16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety
Berlin, Germany, on September 3-4, 2019

Remaining Useful Life Estimation of Industrial Circuit Breakers by Data-Driven Prognostic Algorithms Based on Statistical Similarity and Copula Correlation

Michy Alice[1], Dejan Pejovski[2], Loredana Cristaldi[3]

[1,2,3]*Department of Electronics, Information and Bioengineering (DEIB),Politecnico di Milano, Milan, Italy*
[1]michy.alice@mail.polimi.it
[2]dejan.pejovski@mail.polimi.it
[3]loredana.cristaldi@polimi.it

***Abstract* – Predicting the future behaviour of an item or a complex system based on its past history is the aim of data-driven algorithms. In our paper, we present two algorithms for predicting the Remaining Useful Life (RUL) of industrial circuit breakers (CB) which make use of on-site collected data related to CB's health condition. In the first algorithm, a sub-fleet of CBs is identified by applying the two-sample Kolmogorov-Smirnov Test which relies on statistical similarity between the observations. Once chosen the sub-fleet, the algorithm attempts to exploit correlations between the variation of health condition and sampling time using copulas. The second algorithm models the correlation structure between the time at which a certain degradation level occurs and the item's End of Life (EOL). Both algorithms are used to estimate the item's Remaining Useful Life through the Monte Carlo method. The use of copulas attempts to exploit also the information on the correlation structure in the data in order to obtain a higher accuracy in the estimation.**

Keywords - Remaining Useful Life, Copula correlation, Data-driven prognostics, Health condition variation.

I. INTRODUCTION

In the recent years, condition monitoring (CM) has become available and cost effective for large sets of products; this allowed development of data-driven algorithms which are able to predict the health condition (HC) and the Remaining Useful Life (RUL) of a product. Data-driven algorithms are based on the collected run-to-failure times and do not deal with failure mechanisms, which makes them suitable for analysing systems with complex physical relations between the components. RUL of an item or a system, at a given time instant is defined as the remaining time interval in which it is able to fulfil its required function. Predicting the future behaviour of the product (in terms of HC and RUL) based on the ability to learn from its past history and from the past behaviour of similar products is an essential objective when aiming to reduce maintenance costs and increase the system availability [1].

No health condition is defined as "the extent of degradation or deviation from an expected normal behaviour" [2]. In the case of CB, HC refers to a component profile based on specific parameters which are monitored: degradation of the switching contacts, leakage of the interrupting chamber, SF_6 gas density, temperature of the interrupting chamber, etc.

Relevant scientific papers have been dealing with different aspects of the RUL estimation and maintenance decision making for critical components of the system. Using the fleet concept, in [3] simulation models are presented for maximizing the fleet reliability compared to the target reliability in order to optimize maintenance activities. The importance of fleet size is discussed in [4]: the paper deals with large scale problems reducing them to single items and calculating their reliability individually. An essential need for uncertainty analysis when estimating the RUL is described in [5]. Previous versions of the first algorithm developed in our paper are presented in [6]: the condition monitoring data of a fleet of a product is analysed by statistical methods to extract the usage and degradation profile of the product. This profile is then represented by statistical distributions used later to predict the behaviour of the component.

This paper is structured as follows: in Section 2 a methodology for identifying a suitable sub-fleet is briefly

described. In Section 3 two algorithms for RUL prediction are explained and illustrated on an industrial CBs dataset. Finally, in Section 4 the results are reported and analysed.

II. BASIC CONSIDERATIONS

As proposed in the relevant literature [1], a subset (i.e., sub-fleet) is selected among a given set of products (fleet) that show higher similarity in terms of observed degradation in time, with respect to the item whose RUL estimation is required. The sub-fleet identification is based on a statistical test for grouping those products which present a statistical distribution of their degradation rate similar to the target product. The two-sample Kolmogorov-Smirnov Test (KST) is used in order to decide whether the two samples are drawn from the same continuous statistical distribution or not, i.e., if they belong to the same sub-fleet [1]. The KST uses the maximum absolute difference between the distribution functions of the samples. In general, the test makes use of each individual data point in the samples, independently of their direction and ordering [7]. The confidence level α determines the selectivity of the test.

The information about the past usage of the product is reported as a time series of HC from the initial value of 100% up to 0%. The sampling time and variation of the HC can be calculated as the difference between two subsequent points of the monitored values. Degradation rate is the ratio between the HC variation and sampling time, for i=1, 2,…, n, where n is the number of monitored values [8]. These definitions can be initially used for each CB in the fleet and then combined to obtain vector representations for the whole fleet.

$$\Delta t_i = t_i - t_{i-1}, \tag{1}$$

$$\Delta HC_i = HC_i - HC_{i-1}, \tag{2}$$

$$d_i = \frac{\Delta HC_i}{\Delta t_i}. \tag{3}$$

The correlation structure between sampling time and health condition variation is determined by estimating the underlying copula. A copula is a function which joins (or couples) multivariate distribution functions to their one-dimensional marginal distribution functions, i.e., contains information about the correlation structure between random variables and, in general, can also capture nonlinear relationships [9]. Mathematically speaking, let us assume that F_x and F_y are the cumulative distribution functions of the random variables X and Y. Their joint distribution can be written as [10]:

$$\begin{aligned} &F(x,y) = P(X < x, Y < y) = \\ &P(X < F_x^{-1}(u), Y < F_y^{-1}(v)) \\ &= F(F_x^{-1}(u), F_y^{-1}(v)) = C(u,v) \end{aligned}, \tag{4}$$

where $F^{-1}_x(u) = x$, $F^{-1}_y(v) = y$ and u and v belong to the unit square. In eq. (4), $F_x(x)$ and $F_y(y)$ are the marginal distribution functions of random variables X and Y and C(·) denotes the copula function. The interesting aspect about copulas is that they allow to model separately the marginal distributions and the correlation structure.

III. DESCRIPTION OF THE METHOD

In this paper two algorithms are proposed in order to predict the RUL: the first one is based on KST for sub-fleet identification and copula correlation modelling in order to predict the HC vs time curve, while the second one completely relies on copula modelling in order to estimate the RUL. For both algorithms, the accuracy of the prediction is calculated as the ratio between the sum of correct predictions over the total number of CBs tested as proposed in [1].

A. Algorithm 1

Once chosen the sub-fleet (for which we have a complete or partial HC profile), the prognostic algorithm 1 consists of two main phases:

A). Knowledge extraction from the condition monitoring data of the product as described in section 2: past usage information is extracted by taking for each product the distribution of the sampling time and the distribution of the health condition variation in the related condition monitoring data.

B). Knowledge exploitation: prediction of the future HC profile over time and extracting a confidence interval for the test product RUL. This is done through the steps shown in Fig. 1 (left), assuming that the correlation structure between the random samples of ΔHC and Δt is captured by the copula. The product is not able to fulfil its required function for HC = 0, when it reaches its estimated End of Life (EOF). The algorithm is run for every item in the reference fleet, repeating it a significantly large number of times so that Monte Carlo method can be applied to obtain a 5% confidence interval for the RUL.

Maximum Likelihood Estimation (MLE) is utilized to set the parameters of the copula which fits best the data [11]. In the baseline scenario the independence copula is considered, and the same results are obtained as in the relevant papers [1]. Then other copulas types are used and based on the log-likelihood values, the most suitable one is selected. The final aim is to observe RUL variations obtained by taking into account the dependence between sample time and health condition variation.

The algorithm performance is estimated for a given

observed degradation level (the % of collected data for the test product with respect to its actual lifetime) and different values of α (which defines the desired level of selectivity of the KST) [1].

B. Algorithm 2

Another approach to solve the same problem is proposed in the following section. A set of CM data related to *N*=90 products is considered again. By inspecting the data, a significant pairwise correlation is found between t_i when i = 25, 50, 75 (t_i is the time needed for the product to reach a degradation level equal to i expressed in % [12]) and t_f (the EOL). For instance, the time needed to reach HC=100-25=75% appears to be significantly correlated with the time needed to reach the EOL. This is shown in the scatter plots in Fig. 2 where the correlation is calculated using Spearman's rho (a non-parametric measure of correlation between variables [13]).

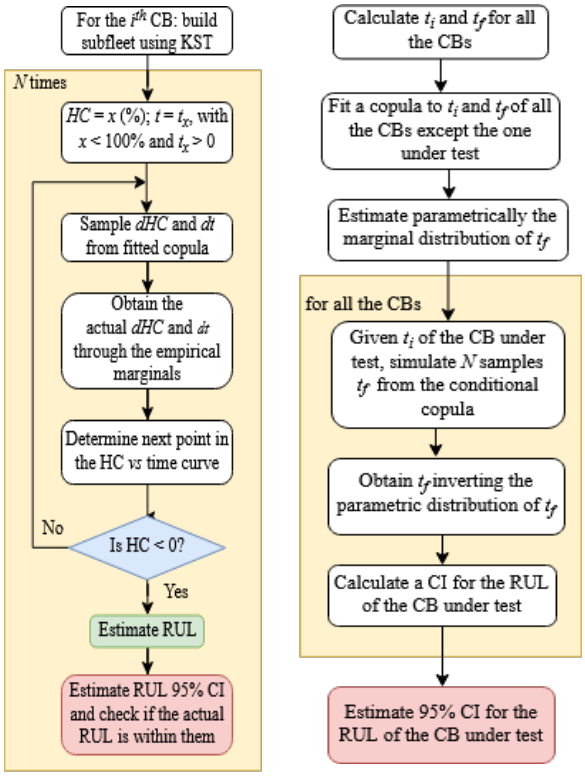

Fig. 1. Flowcharts of algorithm 1 (left) and algorithm 2 (right)

Since from Fig. 2 it becomes clear that some kind of relationship exists, the aim is to predict the RUL (defined as the difference $t_f - t_i$) of a specific CB at a given t_i, by using the information of all the other CBs in the dataset. The algorithm 2 consists of the steps shown in Fig. 1 (right) and it runs for all the *N* CBs in the dataset.

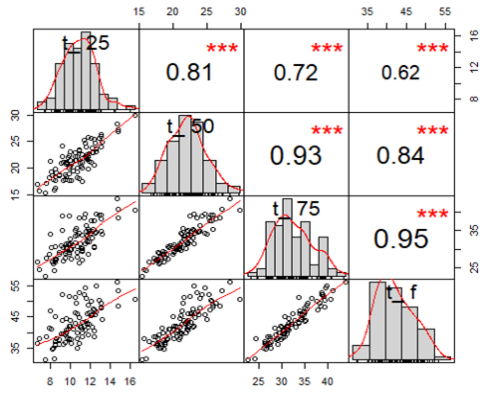

Fig. 2. Pairwise comparison between t_i and t_f

IV. RESULTS AND DISCUSSIONS

In the baseline scenario a naturally assumed correlation is discussed between the sampling time and the variation of health condition. A reasonable expectation is the longer the sampling time, the higher the variation of HC. Since this case is not taken into account in the previous similar work [1], at first it is examined through the analysis of correlation between the sampling time and HC variation from the available dataset. The data, however, indicates that there is almost no correlation pattern between these two variables both when considering a single item (Fig. 3 left), and when considering the entire fleet (Fig. 3 right): indeed the Spearman's rho for the entire fleet is ρ= - 0.025, which indicates that almost no correlation exists. However, at 1% significance level, a bivariate asymptotic independence test based on Kendall's τ (which counts the number of different pairs between two ordered sets and gives the symmetric difference distance [14]) still suggests some possible dependence. Based on the log-likelihood values, a 180° rotated Clayton [10] copula has been selected for modelling the correlation structure.

As it is shown in Fig. 4, some level of correlation and highly visible asymmetry in the upper part exist between the degradation rate and the sampling time.

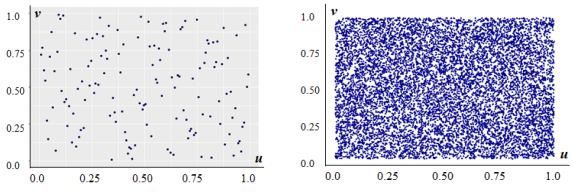

Fig. 3. Pseudo observations of sampling time (u) vs. HC variation (v) for a CB (left) and for the entire fleet (right)

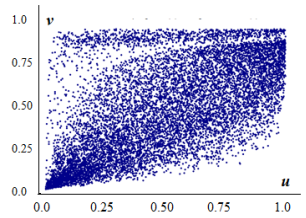

Fig. 4. Pseudo observations of sampling time (u) vs. degradation rate (v) for the fleet

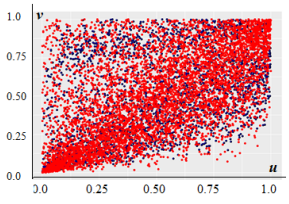

Fig. 5. Pseudo observations of sampling time (u) vs. HC variation (v) for the fleet: red-simulated observations, blue-real observations

The degradation rate and its relationship with the sampling time are considered to be inherent in the model, since they are mathematically linked through Eq. 3. However, the correlation between them is taken into account in a second implementation of algorithm 1 by sampling d (instead of ΔHC) and Δt from a fitted copula (a 180° rotated Tawn [14]): the results are shown in Fig. 5 (in red the simulated data, while in red the observed one) and the algorithm performance is estimated for different levels of α. In both the implementations of algorithm 1 very similar results have been obtained so only the results of the first one are reported in Fig. 6. As it can be seen, the alpha value does not seem to have a notable impact while a significant accuracy (>90%) is reached for observed degradation greater than 50%.

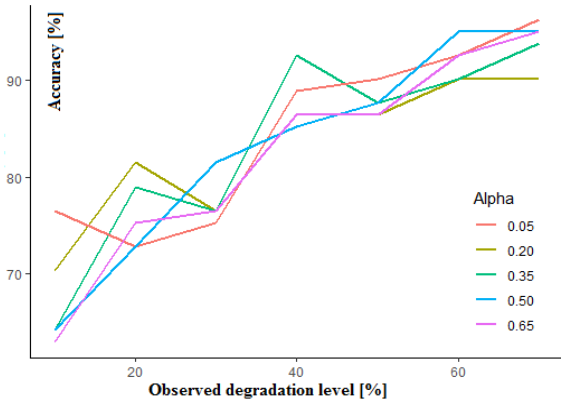

Fig. 6. Algorithm 1 performance N_f= 81

For what it concerns the second algorithm, depending on the (t_i, t_f) pair, Frank and Gaussian copulas [9] are used in the analysis. Independence test at 5% significance level is done prior to fitting each copula through MLE. As it can be seen in Fig. 7, algorithm 2 shows a very high accuracy (>90%) even at low observed degradation level (20%).

A significant advantage of algorithm 1 is the fact that it can be applied to products for which the entire HC time series is not available. On the other hand, algorithm 2 shows higher accuracy even for very low observed degradation levels but requires the complete HC time series for at least some of the products.

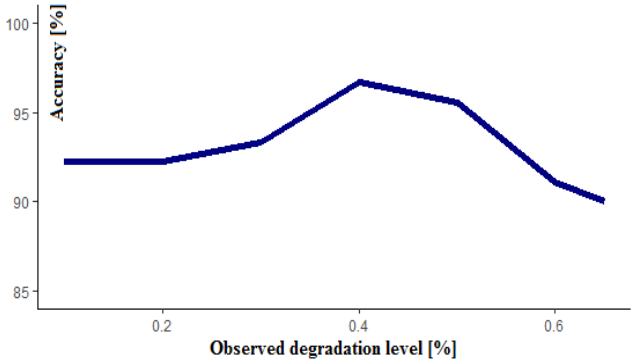

Fig. 7. Algorithm 2 performance N = 90

Another point of comparison between the two models can be the width of their confidence interval (CI). In Fig. 8 and 9 are plotted the CIs of prediction as a percentage of the true RUL value for each algorithm respectively.

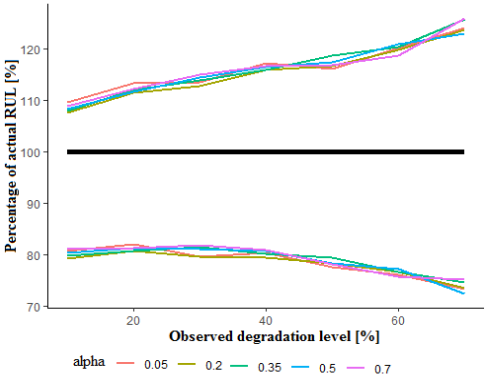

Fig. 8. Actual RUL in case of algorithm 1

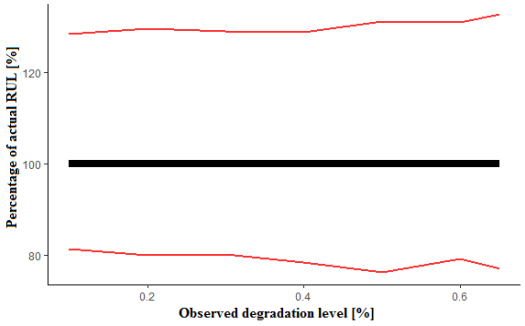

Fig. 9. Actual RUL in case of algorithm 2

The following observations can be made for the first algorithm:

- Regardless of the selectivity value α, the same trend for the confidence intervals is observed;
- As the degradation level increases, the confidence intervals (as a percentage of the actual RUL) increase as well. The overall increase is about 20% (10% on both sides of the CI). This can be explained by the fact that the values to be predicted become smaller and therefore the error weighs more.

For algorithm 2 instead, it can be noted that:

- The confidence intervals do not vary significantly with the observed degradation level;
- The uncertainty on the prediction is, however, larger if compared with the confidence intervals of algorithm 1.

As far as computational complexity is concerned, generally speaking the two algorithms do not pose any particular problems, at least when dealing with limited data such as in this case. Indeed, the most demanding operation is the fitting of the copula. However, this operation is done only once, given the information available on the CBs and then it is updated only when new information become available. To predict the CI of the RUL for a CB, algorithm 2 requires extracting N samples only, while algorithm 1 requires repeating N times the simulation of the entire life of the CB. This means that if the life of a CB is described by m samples, then algorithm 1 needs to sample $m \cdot N$ samples, i.e. it needs m times the number of samples of algorithm 2 to achieve the same purpose.

Overall, a possibility for combined usage of both algorithms seems to exist: algorithm 1 could be used on new products for which historical data is limited by exploiting the fleet and sub-fleet concepts. Algorithm 2 instead can be used as soon as some historical data is collected. Another possible use would be to combine these algorithms through ensemble learning [16] or choose which algorithm to use based on the observed degradation level.

V. CONCLUSIONS AND OUTLOOK

The novelty in the proposed models is the attempt to exploit all the information enclosed in the product HC profile by including the correlation structure in the models. Taking into account the entire HC profiles instead of RUL values only, the models obtain more detailed information, such as Probability of Failure within a predetermined time interval. Considering the appropriate dependences by using the copula approach allows for further exploiting the information contained in the data and improving the performances of the prediction algorithms, particularly evident in the second algorithm. In case of algorithm 1, another advantage is that the sub-fleet is not strictly required to include only products with a known RUL, but also products characterized by a partial HC profile knowledge. A further usage could be explored by combining both the algorithms depending on the data available.

REFERENCES

[1] **Leone, G.; Cristaldi, L.; Turrin, S**.: A data-driven prognostic approach based on statistical similarity: An application to industrial circuit breakers, *Measurement*, No. 108, 2017, pp. 163-170.

[2] **Pecht, M.G**.: *Prognostics and Health Management*, Electronics, John Wiley & Sons, 2008.

[3] **Schneider, K.; Cassady, C. R.:** Fleet Performance Under Selective Maintenance, *Reliability and Maintainability, 2004 Annual Symposium –RAMS*, 2004.

[4] **Yanagi, S.**: An Iteration Method for Reliability Evaluation of a Fleet System, *Journal of Operations Research*, Vol. 43, No. 9, 1992, pp. 885-896.

[5] **Wouters, P.A. A. F.; Schijndel V. A.; Wetzer, J. M.:** Remaining Lifetime Modeling for Power Transformers: Individual Assets and Fleets, *Electrical Insulation Magazine*, IEEE Vol: 27, pp. 45-51.

[6] **Turrin, S.; Subbiah, S.; Leone, G.; Cristaldi, L.:** An Algorithm for Data-Driven Prognostics Based on Statistical Analysis of Condition Monitoring Data on a Fleet Level, *2015 IEEE International Instrumentation and Measurement Technology Conference Proceedings*, Pisa, Italy, 2015.

[7] **Li, M.; Vitanyi, P. M. B.:** Two decades of applied Kolmogorov complexity: in memoriam Andrei Nikolaevich Kolmogorov 1903-87, *Structure in Complexity Theory Third Annual Conference, Washington*, 1988, pp. 80-101.

[8] **Turrin, S.; Subbiah, S.; Leone, G.; Cristaldi, L**.: An algorithm for data-driven prognostics based on statistical analysis of condition monitoring data on a fleet level, *International Instrumentation and Measurement Technology Conference*, 2015, pp. 629-634.

[9] **Peng. W.; Zhang X.; Huang, H-Z.**; A failure rate interaction model for two-component systems based on copula function, *Risk and Reliability*, Vol. 230(3), 2016, pp. 278-284.

[10] **Nelsen, B. R**.; *An Introduction to Copulas*, 2nd Ed., Springer, 2006.

[11] **Jia, X.; Cui, L**.: Reliability research of k-out-of-n: Supply chain system based on copula, *Communications in Statistics – Theory and Methods*, 41:21, 2012, pp. 4023-4033.

[12] **Xi, Z.; Jing, R.; Wang, P.; Hu, C.**; A copula-based sampling method for data driven prognostics, *Reliability Engineering and System Safety*, 132, 2014, pp. 72-82.

[13] **Rebekic, A.; Loncarric, Z.; Petrovic, S.; Maric, S.:** Pearson's or Sperman's correlation coefficient – which to use?, *Poljoprivreda*, 21(2), 2015, 47-54.

[14] **Herve, A**.: The Kendall rank correlation coefficient, *Encyclopedia of Measurement and Statistics*, 2007.

[15] **Eschenburg P**.; *Properties of Extreme-Value Copulas*, PhD thesis, Technical Universituy of Munchen, 2013.

[16] Zhang C., Ma Y. (Eds.), *Ensemble Machine Learning*: *Methods and Applications*. Springer, 2012.