQT.

2.1.1 Number of data sets

The customer can select one data set (e.g. one plant) which should be analyzed within the feasibility study. Further data sets can be considered in the actual project for EPA.

2.1.2 Number of analyzed sensors

The customer data will be analyzed by a data analyst and up to 35 sensors are selected for the further modelling process in Equipment Process Analytics (EPA).

2.1.3 Number of models

There will be up to 6 models built to show the application's potential. The decision on how to distribute the models among the assets will be up to Siemens data analyst.

2.1.4 Results

The results are shown in a demo on EPA[1]. Besides that, a report with screenshots and the most important results are provided to the customer.

[1] There will be no customer installation. Instead an internal EPA installation in Siemens is being used to show the results as demo to the customer.

SIEMENS

3. Project timeline

The project starts with a pre-modelling workshop with the customer earliest 10 days after the receipt of the purchase order. The project includes 1 day kickoff, 10 days work of a data analyst and 1 day of result presentation. The entire project will be executed within 4 weeks. Travel costs are not included.

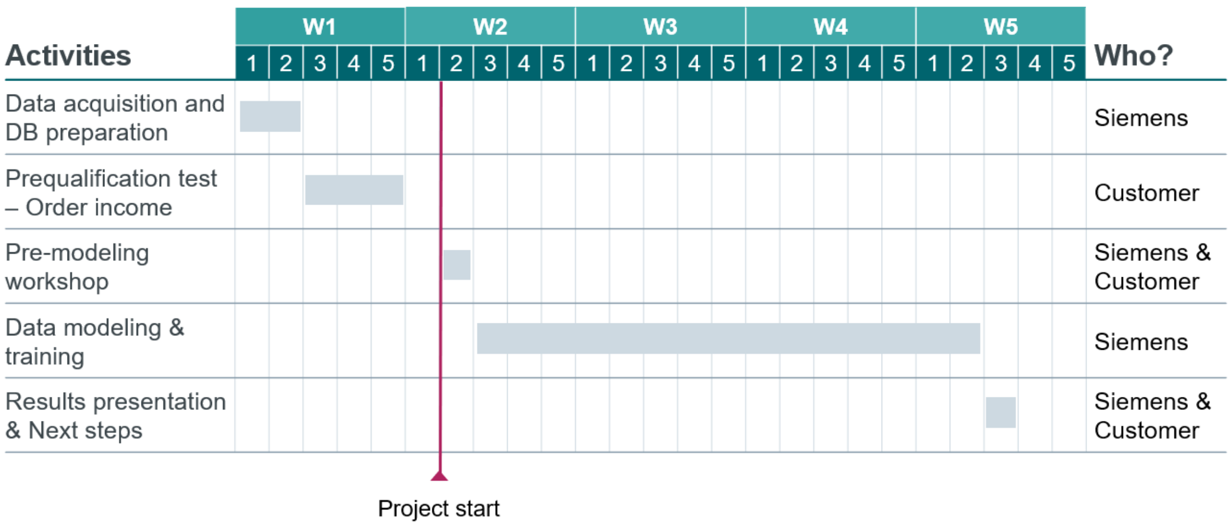

Figure 1: Project timeline

4. Prerequisites

As a tool to support advanced data analysis for predictive maintenance, EPA's full potential can only be achieved if the customers' and Siemens expertise are fully combined. In particular, necessary information should be provided by the customer via workshops or communications during the project execution process to ensure the success of the application.

4.1 Historical data

In order to train the models, there is a need for historical real manufacturing data of the equipment. Historical data of minimum half a year (labeled with normal operation modes) are needed. This dataset includes both historical sensor data as well as DCS alarm data, and this dataset must include all operation modes which should be treated as normal behavior of the equipment. Please be aware that only operation modes which are included in the historical data set can be reasonably monitored with the app. An example for different operation modes would be the distinction between winter and summer time.

SIEMENS

4.2 Failure modes

In order to test different failure types of the equipment, examples labeled with dates and background information should be provided. The more examples of failures are provided, the more accurate the trained models will be.

4.3 Training periods

EPA uses machine learning and therefore needs to know what periods are identified as good to train the models. Training periods for each equipment must be provided.

4.4 Charts and data overviews

An overview of the instrumentation of all sensor tags related to the equipment, which should be monitored, should be provided for a better understanding. This overview can be provided as diagram, sketches etc. Also, a mapping table for the provided sensors to the overview such as P&ID charts are required.

4.5 Data format

4.5.1 Historical sensor data

The historical sensor dataset must be provided in a csv-file with the information about the sensor tags, timestamps and corresponding sensor values. The minimum frequency of data is 5 min. A higher frequency, e.g. 30 secs, is even better to catch more detailed equipment behavior.

The historical sensor dataset must be in the following format:

- For each sensor, the timestamps and sensor values are needed
- Each sensor has its own CSV file
- The name of the files has the names of the sensors (see format below)
- The timestamps are given by unix timestamp in milliseconds
- Each csv file for each sensor must have the format:

unix timestamp 1 in milliseconds, sensor value 1

unix timestamp 2 in milliseconds, sensor value 2...

Example:

Sensor name: Sensor_1 → The sensor values are stored in the file sensor_1.csv

The content of the file is:

```
1466229601290,0.4026583433151245
1466229603290,0.40218815207481384
1466229605290,0.402441322803497
1466229607290,0.40229666233062744
1466229609290,0.40207964181900024
1466229611290,0.40262216329574585
1466229613290,0.4023689925670624
1466229615290,0.40204349160194397
1466229617290,0.40240517258644104
1466229619290,0.402477502822876
1466229621290,0.402586013078689
1466229623290,0.4022243320941925
```

Figure 2: Sensor values

4.5.2 OPTIONAL Historical DCS alarm data

If the customer cannot define stable time periods for training as stated in chapter 4.3 Training periods the historical DCS alarm data should be provided, so that the data analyst can identify training periods for EPA. The dataset should be provided in a csv-file, whose format should be as follow (match the standard OPC AE protocol):

Area, Batch name, Class, Comment, **Date**, Event, Info, Message Duration, Number, Operation, **Priority**, **Source**, **Status**, **Type**

Note:

- The value of 5 under-lined attributes (i.e., Date, Priority, Source, Status, Type) are mandatory to support data processing in EPA.
- The time range of this dataset should match the time range of historical sensor data.

Area	Batch name	Class	Comment	Date	Event	Info	Message Duration	Number	Operation	Priority	Source	Status	Type
Styrene Unit		Tolerance		2016-02-15 09:51:49	Lubr Oil Temp High Alarm LIM1		0	671088824		1	1119_TE_3708/MEAS	C	Tolerance High
Styrene Unit		Tolerance		2016-02-15 09:51:56	Lubr Oil Temp High Alarm LIM1		6000	671088824		1	1119_TE_3708/MEAS	G	Tolerance High
Styrene Unit		Tolerance		2016-02-15 09:51:58	Lubr Oil Temp High Alarm LIM1		8000	671088824		1	1119_TE_3708/MEAS	QS	Tolerance High
Styrene Unit		Tolerance		2016-02-15 09:52:04	Lubr Oil Temp High Alarm LIM1		0	671088824		1	1119_TE_3708/MEAS	C	Tolerance High
Styrene Unit		Tolerance		2016-02-15 09:52:08	Lubr Oil Temp High Alarm LIM1		4000	671088824		1	1119_TE_3708/MEAS	QS	Tolerance High
Styrene Unit		Tolerance		2016-02-15 09:52:11	Lubr Oil Temp High Alarm LIM1		7000	671088824		1	1119_TE_3708/MEAS	G	Tolerance High
Styrene Unit		Operator me	1003101	2016-02-15 09:53:26	1119_TE: Acknowledgment [illegible]	[illegible] or	0	1003101	1119_TE: Acknowled	0	1119_TE_3708/MEAS	C	Operator Message

Figure 3: Example csv

5. Terms and condition

5.1 Scope and execution of the Service

The scope, quality, and all conditions for the consulting services and any other services to be provided by Siemens ("Services") are exclusively defined in the technical proposal and these "Standard Terms and Conditions for Consulting Services of the Division DF for Customers with a Seat or Registered Office Outside of Germany" (collectively referred to as "Contract"). The contract contains the entire agreement between the parties.

The project is done without any interactions in person. Discussions and the presentation of results are done remote via skype or similar.

5.2 General duties of corporation

The customer agrees to answer relevant questions concerning the data during the feasibility study.

5.3 Delay

A positive outcome depends on the fulfilled prerequisites of the customer mentioned in chapter <u>*4 Prerequisites*.</u> The region must ensure that all prerequisites are fulfilled by the customer and that the data is complete before the feasibility study starts. Any delay on delivering the data or needed information has a direct impact on the project's timeline and outcome.

5.4 Data Use Rights

During and after the term of the agreement, Siemens and its business partners may use Collected Data for Siemens' internal purposes (e.g. development or improvement of products or services). On an aggregated basis with other data and in a form that does not identify you and your Users, Siemens shall own and be free to make Collected Data publicly available to you and others (e.g. for information and industry trends, benchmarking data). Use of Collected Data in accordance with this Section will be at our risk.

Collected data include:

- any device data collected from the asset (csv – raw data)
- processed data (KPIs, histograms, events, comments, plans)
- log data (app log, system log)

5.5 Cloud based data storage and analyzing

Siemens is authorized to store and evaluate the data provided by the customer in the cloud (e.g. AWS).

SIEMENS

6. Introduction to EPA

Siemens Equipment Predictive Analytics (EPA) provides decision support for the maintenance of mission critical process equipment in process industries.

With the constant pressure to increase plant efficiency and uptime in a highly competitive environment, condition-based maintenance – as much as needed, but as little as possible - becomes more and more important for a plant's economic success.

Siemens Equipment Predictive Analytics (EPA) is a practical and robust tool that integrates human experience/ know-how and machine learning capability to extract relevant information on equipment status. The information is summarized in an intuitive graphical user interface for easy use by plant operators and maintenance staff. This allows early detection of changes in equipment behavior and enables more solid and reliable decisions regarding the operation of the process and required maintenance work.

6.1 How does EPA work?

EPA enables the users to monitor their equipment's most important sensors grouped by subsystem in one dashboard. The following chapter explains how it works and enables users to assess their sensors more effectively.

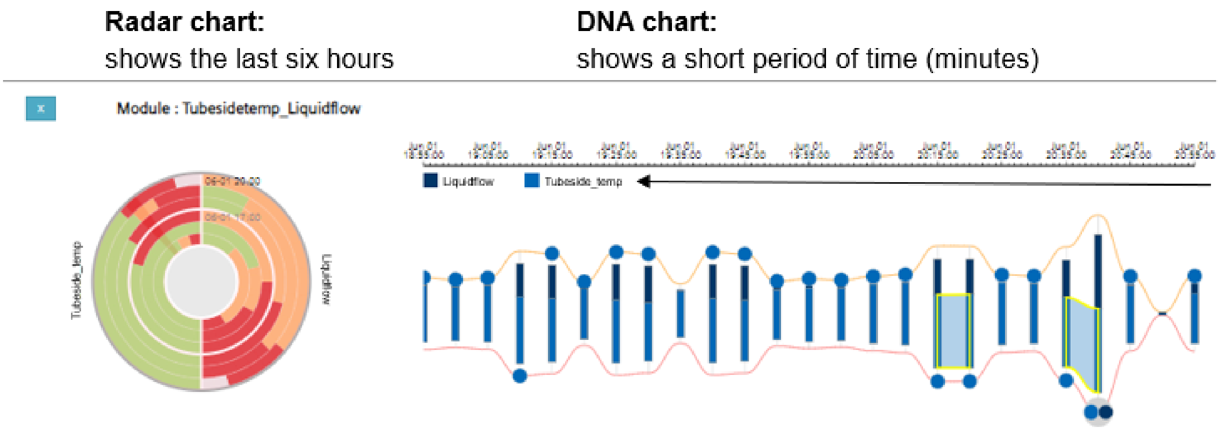

Figure 4: Dashboard overview

6.1.1 Calculate correlations and create models

Based on historical data, the system automatically calculates how sensors correlate with each other in order to define the perfect base for a model. The correlation analysis considers the time-related dependency between sensors, this means also time-shifted dependencies between sensors are considered.

After the models are created the results should be evaluated and adjusted by a domain expert to integrate domain knowledge for a high-quality model which will be able to detect potential risks.

Whenever a model is being built, as the first step, the user should choose the most important sensor or the indicating sensor (KPI) which they usually monitor to determine the equipment status. This sensor is called the target sensor. This could either be a data-driven choice based on correlations, or a choice based on experience/ domain knowledge.

For each chosen target sensor, a model must be built by learning the normal behavior of the sensor from historical data. Afterwards the models evaluate the real-time sensor values for each time point to decide if the sensor is still in normal operation.

If these estimated values are near to the measured sensor values, the equipment is still in normal condition. However, if the difference between these values is large, the equipment's behavior is considered as abnormal. In order to define a tolerance for the expected values a confidence band is calculated by the system. As long as the measured values are within the confidence band, the equipment is still in normal condition. As soon as the measured values exceed the limits of the confidence band, the system throws an alert which informs the operator about the equipment's condition. The user can adjust the tolerated range by configuring the width of the confidence band.

6.1.2 Train models

Since EPA's purpose is to inform the user whenever a target sensor is behaving abnormally, we must know what the normal behavior is. Therefore, the system needs to be trained with real historical manufacturing data[2] of the equipment's healthy periods to learn what a normal behavior looks like. Using this technique also avoids challenges of enumerating complicated rules, especially considering all relevant factors which may affect the result. Moreover, EPA enables the use of several training periods to ensure the system has the highest level of accuracy.

6.1.3 Configure the algorithms

The customer can create new models and adjust existing models by changing the number of correlated sensors within a model or adding training periods to a model themselves. The training to do so can be purchased as a service from Siemens if necessary.

Siemens also offers a service with the first installation to create models for customers. The accuracy of the models is influenced by the combination of sensors, the number of defined training periods and the configuration of the confidence band. As part of the modelling processes the constellation of the variables are changed and the different effects to create models are tested to ensure the best results for our customers. This functionality is provided by Siemens as a service and not part of the application itself.

6.1.4 Create modules

In order to keep the dashboard as clear as possible, it is possible to group several target sensors to modules. These modules can be used to represent a subsystem and helps the operators to organize their target sensors in logical groups.

[2] Data must be provided by customer. Please check prerequisite chapter

6.1.5 Dashboard and alerts

Whenever the system is set up, operators can monitor their subsystems on the dashboard. If an anomaly is detected, the system throws an alert. The alerts are categorized as low and high-risk alerts.

For each target sensor, the dashboard shows the number of alerts over the last six hours and their categorization. In case there are no alerts, the dashboard shows green frequencies within the chart. The color scheme enables operators to recognize anomalies quickly.

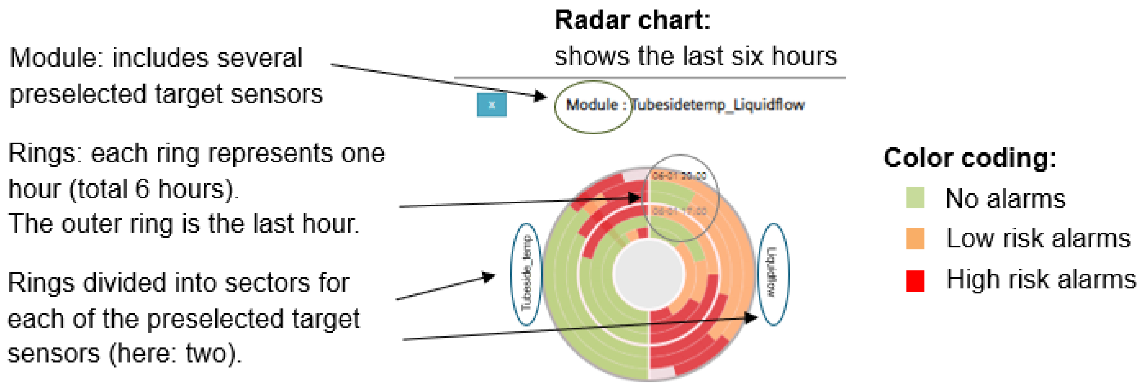

Figure 5: Dashboard description for radar chart

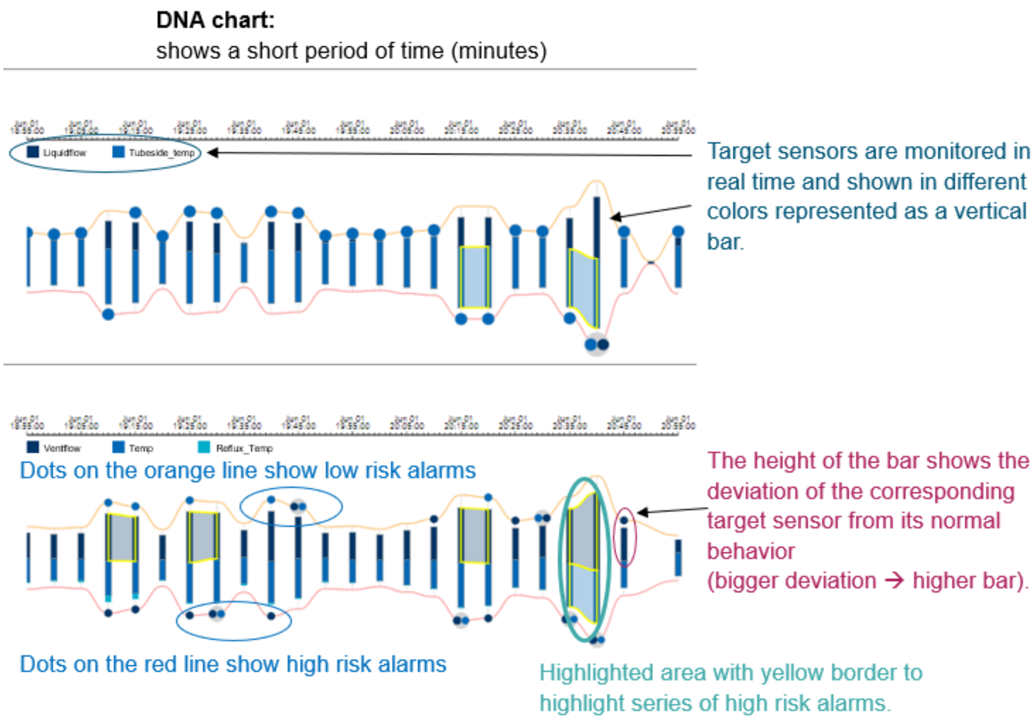

Figure 6: Dashboard description for DNA chart