Application of a Predictive Maintenance Strategy Based on Machine Learning in a Used Oil Refinery

Francesco Negri[a,b], Andrea Galeazzi[b], Francesco Gallo[a], Flavio Manenti[b,*]

aItelyum Regeneration S.p.A., Via Tavernelle 19, Pieve Fissiraga 26854, Lodi, Italy
bPolitecnico di Milano, CMIC Dept. "Giulio Natta", Piazza Leonardo da Vinci 32, Milan 20133, Italy
*flavio.manenti@polimi.it

Abstract

The Itelyum Regeneration used oil re-refining plant in Pieve Fissiraga currently employs a condition-based maintenance strategy for its thermodeasphalting (TDA) section, particularly focusing on the TDA T-401 column. This strategy involves monitoring the real-time pressure differential (ΔP) between the column's top and bottom, which increases in time due to fouling phenomena. Maintenance is scheduled when ΔP exceeds a predetermined empirical threshold, ensuring that the T-401 column operates within normal operations limits. However, this approach has limitations with non-conventional used oils. To address this, a data-driven machine learning algorithm, previously successful in predicting key performance indicators of the PH-401B furnace in the TDA section, was applied to the T-401 column datasets. This algorithm, based on Gaussian Process Regressions, effectively predicts the evolution of ΔP and reduces the time during which T-401 operates in suboptimal conditions. The implementation of this machine learning approach marks a significant improvement in the maintenance strategy, shifting from a static, condition-based approach to a dynamic, predictive one, thus ensuring more efficient and reliable operations, even with non-conventional used oil.

Keywords: Data-driven, Machine learning, Predictive maintenance, Thermodeasphalting, Used oil.

1. Introduction

Maintenance is a crucial aspect of every industrial plant to ensure continuity of operations and safety of the workers. Typical approaches in the European industrial context are corrective, preventive, opportunistic, condition-based, and predictive maintenance (Bevilacqua and Braglia, 2000). The combination of a vacuum distillation column and its feedstock fired heater are critical pieces of equipment in crude oil refineries and used oil re-refineries that suffer from fouling due to the characteristics of the heavy hydrocarbon feed, thus requiring careful maintenance planning (Fuentes et al., 2007; Morales-Fuentes et al., 2014). One of the most robust and efficient processes for the regeneration of used oil is based on the patented Revivoil® technology, and is currently operated in the Itelyum Regeneration re-refining facility in Pieve Fissiraga, Lodi, Italy (Gallo, 2016). The maintenance approach in the case of the TDA section of the process is typically condition-based, using the pressure differential across pieces of equipment as a sentinel key performance indicator to be monitored. Maintenance is planned once the parameter overcomes a warning threshold value. This static approach is typical for the refining industry, where fouling is an ever-present problem. Data-driven approaches have shown good results in modeling fouling in refinery equipment such as heat exchangers, with

better fitting compared with equation-based, mechanistic modeling approaches, which instead often show poor results due to the extremely complex and partially random nature of fouling phenomena (Mei et al., 2023). This work uses said data-driven algorithms to develop a dynamic predictive maintenance strategy that is more effective in avoiding run-time within suboptimal operating regions and it is more adaptable to changes in feedstock composition, a typical situation for used oil waste.

2. Materials and Methods

2.1. Current Maintenance Strategy

The thermodeasphalting (TDA) section of the Revivoil® process works by fractionating dehydrated used oil in a vacuum column. The main products from the TDA T-401 column are three semi-finished base lube oil cuts. The TDA column starts every run with a project-specified pressure differential between top and bottom which is characterized by a dynamic evolution in time, as shown in Figure 1. The graph reports normalized values to preserve industrial secret, which have been cleared of gross errors by applying a robust methodology (Manenti and Buzzi-Ferraris, 2009). The vertical axis shows the pressure differential (ΔP) normalized with respect to the project value, which is the one observed at every start-of-run. The horizontal axis shows the length of run normalized so that the longest run is divided into ten equally and arbitrarily long time periods (τ). Figure 1 shows two curves, each representing a run in which T-401 was processing a certain type of used oil. Used oil type "A" is representative of a typical used oil that is treated in Pieve Fissiraga facility. Used oil type "B" is a non-conventional type of used oil, more volatile than type "A", for which an industrial test was carried out to understand the possible impact on the existing re-refining technology. Both curves show an evolution in time in which the ΔP gradually increases, firstly at a relatively slow pace, followed by a sudden and rapid degeneration, which is a typical fouling behavior for distillation systems working with heavy and unstable hydrocarbon feedstock (Seegulam et al., 2017).
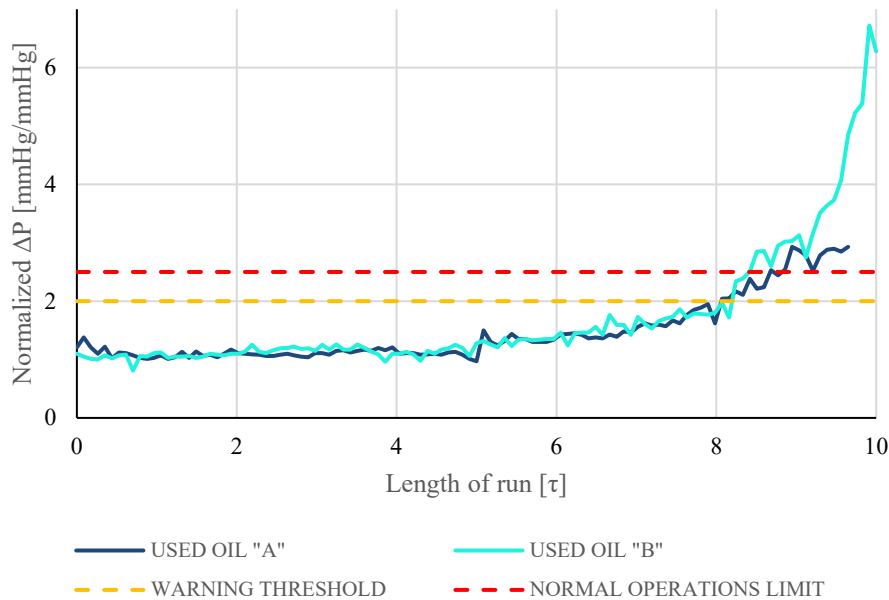

Figure 1. Normalized pressure differential in the TDA column vs time for two types of used oil

The evolution of the ΔP in time may be explained by an initial, gradual reduction of the effective cross sectional area in the column packed beds, which corresponds to a gradual increment in the overall pressure differential. When the ΔP approaches the flooding value in one (or more) of the beds, gas and liquid fail to flow correctly through the column, eventually leading to a unit shutdown (Rocha et al., 1993). The graph in Figure 1 also shows two important thresholds. The first is the warning threshold of 2 mmHg/mmHg, which is the value of normalized ΔP at which, according to the historical condition-based strategy, maintenance must be scheduled. This threshold has been determined by a statistical analysis on the ΔP evolution in time with typical used oil (type "A"). The second is the normal operations threshold of 2.5 mmHg/mmHg, which is the value at which T-401 starts to perform poorly due to suboptimal hydraulic working conditions, thus it is independent from the type of used oil that is being processed. The strategy works well for the case with conventional used oil of type "A", in which T-401 works in an inefficient operating region only for a short amount of time, roughly equal to 0.8τ. However, the strategy proved to be quite poor for the case with the non-conventional used oil of type "B". The different composition of the oil lead to an unexpected fast worsening of the column conditions, with a significant amount of T-401 run time inside the inefficient operating region, roughly equal to 1.6τ.

2.2. Data-driven Algorithm for Predictive Maintenance

Machine-learning-based frameworks have already been successfully implemented in the domain of Pieve Fissiraga re-refinery. Some examples of this are the performance prediction of the PH-401B furnace of the TDA section (Galeazzi et al., 2023a) and the development of a surrogate model for the amine scrubbing section of the plant (Galeazzi et al., 2023b, 2022). The data-driven algorithm shown in the work by Galeazzi et al. (2023a) was already successfully applied to forecast the pressure differential across the PH-401B furnace, which has a dynamic evolution in time due to fouling phenomena. Given the data-driven nature of the algorithm that requires no physical description of the fundamental phenomena governing a given process unit, it is possible to apply it on the T-401 datasets previously shown in Figure 1, to understand the forecast capabilities of the method and the possible implications of its usage inside a predictive maintenance strategy. The algorithm performs a regression of real plant time-dependent datasets by using two different methodologies, a polynomial regression and a Gaussian Process Regression (GPR), then performs future forecasting of the possible evolution of the dataset. The general polynomial relationship between ΔP and time is shown in Eq. (1), where θ are the features of the model, n_P is the selected degree of the polynomial, t is the time, and ε is the residual normally distributed error (Galeazzi et al., 2023a).

$$\Delta P(t) = \sum_{i=0}^{n_P} \theta_i t^i + \varepsilon \qquad (1)$$

The GPR method uses Gaussian Processes to model the dataset. According to this methodology, the time-dependent dataset is modeled as a function of some unknown functions of time f(t), and to specify a prior distribution over each function f(t) that is Gaussian (Galeazzi et al., 2023a). The advantage of this type of approach is that it is possible to model complex relationships in the data without specifying model parameters (Rasmussen and Williams, 2008), a feature that is ideal for modeling a complex phenomenon such as fouling in the vacuum TDA T-401 column. The functional form of

a Gaussian Process expressing ΔP as a function of time is shown in Eq. (2), where m(t) and k(t,t') are respectively the mean and covariance functions (Galeazzi et al., 2023a).

$$\Delta P(t,t') = GP\{m(t), k(t,t')\} \tag{2}$$

The covariance functions included in the method (also referred to as kernels) are the Linear (LIN), the Radial Basis Function (RBF), and the Rational Quadratic (RQ) kernels. The optimal solution for the regression is found through a greedy search in which the kernels are combined through specific operators by following the local optimum at each combination step rather than finding the global optimum, due to computational constraints (Galeazzi et al., 2023a).

3. Results and Discussion

3.1. Application of a Novel Predictive Maintenance Strategy

The aforementioned data-driven algorithm is applied to the dataset shown in Figure 1. Considering a horizontal time axis divided into 10τ periods as shown previously in Figure 1, it was found that an acceptable compromise between prediction accuracy and length of run to be fed to the algorithm is a dataset with a length of 7τ. Forecast accuracy is shown in Figure 2. It is possible to notice that the most probable evolution in time predicted by the algorithm (the orange dashed line in the middle of the green-shaded prediction area) almost never precisely follows the real data from the plant. This is to be expected due to the complexity of the situation that is being analysed, which consists of a gradual increment in ΔP due to fouling that gradually modifies the hydraulic behaviour of the T-401 column, until an unfeasible operating region characterized by flooding is approached. The abrupt change from a feasible to an unfeasible hydraulic operating region translates effectively in an abrupt change in the behaviour of the ΔP as a function of time. The data-driven nature of the algorithm does not allow it to model the future change in the behaviour of the ΔP if no signs of this change are present in past data. Despite this, it is interesting to notice that the natural variability of the dataset generates a 95% confidence interval for the prediction which almost in all cases, shown in Figure 2, fits the real data up to the normal operations limit of 2.5 mmHg/mmHg. This can be explained by observing that the time period of 7τ includes both a "clean" operating condition which oscillates around the start-of-run value of ΔP, and a later different operating condition in which the ΔP starts increasing at a certain rate due to fouling phenomena. This effective change in the underlying physics generates a variability in the data that is noticed by the method, which then proposes a more conservative confidence interval for the forecast. The reason why the algorithm overall predictive capabilities are remarkably different between type "A" and type "B" oil is due to the unexpected early occurrence of the unfeasible hydraulic working condition for type "B" oil, with no past data showing the sign of this change. A similar behaviour would be shown for oil "A" as well, if enough time was given to the system. Defining a robust and adaptable strategy for maintenance in this domain is crucial. One may define a predictive-maintenance approach based on ΔP forecasting combined with a proper reference threshold, namely the normal operations limit. In this case, a dataset having a length of 7τ is used as input, and the algorithm is rolled daily. Daily rolling stops when the 2.5 mmHg/mmHg normal operations limit falls within the 95% forecast confidence interval. This threshold is used since it is representative of T-401 hydraulics and thus it is independent from the oil type. Maintenance can be scheduled for the day in which it is forecasted with a 95% confidence interval that the 2.5 mmHg/mmHg threshold will be overcome. Table 1 reports T-401 run time in the suboptimal operating region above the normal operations limit under different

strategies. It is possible to see that in three cases out of four, the algorithm allows the length of the suboptimal operating region to be reduced down to zero. In those three cases, the algorithm contains within the 95% confidence interval the moment in which T-401 overcomes the normal operations limit, meaning that the prediction can be reliably used to schedule the maintenance much in advance compared to what usually happens with the condition-based approach. The only case in which the prediction does not provide benefits is the polynomial regression for used oil of type "A". It is interesting to notice that Gaussian Process Regressions proved to be effective both with type "A" and type "B" oil.
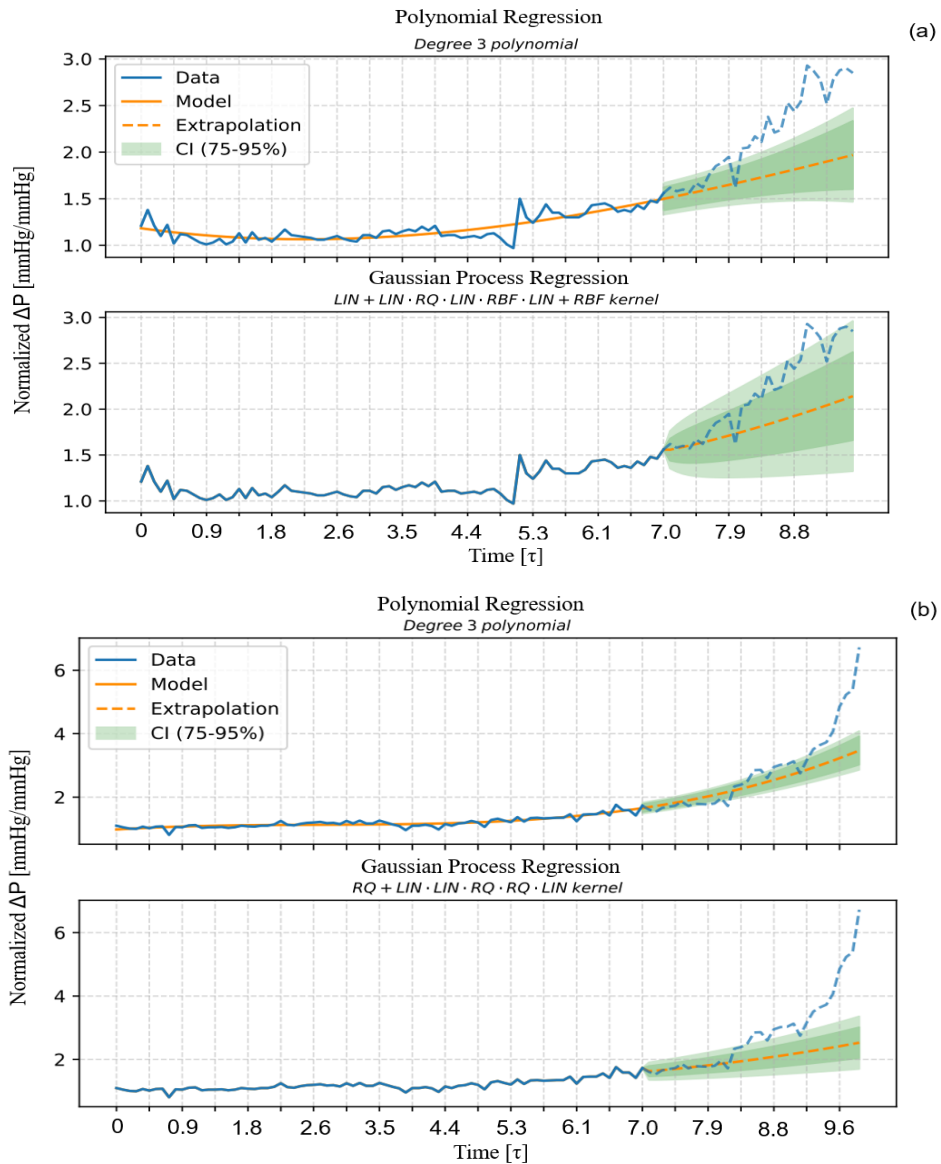

Figure 2. Forecasts on the ΔP time profile for used oil of type "A" (a) and type "B" (b)

Table 1. T-401 run time in suboptimal operating region under different strategies

Used oil type	Condition-based approach	Polynomial predictive approach	GPR predictive approach
Type "A"	0.8 τ	0.8 τ	0.0 τ
Type "B"	1.6 τ	0.0 τ	0.0 τ

4. Conclusions

The work has shown the limitations of a conventional condition-based maintenance strategy applied to a thermodeasphalting section in a used oil re-refinery. This approach is acceptable for typical used oil due to the consistency of its composition, but it fails when non-conventional, more volatile feedstock is processed, leading to long periods of run time inside a suboptimal operating region. It is possible to solve this problem by applying a data-driven algorithm to T-401 datasets. The algorithm predicts correctly the evolution in time of the pressure differential between top and bottom of the column up to the normal operations limit of 2.5 mmHg/mmHg, inside a 95% confidence interval. This allows to schedule the maintenance in advance and avoid altogether suboptimal run time periods. Gaussian Process Regressions have proven to be systematically better for this purpose, even with non-conventional used oil feedstock. Prediction capabilities of the algorithm may be upgraded in the future with hybrid modelling.

References

M. Bevilacqua, M. Braglia, 2000. Analytic hierarchy process applied to maintenance strategy selection. Reliability Engineering and System Safety 70, 71–83.

M.J. Fuentes, R. Font, M.F. Gómez-Rico, I. Martín-Gullón, 2007. Pyrolysis and combustion of waste lubricant oil from diesel cars: Decomposition and pollutants. Journal of Analytical and Applied Pyrolysis 79, 215–226.

A. Galeazzi, F. de Fusco, K. Prifti, F. Gallo, L. Biegler, F. Manenti, 2023a. Predicting the performance of an industrial furnace using gaussian process and linear regression: A comparison. Computers & Chemical Engineering 108513.

A. Galeazzi, K. Prifti, C. Cortellini, A. Di Pretoro, F. Gallo, F. Manenti, 2023b. Development of a surrogate model of an amine scrubbing digital twin using machine learning methods. Computers and Chemical Engineering 174.

A. Galeazzi, K. Prifti, F. Gallo, F. Manenti, 2022. A Methodology for The Optimal Surrogate Modelling of Digital Twins Using Machine Learning. Computer Aided Chemical Engineering 51, 1543–1548.

F. Gallo, 2016. Procedimento di rigenerazione di olii usati. ITUB20151298A1.

F. Manenti, G. Buzzi-Ferraris, 2009. Criteria for Outliers Detection in Nonlinear Regression Problems. Computer Aided Chemical Engineering 26, 913–917.

X. Mei, H. Kiyomoto, S. Kato, Y. Kansha, 2023. Data-Driven Soft Sensor for Crude Oil Fouling Monitoring in Heat Exchanger Networks. IEEE Sensors Journal 23, 26336–26346.

A. Morales-Fuentes, M. Picón-Núñez, G.T. Polley, S. Méndez-Díaz, 2014. Analysis of the influence of operating conditions on fouling rates in fired heaters. Applied Thermal Engineering 62, 777–784.

C.E. Rasmussen, C.K.I. Williams, 2008. Gaussian processes for machine learning, 3. print. ed, Adaptive computation and machine learning. MIT Press, Cambridge, Mass.

J.A. Rocha, J.L. Bravo, J.R. Fair, 1993. Distillation Columns Containing Structured Packings: A Comprehensive Model for Their Performance. 1. Hydraulic Models. Industrial and Engineering Chemistry Research 32, 641–651.

N. Seegulam, F. Coletti, S. Macchietto, 2017. Effect of Fouling on Control and Energy Recovery in an Industrial CDU Column. Computer Aided Chemical Engineering 40, 1555–1560.