D2.5 Final report on equipment degradation modelling

Lead Author: Lenzing
Involved Partners: INEOS, P&G, DIVIS, UVA

May 2019
www.spire2030.eu/copro
Deliverable 2.5
Final report on equipment degradation modelling

Project Details

**PROJECT TITLE**
Improved energy and resource efficiency by better coordination of production in the process industries

**PROJECT ACRONYM**
CoPro

**GRANT AGREEMENT NO**
723575

**INSTRUMENT**
RESEARCH AND INNOVATION ACTION

**CALL**
H2020-SPIRE-02-2016

**STARTING DATE OF PROJECT**
NOVEMBER, 1ST 2016

**PROJECT DURATION**
42 MONTHS

**PROJECT COORDINATOR**
PROF. SEBASTIAN ENGELL (TUDO)

The CoPro Project

The goal of CoPro is to develop and to demonstrate methods and tools for process monitoring and optimal dynamic planning, scheduling and control of plants, industrial sites and clusters under dynamic market conditions. CoPro pays special attention to the role of operators and managers in plant-wide control solutions and to the deployment of advanced solutions in industrial sites with a heterogeneous IT environment. As the effort required for the development and maintenance of accurate plant models is the bottleneck for the development and long-term operation of advanced control and scheduling solutions, CoPro will develop methods for efficient modelling and for model quality monitoring and model adaption.

The CoPro Consortium

<table>
<thead>
<tr>
<th>Participant No</th>
<th>Participant organisation name</th>
<th>Country</th>
<th>Organisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (Coordinator)</td>
<td>Technische Universität Dortmund (TUDO)</td>
<td>DE</td>
<td>HES</td>
</tr>
<tr>
<td>2</td>
<td>INEOS Manufacturing Deutschland GmbH (INEOS)</td>
<td>DE</td>
<td>IND</td>
</tr>
<tr>
<td>3</td>
<td>Covestro Deutschland AG (COV)</td>
<td>DE</td>
<td>IND</td>
</tr>
<tr>
<td>4</td>
<td>Procter &amp; Gamble Services Company NV (P&amp;G)</td>
<td>BE</td>
<td>IND</td>
</tr>
<tr>
<td>5</td>
<td>Lenzing Aktiengesellschaft (LENZING)</td>
<td>AU</td>
<td>IND</td>
</tr>
<tr>
<td>6</td>
<td>Frinsa del Noroeste S.A. (Frinsa)</td>
<td>ES</td>
<td>IND</td>
</tr>
<tr>
<td>7</td>
<td>Universidad de Valladolid (UVA)</td>
<td>ES</td>
<td>HES</td>
</tr>
<tr>
<td>8</td>
<td>École Polytechnique Fédérale de Lausanne (EPFL)</td>
<td>CH</td>
<td>HES</td>
</tr>
<tr>
<td>9</td>
<td>Ethniko Kentro Erevnas Kai Technologikis Anaptyxis (CERTH)</td>
<td>GR</td>
<td>RES</td>
</tr>
<tr>
<td>10</td>
<td>IIM-CSIC (CSIC)</td>
<td>ES</td>
<td>RES</td>
</tr>
<tr>
<td>11</td>
<td>LeiKon GmbH (LEIKON)</td>
<td>DE</td>
<td>SME</td>
</tr>
<tr>
<td>12</td>
<td>Process Systems Enterprise LTD (PSE)</td>
<td>UK</td>
<td>SME</td>
</tr>
<tr>
<td>13</td>
<td>Divis Intelligent Solutions GmbH (divis)</td>
<td>DE</td>
<td>SME</td>
</tr>
<tr>
<td>14</td>
<td>Argent &amp; Waugh Ltd. (Sabisu)</td>
<td>UK</td>
<td>SME</td>
</tr>
<tr>
<td>15</td>
<td>ASM Soft S.L (ASM)</td>
<td>ES</td>
<td>SME</td>
</tr>
<tr>
<td>16</td>
<td>ORSOFT GmbH (ORS)</td>
<td>DE</td>
<td>SME</td>
</tr>
<tr>
<td>17</td>
<td>Inno TSD (inno)</td>
<td>FR</td>
<td>SME</td>
</tr>
</tbody>
</table>
Deliverable 2.5
Final report on equipment degradation modelling

Abstract
This report provides a description of the work carried out for Work package 2.4: Monitoring of equipment degradation. It covers the general fundamentals of equipment degradation and a more detailed description of the two degradation types fouling and cooking. A methodology for database equipment degradation monitoring and modelling is presented and exemplified on an actual plate heat exchanger from the Lenzing use case in COPRO. The methodology contains data preprocessing, data selection, the generation of a clean equipment performance model, online degradation monitoring and modelling to predict future equipment degradation behaviour. The results show that the state of equipment degradation can be predicted with a mean square error of 3% over a time horizon of 55 days. Furthermore, an approach is presented to monitor and predict the cooking index of a naphtha cracker from the INEOS use case in COPRO.

Revision History
The following table describes the main changes done in the document since it was created.

<table>
<thead>
<tr>
<th>Revision</th>
<th>Date</th>
<th>Description</th>
<th>Author (Organisation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>V0.1</td>
<td>20 February 2019</td>
<td>Creation</td>
<td>C. Jasch (LENZING)</td>
</tr>
<tr>
<td>V0.2</td>
<td>25 March 2019</td>
<td>First draft ready including contributions from divis for INEOS case</td>
<td>C. Jasch (LENZING)</td>
</tr>
<tr>
<td>V0.3</td>
<td>01 April 2019</td>
<td>Review</td>
<td>S. Wenzel (TUDO)</td>
</tr>
<tr>
<td>V0.4</td>
<td>02 April 2019</td>
<td>Revised after comments from TUDO</td>
<td>C. Jasch (LENZING)</td>
</tr>
<tr>
<td>V0.5</td>
<td>16 April 2019</td>
<td>Revised after comments from TUDO</td>
<td>C. Jasch (LENZING)</td>
</tr>
<tr>
<td>V0.6</td>
<td>26 April 2019</td>
<td>Added P&amp;G contribution</td>
<td>S. Däubener (divis)</td>
</tr>
<tr>
<td>V0.7</td>
<td>26 April 2019</td>
<td>Review of the added P&amp;G Use Case</td>
<td>S. Wenzel (TUDO)</td>
</tr>
<tr>
<td>V0.8</td>
<td>26 April 2019</td>
<td>Revised after comments from TUDO</td>
<td>S. Däubener (divis)</td>
</tr>
<tr>
<td>V0.8</td>
<td>21 May 2019</td>
<td>Final approval</td>
<td>S. Engell (TUDO)</td>
</tr>
</tbody>
</table>
THIS DOCUMENT IS PROVIDED "AS IS" WITH NO WARRANTIES WHATSOEVER, INCLUDING ANY WARRANTY OF MERCHANTABILITY, NONINFRINGEMENT, FITNESS FOR ANY PARTICULAR PURPOSE, OR ANY WARRANTY OTHERWISE ARISING OUT OF ANY PROPOSAL, SPECIFICATION OR SAMPLE. Any liability, including liability for infringement of any proprietary rights, relating to use of information in this document is disclaimed. No license, express or implied, by estoppels or otherwise, to any intellectual property rights are granted herein. The members of the project CoPro do not accept any liability for actions or omissions of CoPro members or third parties and disclaims any obligation to enforce the use of this document. This document is subject to change without notice.

Table of contents
1 Executive summary ...................................................................................................... 6
2 Introduction ................................................................................................................. 6
3 Equipment Degradation ............................................................................................... 7
   3.1 Fouling ..................................................................................................................... 8
   3.2 Coking ....................................................................................................................... 9
4 Lenzing use case ......................................................................................................... 10
   4.1 Equipment degradation in Lenzing ........................................................................ 10
   4.1.1 Degradation of the evaporators ........................................................................... 11
   4.1.2 Degradation of the plate heat exchangers ........................................................ 12
   4.2 Equipment degradation modelling ........................................................................ 13
   4.2.1 Data pre-processing ............................................................................................ 13
   4.2.2 Data selection and clean model generation ....................................................... 16
   4.2.3 Degradation monitoring .................................................................................... 18
   4.2.4 Degradation modelling ..................................................................................... 20
   4.3 Outlook .................................................................................................................. 23
5 INEOS use case ........................................................................................................... 24
   5.1 Degradation of the cracker .................................................................................... 24
   5.2 Degradation modelling .......................................................................................... 24
   5.3 Conclusion ............................................................................................................ 26
6 P&G use case ............................................................................................................. 28
   6.1 The maintenance use case .................................................................................... 28
   6.2 Methods for developing a maintenance indicator ............................................... 28
   6.3 Conclusion ............................................................................................................. 30
7 References ................................................................................................................ 30
Deliverable 2.5
Final report on equipment degradation modelling
1 Executive summary

As a part of COPRO, novel data processing methods are developed and tested for process and product quality monitoring. Especially in the process industries with its high employment of equipment, the ability to monitor and to predict the equipment degradation is a great advantage. Therefore, this workpackage is concerned with equipment degradation monitoring and modelling.

This report presents a summary of the work that Lenzing, INEOS, P&G, divis and UVA conducted during the first 30 month of the project on Task 2.4: Monitoring of equipment degradation. The report is about how certain equipment and plants on the Lenzing and INEOS site suffer from degradation and how this degradation can be monitored and predicted. In the Lenzing use case the focus is particularly on the degradation of evaporators and heat exchangers caused by fouling. While in the INEOS use case the focus is on the degradation of the naphtha cracker caused by coking. While for the previously mentioned use cases the specific cause of degradation is known, in the P&G use case general abrasion is assumed without a specific cause for it.

First, equipment degradation and its effects are described in general and individually for the Lenzing and INEOS use case. Later, the developed methodologies to monitor and model the state of equipment degradation, in order to forecast it, are presented and explained in details using actual equipment data from the Lenzing, INEOS and P&G use cases. The methodologies consist of different steps. In the first step, data collection and data preprocessing are carried out in order to exclude outliers and to collect relevant data. In the second step, the data is classified and selected to set up a data driven model used to describe “clean” equipment performance without degradation. In a last step, the difference between the model and the actual measurement data is used to determine the state of equipment degradation and to predict the future degradation behaviour.

The presented methodologies enable users to monitor the state of equipment degradation for different process equipment in real time and to build efficient degradation models to predict the future fouling behavior based only on ordinary process data. These models can be used to optimize the equipment cleaning scheduling and therefore improve the energy efficiency in process industries.

2 Introduction

In a competitive market with a high level of system complexity and strict safety requirements, an efficient asset management is of paramount importance for the process industries in Europe. According to [1], 15–40 % of the manufacturing costs across many industries including the process industries are attributed to maintenance. In the last decades, there has been a change in the maintenance strategy. The old and relatively simple strategy “fix it when it breaks” is no longer appropriate for most companies. Since the resulting unscheduled downtimes are obstructive for the often times tight production schedules and the resulting breakdowns represent a major threat to health, safety and environment. Also, especially in resource intensive industries with high outputs like the process industries, the focus is not just on equipment reliability but also on equipment performance. Most process equipment shows a degradation in performance such as a decrease in overall capacity or an increase in specific energy consumption over time even before reliability is affected.
These challenges provide a motivation to change the strategy and to perform maintenance before equipment breakdowns occur. The simplest approach is to perform maintenance at pre-established consistent intervals that are often defined in terms of operating hours. While this approach provides relative high equipment reliability, it rarely represents the optimal solution with respect to the changes in equipment performance. The development of the equipment performance is affected by the complex production environment and the frequently changing process conditions, which cannot be represented in pre-established consistent maintenance intervals. These intervals lead to either excessive maintenance costs when maintenance is performed too often or to poor overall equipment performance when it is performed not often enough. To define optimal maintenance intervals, a method to monitor the state of equipment health (or state of equipment degradation) during operation is necessary. Based on this information, a model for the equipment degradation process can be obtained in order to predict the future state of equipment degradation [2].

This report presents methodologies to monitor the degradation levels of different process equipment during operation and to develop data-driven equipment degradation models necessary for an efficient maintenance scheduling. The report focuses on two different causes for equipment degradation: cooking and fouling. The process of equipment degradation as well as the methodologies for the degradation level monitoring and modelling are developed, explained and tested within the INEOS and LENZING use cases inside the CoPro project.

3 Equipment Degradation

Every process equipment is unreliable in the sense that it degrades and eventually fails. Although there is the possibility that process equipment fails suddenly without any evidence of degradation, usually most of them go through different stages of degradation. These stages can be described by a function of the operating performance over time, illustrated in the degradation cycle in Figure 1.

![Equipment Degradation Cycle](image)
During the first stage, the equipment runs smoothly and almost no degradation of the operating performance can be detected. At point P, the second stage starts where a change in performance is detectable and equipment degradation is accelerating. In the last stage starting at point F, the performance decreases rapidly and a failure impends. In this stage, often equipment maintenance is not sufficient and costly repairs or replacements are necessary. The actual trend of the operating performance highly depends on the observed process equipment and on the cause for equipment degradation.

In the process industries the causes for equipment degradation are numerous. Some of the most common causes are corrosion, wear, aging, improper operation, improper environment and the deposition of unwanted material inside the process equipment often known as fouling. In the following, we describe two special variants of fouling in more detail: the fouling of heat transfer equipment in a viscose fibre plant and the fouling in a naphtha-cracker also known as coking. [3]

### 3.1 Fouling

Fouling can be described as the unwanted build-up of deposits on a surface. These deposits may be caused by sedimentation, crystallization, biological growth, chemical reactions, corrosion products, or a combination of these effects. Fouling is a major problem all around the process industries. For instance, it can lead to product quality problems such as contamination. Pipes and valves can clog which leads to pressure drops and throughput problems. But most important, fouling decreases the effectiveness of the heat transmission and thereby is responsible for a significant amount of wasted energy. Furthermore, the costs due to downtime because of necessary cleanings have to be taken into account. In [4], Müller-Steinhagen estimated the costs of heat exchanger fouling at about 0.25% of the GDP of industrialized countries. [5]

In many cases, the deposits have a very different chemical composition compared to the process fluids due to preferential deposition of some component of the material. But in some cases, it is the deposition of the processed material itself. Despite the different causes for fouling, the fouling process in heat exchangers itself generally involves five different phases: [6]

1. **Initiation:** The delay period before fouling can be detected
2. **Transport:** The phase where the fouling components are transported from the bulk of the fluid to the heat transfer surface. The mass transfer rate \( \dot{m} \) described in (1) depends on the transport coefficient \( k_t \) and the concentration of the fouling components within the bulk \( C_b \) and the surface \( C_s \).
   \[
   \dot{m} = k_t \cdot (C_b - C_s) \tag{1}
   \]
3. **Attachment:** In this phase adsorption takes place and the fouling components attach to the surface
4. **Removal:** The removal of the deposits generally occurs by shear stress, thermal stress or by diffusion
5. **Aging:** The deposits on the surface undergo an aging process where the fouling material is usually transformed to a more cohesive form. The transformation changes the chemical and physical nature of the deposits.
Depending on the fouling mechanism and conditions as well as the equipment design, the rate of fouling may be linear, falling, asymptotic or it may follow a sawtooth rate, where significant removals take place. In Figure 2, the possible fouling curves are illustrated.

![Figure 2: Possible fouling curves](image)

Since fouling is a complex process, there are numerous variables that affect fouling. Often, four key variables can be identified:

1. The composition of the bulk phase is crucial for the type of fouling that takes place.
2. The temperature of the bulk phase and surface determines the rate of reaction, higher temperatures usually promote fouling and the ageing effect.
3. The fluid velocity determines the shear stress on the deposit/liquid interface and strongly affects the transport and removal phase. Higher fluid velocity impedes fouling and may even result in fouling detachment.
4. The surface conditions and roughness effects especially the attachment phase. A higher roughness encourages the fouling rate. The coating of surfaces to make them smoother is a common strategy for fouling mitigation.

### 3.2 Coking

Steam cracking is a thermal process to produce light olefins such as ethylene and propylene from paraffins [9]. At INEOS in Köln, naphtha is cracked into several hydrocarbons with short chain length that serve as the raw material of the follow-up plants. This technology has been developed in the early 1940s, when the first commercial plant came into operation [10].

The core pieces of equipment of a steam cracker are the furnaces. In tubes (“coils”), a mixture of naphtha and steam is heated to a temperature of up to 900°C using fuel gas leading to the “cracking” of the long chain paraffin molecules to short chain olefin molecules such as ethylene and propylene that are the building blocks of most polymers. A byproduct of the cracking process is the formation of coke. During normal operation, coke is deposited on the inside of the metal tubes of the cracker...
furnaces. The deposits lead to shorter residence times, different cracking kinetics with different cracked gas compositions, higher energy demand and equipment strain. The latter is visible through a higher tube metal surface temperature of the coils. Thereby, the furnace must be taken out of operation in order to remove the coke layer by burning it off.

This process, known as “decoking”, stresses the furnace equipment and is a time of no-production. If the furnace of a cracker is the bottleneck of the utilization, decoking implies that the load of the whole cracker has to be reduced. At INEOS in Köln, a typical run-length of 60 days is defined per furnace without having any indication of the coke deposition inside the coils. Both the monitoring of the coking process to define an optimal run-length as well as an optimal decoking scheduling have a high potential to increase the utilization and the economic performance of the plant. The crucial step to address the monitoring and the scheduling is the development of a suitable model that is able to estimate/predict the coke deposition inside the coils based on the existing measurements in the cracker complex.

4 Lenzing use case

Lenzing operates in Upper Austria the largest fully integrated viscose fibre production site worldwide. The production of viscose fibers is a chemical-technological process that proceeds in multiple steps and is associated with a high employment of process equipment. Especially in the recovery section, a large number of different and redundant equipment is used to recover the abrasive and aggressive spinbath liquid and co-products. In accordance, the maintenance costs are relatively high and equipment degradation is a major challenge. In the COPRO project, Lenzing developed and currently tests a new data-driven method to determine the degradation state of certain equipment types and to model their equipment degradation behaviour. The results will later be used to optimize the cleaning scheduling of the equipment.

4.1 Equipment degradation in Lenzing

The main operating resource in the viscose fibre production is the spinbath liquid. It consists of sulfuric acid, sodium sulfate and zinc sulfate. During the production process other sulphuric compounds like H₂S and CS₂ dissolve in it. Furthermore, the spinbath liquid contains a significant amount of organic impurities deriving from the organic viscose. As a result, any equipment that comes into contact with the spinbath liquid tends to foul. Often, a significant fouling layer on the surface of the equipment can be observed after some operation time. In Figure 3, a typical fouling layer from the spinbath liquid on the inside of a pipe is illustrated.
In the recovery section of the Lenzing production site, the equipment that is affected the most by fouling are the evaporators and heat exchangers used to re-concentrate and reheat the spinbath liquid. Due to fouling, these process units show a decrease in equipment performance that leads to overall higher specific energy consumption.

### 4.1.1 Degradation of the evaporators

In Figure 4, a simplified illustration of a multi-stage evaporator used at Lenzing can be seen. The evaporator is equipped with a condenser unit to create a vacuum and uses live steam as main heat source.
Although fouling can be observed in all parts of the evaporator, the areas that are affected the most by fouling are the live steam heat exchangers, the vapor steam heat exchangers and the condenser. The live steam heat exchangers are shell-and-tube counter flow exchangers, which heat the spinbath up to the maximum temperature before it enters the evaporation chambers. Here, the highest temperatures are reached and thereby these exchangers are affected the most by fouling. The vapor steam heat exchangers are also of the shell-and-tube counter flow type. These exchangers suffer from fouling on both sites since the vapor steam contains sulfuric compounds that accumulate during condensation. Some evaporators in Lenzing use river water to operate the condenser used to create the vacuum. These condenser show a special type of fouling – the bio-fouling – since algae-growth is accelerated due to the warmer temperatures in the condenser.

Despite the different fouling areas and procedures in the evaporator the result is still the same: equipment performance degradation. More precise the heat transfer efficiency decreases and therefore the specific energy consumption of the entire evaporator increases due to fouling.

### 4.1.2 Degradation of the plate heat exchangers

In the recovery section, Lenzing uses a network of plate heat exchangers to reheat the spinbath liquid after the co-product sodium sulfate is extracted via crystallization. The network consists of more than 20 different counter flow plate heat exchangers that use warm waste water or condensate streams as heat source. Fouling takes place primarily on the spinbath side of the heat exchangers due to the impurities mentioned above. Figure 5 shows a fouling layer on the spinbath side of a heat exchanger plate.

![Figure 5: Fouling layer on the spinbath side of heat exchanger plate](image)

The results of the fouling are the same as for the evaporators. The heat transfer is hindered by the fouling layer and equipment performance degradation in form of higher heat consumption can be observed.
4.2 Equipment degradation modelling

One of the problems with equipment degradation in process industries is that most of the times it is not possible to directly measure the degradation state of equipment during production. Even if the performance of the equipment is directly measurable, the specific performance degradation caused by fouling is not. Since specific equipment performance depends not only on the state of equipment degradation but also on input parameters, load and ambient conditions. The key concept of the method for equipment degradation modelling presented in this report is the ability to compare and contrast equipment performance in the clean condition (day-zero condition) against performance in the current degraded condition during normal operation.

The workflow for the equipment degradation modelling of the evaporators and plate heat exchangers used in Lenzing is illustrated in Figure 6. The rest of the chapter reviews each step of this workflow. Each step will be explained in general first and then a selected example for this step from actual equipment in the Lenzing use case will be described.

![Workflow for equipment degradation modelling](image)

**4.2.1 Data pre-processing**

The most important and time-consuming step in equipment degradation modelling is the pre-processing of the data. In this step, the relevant raw data for the selected equipment is collected, filtered and modified before it can be used in any sort of modelling. The collected dataset can contain raw measurement data from the equipment (temperatures, flows etc.), process information
Deliverable 2.5
Final report on equipment degradation modelling

(product type, location etc.) and calculated or derived data from the first two sections (operation state, load, transmission coefficients, specific consumptions etc.). Whether or not data is relevant for the equipment degradation is up to the specific equipment and usually human analysis and decisions based on proper process knowledge is necessary. There are many reasons for data not being relevant like:

- Gross measurement errors
- Errors in the procedure used to obtain the data
- Data that covers operation points outside of the scope

Getting rid of the non-relevant data is crucial for the model generation in the subsequent steps. A model can only be as good as the data used to train it. Following procedures can detect and eliminate irrelevant data:

- Filtering data that is out of the relevant operating range based on expert knowledge
- Removal of outliers
- Steady-state detection
- General filtering operations
- Selecting subsets of data based on clustering algorithms

It is important to keep in mind that the removal of irrelevant data will most likely restrict the validity region of the later generated models. This means that the equipment degradation model can only give valid predictions as long as the equipment is operated in the relevant operation window. Model extrapolation to regions not included in the training dataset will most likely produce poor predictions. The modification can include a variety of statistical and signal processing methods like averaging, smoothing, centering, de-trending and scaling [11].

---

Example of a Lenzing plate heat exchanger

In Figure 7 raw time series data of a specific plate heat exchanger used to reheat spinbath liquid (heat sink) over a period of two month is depicted. For reasons of clarity, only the measured flows and the calculated heat transmission coefficient are depicted. The raw data, especially the flow of the heat source and the resulting heat transfer coefficient, contains lots of noise. There are several areas where the inflow is very low which can be classified as outside of the relevant operating range and which often can be caused by shut downs. Furthermore, single outliers can be found over the entire dataset stemming from measurement errors or instable operation points. Datasets such as depicted in Figure 7 are of no proper use for training a degradation model.

In Figure 8 the same dataset is displayed after data pre-processing. The pre-processing steps are performed by an in-house developed algorithm in MATLAB. First the raw data is divided into 30-minute-intervals. Every single measurement within these intervals is checked if the measured flows and temperatures are within certain on process knowledge based ranges to make sure only data within a relevant operation window is selected and shut downs are neglected. For the steady-state detection, first the average flow rate of the heat source of all data points within an interval is calculated and the deviation from this average for every data point is determined. The heat source flow rate is selected since it is the main control variable for the plate heat exchangers. For a data point to be in a steady state, only a specified maximal deviation from the calculated average is allowed. The maximal allowed deviation is based on process knowledge and is individually defined.
for each equipment. If all data points within an interval are in steady state and inside the selected operating window, all points are condensed to a single 30 minute average data point. If not, the data points are dismissed and the next interval is checked.

Figure 7: Raw data of flow measurements and calculated heat transfer coefficients of a plate heat exchanger before data pre-processing

Figure 8: Data of flow measurements and calculated heat transfer coefficients of a plate heat exchanger after data pre-processing
4.2.2 Data selection and clean model generation

In order to define equipment degradation, it is important to define equipment performance in day zero condition. In this condition, the equipment has not yet suffered from degradation. Day zero conditions can only be found after equipment is cleaned or maintained or when new equipment is put into operation. Information about when cleanings or maintenance were performed can be found for instance in resource planning software like SAP or directly in the distributed control system (DCS). Cleaning times can be detected directly from the measurement data via pattern recognition, too. A precise detection of the time when equipment cleaning or maintenance has finished and normal operation starts again is important. Depending on the equipment and the degradation mechanism, it can be assumed that day zero conditions last for a certain period. In this period, no signs of equipment degradation can be found (see Figure 2). The decision how long equipment stays in day zero condition requires knowledge about the affected equipment and process.

Afterwards, the pre-processed dataset can be divided into a clean (zero day) set and a dirty (degraded) set. For a better model validity, taking clean data from several cleaning or maintenance cycles is recommended to cover an operating window as broad as possible within the resulting data. The clean dataset is split into a part used for training (fitting) of the clean model and into a part used for validating the model. The split is random and a typical ratio of the size of the training and validation sets is 70%/30%.

Models used to predict the clean equipment performance can be either first principles models or any sort of statistical models. In cases where the equipment performance can be defined by a single performance value, Multiple-Input-Single-Output model structures are used. The actual model function can be either linear or nonlinear and depends on the process and the available input data. For model parameter estimation least squares approximation or genetic algorithms can be used. The model quality can be assessed by evaluating the generated model on the part of the dataset selected for validation. Here, a measure for the accuracy of the model is calculated by means of a scoring function and the decision whether a model predicts the data well can be made. A list of commonly used scoring functions can be found in [11].

Example of a Lenzing plate heat exchanger

For the Lenzing plate heat exchangers a pattern recognition algorithm is used to identify the cleaning times and the results are manually validated with data from the SAP system of the company. It is assumed that fouling effects on the equipment performance can be neglected up until 24 hours after the end of a cleaning procedure. The splitting of the pre-processed dataset into clean data and by fouling affected (dirty) data is carried out by a selection algorithm in MATLAB. In Figure 9, the same dataset from 4.2.2 is depicted and three clean data intervals after cleanings are highlighted. The obtained clean data points after several cleanings are split random into a section used for training the clean model and into a section used to validate the clean model.
For the plate heat exchangers used in the Lenzing use case the key equipment performance value is the heat transfer coefficient ($HTC$). Since the $HTC$ is highly affected by the flow conditions on both sides of the heat exchanger plates, the heat source flow rate and the heat sink flow rate are the model inputs. The following polynomial regression model is used to obtain the clean heat transmission coefficient for varying flow rates in the plate heat exchanger:

$$HTC = t_1 + t_2 \cdot f_{\text{hot}}^2 + t_3 \cdot f_{\text{hot}} \cdot f_{\text{cold}} + t_4 \cdot f_{\text{hot}} + t_5 \cdot f_{\text{cold}}$$

In (2) $f_{\text{hot}}$ is the flow rate of the heat source, $f_{\text{cold}}$ is the flow rate of the heat sink and $t$ are the model parameters. For the model parameter estimation the adaptive Nonlinear Least-Squares solver “NL2SOL” of the OPTI Toolbox for MATLAB is used. In Figure 10, the resulting non-linear regression model as well as the clean data points used for training and validation are illustrated. The goodness of the fit ($R^2$) in the illustrated example is around 90% and the relative Mean-Square-Error is below 3%. The selected clean data covers a broad operation range concerning the flow of the heat source but only a small range of the heat sink flow. The data indicates that the heat exchanger is mostly operated with a heat sink flow between 144.5 m$^3$/h and 143.0 m$^3$/h. It is important to note that the use of model predictions outside of observed operation window will probably lead to poor results. According to the generated clean model an increase of the heat source flow will lead to an increase of the $HTC$. This behaviour is coherent with the physics behind the $HTC$ since an increase of flow velocity usually results in an improved heat transmission.
4.2.3 Degradation monitoring

The base principle used to for the equipment degradation monitoring presented in this report is to compare the equipment performance ($P_t$) at any given time ($t$) with the possible equipment performance ($P_{clean}$) for the same operating point without the negative effects of degradation. The effects of equipment degradation are assumed to be responsible for any sort of discrepancy between those two stages. Therefore, the state of equipment degradation ($SED$) is defined as:

$$SED = 1 - \frac{P_t}{P_{clean}}$$  \hspace{1cm} (3)

The counterpart to the $SED$ is the state of equipment health ($SEH$) and is defined as:

$$SEH = \frac{P_t}{P_{clean}}$$  \hspace{1cm} (4)

The equipment performance $P_t$ can be either directly measured or can be derived from all sorts of measurement data. Whereas the corresponding clean equipment Performance $P_{clean}$ is the output of the clean model generated beforehand using the relevant input parameters measured at the same time. Since the calculation of the $SED$ and the $SEH$ contains the predictions from the clean models it is important to stay in the range of validity of the models. Calculations performed for operation points outside of the model range will lead to poor results for the $SED$ and $SEH$.

By calculating the $SED$ and $SEH$ for the pre-processed dirty data set a description of the degradation behaviour over time for several cleaning cycles can be obtained while at the same time the range of validity of the model is not violated. The computational effort necessary for the calculations is low. Hence, the degradation monitoring can be performed in real-time.
Example of a Lenzing plate heat exchanger

In Figure 11 the heat transmission coefficient of a plate heat exchanger is depicted over several months and several cleaning cycles. The figure contains the actual (dirty) HTC obtained from measured data as well as the corresponding clean HTC predicted by the generated model in 4.2.2 and the times when cleanings have been performed.

![Figure 11: Heat transmission coefficient of a Lenzing plate heat exchanger obtained from actual measurements affected by equipment degradation (blue) and predicted by the clean model (red) and cleanings (green)](image)

Figure 11 shows that directly after cleaning there is no difference between the measured and the predicted HTC, since the data points in this area have been used to fit the model. The discrepancy between the prediction and the measured HTC is growing with the time since the last cleaning was performed. It is assumed that the equipment degradation due to fouling is responsible for this development. The measured HTC decreases with the time since the last cleaning but a lot of noise and different plateaus can also be observed.

In Figure 12 the state of equipment degradation as well as the state of equipment health is illustrated for the same timespan as in Figure 11. The trend of the SED shows a steady increase with the time since the last cleaning starting around 0% after a cleaning and ending between 20% and 30% before the next cleaning is performed. As expected, the trend of the SEH shows an exact contrary behaviour to the SED starting at 100% after a cleaning and ending between 80% and 70%. Based on these observations it can be determined that the specific plate heat exchanger loses up to 30% in equipment performance over a period of a cleaning cycle due to fouling. The data also shows that after a cleaning was performed the negative effects of fouling disappear and the equipment performance recovers to its initial level.
4.2.4 Degradation modelling

For an efficient maintenance strategy, only monitoring the equipment degradation is not sufficient. Plant managers and supervisors need the ability to predict the degradation behaviour of their equipment in advance to set up a proper maintenance scheduling. The necessary prediction horizon depends on production constraints, extent of the scheduling task and the resources necessary for maintenance. Predictions can be made over general time, operation time, production volume or any other relevant reference value depending on the actual process equipment and degradation mechanism.

To set up a suitable equipment degradation model, the obtained dataset of the SEH (or SED) from the degradation monitoring over several cleaning cycles is used. In a pre-processing step the entire dataset is extended by an additional value containing the reference value that is relevant for the prediction. The corresponding value at the end of the last cleaning is thereby the reference point for each data point. In case there are different types of cleanings and the SEH after a cleaning depends on the cleaning type, an additional case distinction is necessary to find a reference point. With the new obtained information, the data of different cleaning intervals can directly be compared to each other.

The extended dataset is split into a part used for training (fitting) the degradation model and into a part used for validating the model. The split is performed as explained in 4.2.2. Models used to predict the equipment performance degradation are often regression models with Single-Input-Single-Output model structures describing the SEH (or SED) as a function of the reference value. The actual model function can be either linear or nonlinear and depends on the equipment, the degradation mechanism and the available input data. The model quality can be assessed by...
Deliverable 2.5
Final report on equipment degradation modelling

Evaluating the generated model on the part of the dataset selected for validation. Here, the accuracy of the model is calculated in a scoring function and the decision whether a model predicts the data well can be made.

Example of a Lenzing plate heat exchanger

In the SEH of a plate heat exchanger as a function of the timespan since the last cleaning for several cleaning cycles is illustrated. As there is only one type of cleaning for the plate heat exchangers used in Lenzing, the results of the different cleaning cycles are directly comparable and no additional case distinction is necessary. The timespan since the last cleaning is selected as relevant reference value in Figure 14. Results obtained when selecting the operation time or the amount of heat transferred since the last cleaning show a slightly higher variance. Based on these studies it can be assumed that the fouling of the plate heat exchangers is most of all affected by the timespan since the last cleaning and less on the actual mode of operation. The results indicate similar to Figure 12 a clear decrease of the SEH with increasing timespan since the last cleaning. Clustering of the results from different cleaning intervals cannot be observed. The results range up to a timespan of 55 days since the last cleaning and the spread of the results is sufficient.

![Figure 13: State of Equipment Health (SEH) as a function of the timespan since the last cleaning for several cleaning intervals](image)

For the modelling of the degradation as a function of the timespan since the last cleaning, the obtained results of the SEH are again split into a training dataset and into a validation dataset. The solver “NL2SOL” from the OPTI Toolbox for MATLAB is used to determine the best model architecture and fit the model parameters. The following second degree polynomial equation scored the highest goodness of fit for the development of the SEH over the timespan since the last cleaning:

\[ SEH = a \times t^2 + b \times t + c \]
Deliverable 2.5
Final report on equipment degradation modelling

\[ SEH = m_1 z_c^2 + m_2 z_c + m_3 \] (5)

In (5) \( z_c \) is the timespan since the last cleaning and \( m \) are the model parameters. The resulting curve fit as well as the data points used for training are depicted in Figure 14.

As indicated in Figure 14, the decline of the SEH and therefore the equipment degradation mechanism itself slows down over the timespan since the last cleaning. The growth of the fouling layer inside the plate heat exchangers leads to higher flow velocity and thereby the fouling layer itself experiences more sheer stress. This process favours the fouling removal phase and is responsible for the flattening of the SEH curve. It can be assumed that at longer cleaning intervals the state of equipment health or degradation reaches a certain level where an equilibrium between fouling and fouling removal is established and no further changes can be observed.

The goodness of the fit (R²) of the illustrated model is around 78% and the validated relative Mean-Square-Error is at 3%. At first a R² of 78% seems a relative poor fit but considering the measurement uncertainty as well as that data from several cleaning cycles over a period of nine month is used, the fit is satisfying enough. Also for the cleaning scheduling a prediction of the SEH on a daily basis is well enough since the cleanings are planned on a daily to weekly basis.
4.3 Outlook

The presented methodology for the equipment degradation modelling in chapter 4.2 allows Lenzing to monitor and predict the degradation of its evaporators and its plate heat exchangers caused by fouling. The process has been applied to a number of different evaporators and heat exchangers and will be fully rolled out to all relevant equipment during the next couple of months. The information about the current state of equipment health and the prediction of the future state will be used to optimise the cleaning scheduling of the affected equipment. Since the cleaning capacity is limited by production and resource constraints, the results of the predictions from the degradation modelling will in the first place help to set up an efficient cleaning sequence for the redundant equipment. In this sequence, the equipment that has the highest SED will receive the next cleaning. In the next phase, the predictions about the performance degradation development will be combined with cost functions and prices for the cleaning. The goal here will be to find the cleaning cycle interval with the lowest operational cost for every equipment. These operational costs will contain the cost for cleaning as well as the costs for additional energy consumption caused by the specific equipment performance degradation. In Figure 15, an example of the different cost functions and the optimal cleaning cycle interval for an evaporator are illustrated.

![Figure 15: Exemplary cost functions and optimal cleaning cycle interval of an evaporator](image-url)
5 INEOS use case

INEOS in Köln operates a typical integrated petrochemical complex, processing mainly naphtha in steam crackers as major feedstock to produce a large number of important base chemicals.

5.1 Degradation of the cracker

The degradation of the cracker is best described by the amount of coking in the pipes. While the crackers run, coke as a waste product sediments on the inside of the coils and reduces their diameter. To derive a measure for the diameter the Darcy-Weisbach equation is used

$$\Delta p = \lambda * \frac{L}{D} * \frac{\rho}{2} * \frac{V^2}{A}, \quad (6)$$

where $\Delta p$ is the difference between input and output pressure, $\lambda$ flow coefficient, $L$ length, $D$ is the inner coil diameter, $\rho$ is the density of the naphtha, $V$ is the volume flow rate and $A$ the cross sectional area of the coil.

With transposing this equation and neglecting constants, the following coking index which describes the change of the diameter can be derived:

$$coking \ index = \frac{\Delta p}{\frac{V^2}{2} \rho} \quad (7)$$

5.2 Degradation modelling

Next to the development of a coking index the naphtha concentration was focused. The composition of naphtha influences the rate of the coke formation and therefore should be considered for the prediction of the coking index. There are no online measurements of the naphtha feed composition entering the cracker, and the only available information about the decomposition is given by the delivery notes. There are three different naphtha sources which fill up the tanks in the INEOS use case: batches, ship deliveries and a recycle stream from the production plants. The naphtha is stored in five different tanks, where four of them feed a sixth one (see Figure 16). Partly, the tanks can be refilled and emptied simultaneously. There are several challenges during the load mapping process due to inconsistent noted arrival and refilling dates, approximations in the mass balances and measurements which are either given as mass or volume flows. An approximate naphtha composition was derived, which was close to some early measurements in 2017 but proven to be inconsistent with measurements taken in 2018.
Due to this inconsistency the naphtha composition was not included as an input to the coking index model, and the density of the naphtha as an indicator of its composition was used, for which an online measurement is available. There are seven different input parameters remaining: amount of pre sulfur, steam, density of the naphtha, cracked gas volume flow, sulfur feed, naphtha feed and cumulative naphtha feed. A total of 27 coking cycles (see Figure 17) were observed. For the modelling preparation, several different data pre-processing methods have been tried: from smoothing techniques like sliding mean to change based methods like using only gradient values. With these pre-treated data sets multiple modelling approaches were tested, because of the fast calculation, simplicity and extrapolation properties a linear model was chosen. Furthermore, different linear model modifications were tested, e.g., the use of all coking cycles as input, the use of variable interactions up to cubic terms and reducing the model again with Akaike information criterion as well as the use of the starting value of each coking cycle as the intercept. While increasing the complexity of the models, no mentionable improvement could be found by applying these techniques, because the improvement was in no relation to the limitations which came along with each model specification. Therefore, only two approaches will be described, which use different training and test data settings, which were found most convenient for the modelling approach and the available data was used without further pre-processing.
The first one uses only the data from one coking period each. While ramping up the furnaces, the data from the first 48 hours was collected and a linear model was trained on this data set. With this model the remaining runtime is predicted. After the first 96 hours the model was retrained on all data which was observed till this time point and again the whole time horizon was predicted. With evolving time the prediction accuracy gets better (see Figure 18). Therefore, information gained from previous sections was used to predict the next instances in a second approach. A rolling horizon model was tested, where the data from maximal 2160 previous hours was used to learn a model and the next two days were predicted based on this model. Afterwards, the model was retrained again and data earlier than three months past is discarded. Figure 19 shows the moving horizon approach and Figure 20 shows the result for the intersection 26. The prediction value is always underestimating the coking index. This suggests that the slope of the coking index learned on data of a previous coking period is not necessarily meaningful for the current period. This might be because of multiple unmeasured effects, like small damages in the coil walls due to strong decoking or because of different naphtha decompositions. Since there is not more information available no further root cause analysis is applicable.

5.3 Conclusion

In the INEOS use case it was tried to monitor and model the degradation of a naphtha cracker caused by cooking. Two different approaches for the degradation modelling were tested. The first one using data from one coking period each and the second one using a rolling horizon model. In both approaches the obtained results were not accurate enough to predict the future degradation behaviour. The lack of significant information like the naphtha decomposition is considered to be the cause for the weak prediction accuracy.
Figure 18: The intersection 26 with the calculated coking index (black) and the prediction based on previously observed data.

Figure 19: Moving horizon approach. The data used for learning a model is extended until the data of three months is used. Afterwards, with each new observation the oldest observation is neglected.

Figure 20: Moving horizon approach for the same coking cycle used in Figure 17. The prediction results are worse than before.
6 P&G use case

P&G is the world largest producer of consumer goods. P&G sells products under more than 50 different brands such as Pampers, Tide, Dash, Always, Swiffer. All these products are produced in more than 100 plants around the world with 100,000+ employees. P&G serves 5 Billion people every day. This work focuses on European plants and more precisely on plants in France with potential to be extended to the entire P&G network.

The focus in CoPro is on the laundry detergent, more precisely on the production of the washing pouches. Two different use cases arise in the CoPro framework: the optimization of the operation of the production lines with respect to time and production capacity efficiency and related to this, the prediction of necessary maintenance actions. While the first is described in deliverable 4.3 the latter is part of the equipment degradation task and further described in this document.

6.1 The maintenance use case

Maintenance actions require a production stop which is currently planned at fixed time points. If unscheduled production stops due to equipment degradation are necessary, not only the maintenance action takes time, but also the root cause analysis. Therefore, the goal is to predict the health state of components to be able to generate an optimal production plan and to schedule maintenance actions as needed. The production is conducted on multiple lines which have approximately the same setup but can produce independently of each other different products. In each line, there are multiple modules which are again composed by two top and one bottom filling unit which are run by separate pumps each.

The available process data consists mostly of motor data from the pumps which are used during the filling process of the pouches’ compartments. Each filling data depends on the current line speed but in general one filling cycle happens in the area of milliseconds. The frequency at which the data is collected goes from every 5 minutes down to one second.

6.2 Methods for developing a maintenance indicator

To avoid unnecessary maintenance and production stops, various approaches for predictive maintenance were tested. Since no direct indicator measurements about the process quality are available, e.g. no rejection rate or measured abrasion of the machinery, this task needs to be conducted as semi-unsupervised – semi, because the failure times are known. Different approaches were conducted: first, based on a classification approach, where the underlying assumption is that a portion before a failure is already suspicious and therefore labelled as not ok and data before this time window is labelled as ok and second based on unsupervised anomaly detection. While the first approach showed no satisfying results due to missing information in the data, the second approach seemed promising and was therefore extensively tested. The underlying assumption with unsupervised anomaly detection is, that close to a failure the occurrence of anomalies is more frequent than compared to the rest of the runtime. To account for the multidimensionality in the data, a principle component analysis was used to derive the top principle components which were used to transform the time series accordingly. A regression model was learned on these composed time series where variables of a pre-defined window served as inputs and the output was the next instance. If the deviation from prediction to expected value passes a certain threshold based on the Rosner test, these instances are labelled as anomalies. Figure 21 shows an exemplary first principle
component with anomalies marked in red. The accumulation of detected anomalies is not increasing at the end – which marks the failure timestamp – but it is high during high values of the first principle component.

![Figure 21: The transformed time series according to the first principle component over the course of 3 days in black with automatically detected anomalies marked as red lines.](image1)

But because of the data availability and not listed changing production actions this as well lead to non-satisfying results. Another approach is based on the assumption that the actual filling curve can be derived from the currently observed data which is based on a standard curve/pattern. In Figure 22

![Figure 22: Multiple cycles of the actual velocity layered above each other to identify a standard curve.](image2)
multiple curves for the actual velocity layered above each other to derive a standard curve are shown. The smoothed deviation from the actual measured curves over time compared to the standard curve is shown in Figure 22.

![Smoothed Error](image)

**Figure 23: Smoothed deviation from the standard curve over the course of 3 days. It seems that during this period there are not just one but at least two standard curves.**

There are two different levels visible. This effect is currently further investigated with higher frequency data and added product information. Assumingly, more than one standard curve would be needed due to different production settings.

### 6.3 Outlook

Summing up, the ability to reconstruct a high frequency process based on a suited sampling rate is crucial for this task. The process needs to be reflected by the given data which is currently not the case as the Nyquist-Shannon sampling theorem [12] implies. Therefore, the main work is focused on improving data quality which includes investigating an alternative data collection method.

### 7 References


