International Journal of
Production Management
and Engineering

http://dx.doi.org/10.4995/ijpme.2013.1517
Received: 2013-05-16 Accepted: 2013-05-31

Survival function of a power transformer and a switch by means of non-parametric estimators

Aparicio Ruiz, P., Rodríguez Palero, M., & Onieva Giménez, L.

Dpto. de Organización Industrial y Gestión de Empresas II, Escuela Superior de Ingeniería
Avda. de los Descubrimientos s/n, Isla de la Cartuja, 41092 Sevilla - Spain
pabloaparicio@us.es
mariarodriguez@us.es
onieva@us.es

Abstract: The aim of this work is the estimation of the survival function of a power transformer and a switch for a medium voltage substation, which provides the empirical reliability of this transformer and its switch. The statistical analyses of recurrent events are used for this estimate. In this study, have been applied several estimators: in the presence of correlation, under a model of maximum likelihood and assuming a gamma frailty model. This work is part of a project applying various techniques of maintenance. These techniques are based on the reliability of the electrical substations belonging to future Smart Grid.

Key words: *Survival function, non-parametric estimators, transformer.*

1. Introduction

Nowadays, the power distribution companies are in a changing and fiercely competitive scenario. The traditional energy distribution will change in a short space of time towards a new paradigm. The situation which for years has been the policy framework of the electricity distributors, based on regulated monopolies typically directed by a major producer and countless captive customers. It is expected to disappear soon.

The success of the current distribution is to adapt to the new situation where there will be many producers, smaller, and also a host of customers, with much higher requirements for quality of service and performance. In this new energy framework, the traditional distribution networks start to become obsolete. The energy distribution must become more pro-active, and must develop and use new tools, new concepts and new responses, to maintain and improve service quality by developing and incorporating new technologies and techniques. One of the aspects to be included in this new situation is the new maintenance management techniques.

One of the aspects to be included in this new situation is the new maintenance management techniques. The standard UNE-EN 13306:2011 define maintenance as the combination of all technical, administrative and managerial actions during the life cycle, these are realised by means of maintenance planning, control and supervision of maintenance, improvement of methods in the organization including economics aspects. One of the keys to improving electric service is to increase the continuity of electric service, by carrying out a continuous supply in time, and even wiping out or minimizing interruptions of supply to end customers. At the same time, due to this aim, electrical substations should increase their reliability indices. Given the powerful network of generation and distribution of electric utilities networks, actions must be carried out globally increase the reliability of all components of the network.

This work is included in one of largest whose purpose is to evaluate the reliability of electrical substation medium voltage distribution through the application of RCM (Reliability Centered Maintenance). This document describes the work that has been developed for estimating the reliability function of a power transformer and its switch, considered critical element in substations via a non-parametric estimation. The goal is to plan and carry out maintenance plans to improve the reliability obtained.

2. Antecedent

As mentioned above, this work is included within a major project. The development of such work has been performed to identify the functional blocks in which a substation may be divided. Once these components are identified, an analysis of the different failure modes that can occur in each element. For the determination of elements and critical failures, has done a failure modes and effects analysis (FMEA). This methodology looks set for each equipment, the failure mode, its causes and effects. In the beginning of this paper, this study was performed once identified the critical elements of the substation. Once identified failure modes, the study was completed with a preliminary hazard analysis (PHA).

Following these studies identified the critical elements in the functioning of substations: the power transformer and switch. Just as the most common failure modes of these critical elements.

In the reliability analysis of transformers, It is interesting to know the independence or dependence of the parameters that indicate the operation of the transformer. Throughout the development of the works have been collected different control parameters of transformers under study. The parameters that were collected for the transformers are shown in Table 1.

In this table, It also indicates the abbreviations used to identify this parameters. As seen below, the analysis of the reliability of the transformers was performed for different levels of one of these parameters.

To study the dependence between variables, were monitored and were calculated the correlation matrix for two specific computers. Analyzing the values collected by two transformers, in two medium voltage outdoor electrical substations.

Table 1. Control parameters used in the power transformer.

Code	Description
MT1	Oil temperature
MT2	Air temperature
MT3	Humidity in the Oil
MT4	Presence of gases in the oil

The correlation matrix is obtained from the analysis of time series, that are obtained from the collected control signals. The data set must be uniform with respect to the time interval to which they belong. In this way, we can know the evolution of the parameters at the same operating conditions.

For the realization of the matrix of each equipment, have been analized the time series in the significant parameters that have been collected in the project's development.

In Table 2, is shown graphically by means a color code, the level of correlations between the four measure taken for the power transformer. The darker color in the cell represents the higher correlation between the measures adopted. In the same way, on the right-hand side of the Table 2 is shown in detail the interval for each color level. For example, for set variables MT2 and MT3, the magnitude of correlation index between both, is defined in the positive range between 0.80 and 0.90, where this interval is said to be right-open. That is, the limit value range on the right does not belong to it. It is noteworthy that the highest correlation occurs between the oil temperature (MT1) with other input parameters. It can be seen that there is a high degree of correlation, positive in this case, of all actions taken and the indicated (MT1).

All this suggests a high degree of relationship between them. It can also be noted that the lower correlation occurs for the air temperature (MT2) relative to the humidity (MT3) and gases in the oil in the power transformer (MT4). This indicates that an increase or decrease in this parameter has an effect of the same sign, in case of positive correlation, but to a lesser extent than produced by a proportional effect.

For the election of parameters used in the reliability analysis was used the parameter greater correlation with others, because it is understood that this will be

Table 2. Correlation Matrix.

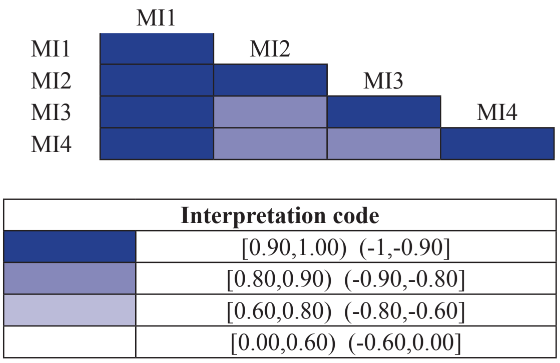
	MT1	MT2	MT3	MT4
MT1	■			
MT2	■	■		
MT3	■	▓	■	
MT4	■	░	▓	■

Interpretation code	
■	[0.90,1.00) (-1,-0.90]
▓	[0.80,0.90) (-0.90,-0.80]
░	[0.60,0.80) (-0.80,-0.60]
	[0.00,0.60) (-0.60,0.00]

Table 3. The control parameters used in the switch.

Code	Description
MI1	Humidity in SF6
MI2	Number of operations in servomotor
MI3	Wear poles
MI4	Number of operations

Table 4. Correlation matrix in the parameters list of switch.



	MI1	MI2	MI3	MI4
MI1	■			
MI2	■	■		
MI3	■	▩	■	
MI4	■	▩	▩	■

Interpretation code	
■	[0.90,1.00) (-1,-0.90]
▩	[0.80,0.90) (-0.90,-0.80]
▨	[0.60,0.80) (-0.80,-0.60]
□	[0.00,0.60) (-0.60,0.00]

the one that best characterizes the state of the power transformer and therefore the most significant to estimate the reliability thereof.

In the case of the switch, the parameters considered are shown in Table 3.

Analogously, was studied the correlation matrix of the parameters used in the switch case, the results obtaining are shown in Table 4. In this are shown the correlations for measures relating to the switch. The interpretation of the color code is similar to the previous. The first column, corresponding to the first measure, presents the highest correlation with other measured parameters. As can be seen, the measure is identified by the code MI1, reflected in the table with the darker color, is humidity in SF6. This suggests, that the measure is the more effect it has on the other, so choose it is convenient to characterize the state of the switch.

Described below the methodology used to esti-mate the reliability function of the power transformer in response to most frequently identified failure as the failure mode 1 and mode 2 failures and the reliability function of the switch, in this case for a single failure mode, identified as mode 3. The aim of this comparison is to identify the faults that have more influence on the values of system reliability. The models obtained can also be used to describe the occurrence hazard of an event of interest.

3. Methodology

The reliability analysis of a system includes the concepts, tools and techniques, which study the time of occurrence until an event occurs. These studies are currently widely applied in scientific research in various fields such as health (González and Peña, 2004) or the financial markets (Fuentelsalz, 2004).

In the case at issue in this work, assuming as initial time instant immediately following the first recorded failure, consider n types of modes of transformer failure, which are considered independent of each other, is called by the variable T_{ij} randomly while operating the transformer to the jth failure for the ith failure mode. This random variable is a function of probability of failure is unknown, which is defined according to equation (1).

$$F(t) = P(T_{ij} \leq t) \tag{1}$$

Another way of expressing the same distribution is through the reliability function, R(t). Which determines the cumulative probability that the event under study occurs after time t. This dis-tribution and the relationship with the above equations are defined by (2) and (3).

$$R(T) = P(T_{ij} \geq t) \tag{2}$$

$$R(t) = 1 - F(t) \tag{3}$$

The failure probability function, F(t), can be expressed also by their failure density function f(t) as shown in (4)

$$F(t) = \int_0^t f(u)\,du \tag{4}$$

This density function can also relate to the previous equations according to the expression (5), (Fuentelsaz *et al.*, 2004).

$$f(t) = \lim_{\Delta t \to 0} + \left(\frac{P(t \leq T_{ij} < t + \Delta t)}{\Delta t} \right) = \frac{dF(t)}{dt} = \frac{-dS(t)}{dt} \tag{5}$$

Another feature of interest is the cumulative hazard function, H(t), given in definition of the above function as expressed in equations (6) and (7), (Fuentelsaz *et al.*, 2004).

$$H(t) = \int_0^t h(u)\,du \tag{6}$$

where h(t) is the failure rate function,

$$r(t) = \lim_{\Delta t \to 0} + \left(\frac{P(t \leq T_{ij} < t + \Delta t)}{\Delta t} \right) = \frac{f(t)}{s(t)} \quad (7)$$

In cases in which can be seen in more than one occasion an occurrence of interest, in each individual or system; specific techniques to estimate for recurring events should be used. Techniques for reliability analysis differ in the event that the behavior of the variables studied, regarding covariates, follow a known probability distribu-tion or not. In the case where the variable is not following according to a known probability distribution, is used non-parametric estimates. For this reason, the estimation of the survival function through traditional methods, such as the product limit estimator of Kaplan-Meier (1958) or other more modern, ceases to be useful because it operates with unique events, besides of independence in the occurrence of events.

In the present paper, we perform a non-parametric estimation of the time between the occurrence of faults for a power transformer, assuming that there are recurring events and interdependence between them. The variable of interest is the random variable of time between two instants when the power transformer chang-es its state from available to unavailable associ-ated with each failure mode. It is intends to conduct an estimations of transformer reliability function.

This problem has been approached by Wang and Chang (1999) and Peña *et al.* (2001), which present different estimators for the reliability function for recurring events.

To determine the reliability of the power trans-former and the switch have been used estimators previously discussed, which are described in more detail in the following sections.

3.1. WC estimator

Wang and Chang (1999) propose an estimator for the case of recurrent events in the presence of correlation, for the particular case that the observations are independent and identically distributed. The authors present a model of fragility where the estimator for the reliability function is given by the expression (8),

$$\hat{R}(t) = \prod_{Tj \in T, Tj \leq t} \left(1 - \frac{d*(Tj)}{R*(Tj)} \right) \quad (8)$$

where d*(t) is the sum of the proportions of the devices in which the inter-occurrence times are equal to t, when there is at least one fault, and R*(t) represents the average of devices that are in risk at instant t.

3.2. PSH estimators

In Peña *et al.* (2001) are presented two possible estimators. The first is an estimator non-parametric maximum likelihood for a model with variables independent and identically dis-tributed, called IIDPLE (Independent and Iden-tically Product Limit Estimator). This estimator is an extension of the product limit estimator for recurring events through counters processes, and is expressed through two functions doubly indexed to time scales: calendar time, s, and inter-occurrence time, t. IIDPLE expression of the estimator is presented in (9),Infrastructure

$$\hat{R}(t) = \prod_{w \leq t} \left(1 - \frac{N(s, \Delta w)}{Y(s, w)} \right) \quad (9)$$

where N(s,t) represents the observed events number that occur in the time interval [0, s], where inter-ocurrence times do not exceed the time unit, t, and Y(s,t) represents the number of observed events in the calendar time [0, s], where interocurrencia times are at least, t, time units.

The second estimator proposed by the same authors is used to determine the distribution of the times of occurrence when times are correlated according to a gamma frailty model with shape and scale parameters equal to α and unknown. The proposed estimator is presented in expression (10),

$$\hat{S}(s,t) = \left[\frac{\hat{\alpha}}{\hat{\alpha} + \hat{h}_0(s,t)} \right]^{\hat{\alpha}} \quad (10)$$

Where $\hat{\alpha}$ and $\hat{h}_0(s,t)$ are estimators of the scale parameter, α, and the marginal accumulative risk function of h_0(t), respectively.

4. Reliability assessment

To perform the desired estimates, it is necessary to have a database that is sufficiently extensive and debugged, which affect a greater accuracy in the results.

To estimate the reliability of the power transformer, were analyzed the faults in one of the substation transformers involved in the project, during the development of this activity.

Table 5. The statistical descriptions for the failure mode 1.

Fault 1. type	Frequency	%	Median	Interquartile	Skewness	Kurtosis
LEVEL I	0.724138	72.41%	6.00	11.00	1.711362	2.016715
LEVEL II	0.206897	20.69%	53.50	70.00	0.000000	-0.867089
LEVEL III	0.068966	6.90%	182.50	207.75	NC	NC

For each faults have been analyzed times of calendar and intercurrence between successive failure modes, as well as those times between failures categorized with the same failure mode.

For the power transformer we have been analyzed the two more frecuently failure modes. These modes are failure mode 1 and failure mode 2. For each failure occurred in the transformer under study, we also analyzed a control parameter. The parameter collects oil temperature at the last moment just before the occurrence of the fault. For this parameter have been defined three levels of temperature level I, II and III, of low to high temperature. Level I represents the range closest to the right operational level, but this is higher than the suitable operating range.

The tables shown (Table 5 and 6) the characteristic descriptive statistics of the observations collected for the two failure modes analyzed in the transformer corresponding temperature units (c.t.u.).

As shown in Table 5, over 70% of the observations collected for the failure mode 1 correspond to the first temperature level. For these observations, 50% of the data presented a value below than 6 c.t.u. and 75% lower than 11 c.t.u., as indicated by the corresponding values of the median and interquartile columns of the first row.

The following values in the row make reference to the skewness and kurtosis of the distributions. According to the data, the distribution is skewed to the right and is more pointed than a normal distribution given the value corresponding skewness and kurtosis. Regarding the temperature level II, the population of data that is available represent 20% of the population, being 53.5 c.t.u. the value which leaves 50% of the population to the left. Furthermore, 75% of the population, takes values below 70 c.t.u., in this case, the distribution is symmetrical. Finally, the observations for the third temperature range are only 6% of the population. For these observations, the 50% of the observations have measured values below 182.50 c.t.u. and 75% below 207.75 c.t.u.

According to observations made to the failure mode 2, 47% correspond to the first level of temperature, as is shown in the frequency columns. In this case, the median of the population corresponds to a level of 12.5 c.t.u. and third quartile to 19.5 u.t.c. A similar analysis can be done to the population values corresponding to the temerature levels II and III. As shown, the distribution in the three cases are asymmetric to the right and are much more pointed than a normal distribution.

In the case of the switch, a similar analysis was performed. Were collected failure modes, and were analyzed the most common failure mode, which was identified as failure mode 3. In this case the monitored parameter which has been used is the humidity level in SF6. To characterize this data was divided into three humidity levels.

Regarding the recorded values are considered correct from the operational level to the most extreme value, which was divided into three intervals. By analogy with the above treatment, is called level I, II and III in an increasing order.

In Table 7 are shown the descriptive statistics of the population of data collected for the switch. In this case, over 60% of the observations correspond to the first level of humidity. While the rest of the population was evenly divided between levels II and III. As in previous cases, once the data have been sorted, It was observed that the distribution proved to be skewed to the right and slightly more pointed than normal.

In all cases, we examined whether the available data were censored or not. There was a censored failure

Table 6. The statistical descriptions for the failure mode 2.

Fault 2. type	Frequency	%	Median	Interquartile	Skewness	Kurtosis
LEVEL I	0.476190	47.62%	12.50	19.25	2.961455	9.059048
LEVEL II	0.285714	28.57%	23.00	40.25	2.138618	4.658857
LEVEL III	0.238095	23.81%	119.00	191.00	2.099625	4.477970

Table 7. The statistical descriptions for the failure mode 3.

Fault 3. type	Frequency	%	Median	Interquartile	Skewness	Kurtosis
LEVEL I	0.636364	63.64%	6.00	39.00	1.572202	1.367747
LEVEL II	0.181818	18.18%	46.00	67.50	NC	NC
LEVEL III	0.181818	18.18%	202.00	213.50	NC	NC

rate of 12 percent compared to the total available data.

In addition, has been tested the goodness of fit of the failure times, which allowed to test the suitability of non-parametric techniques, where the null hypothesis of the law that follows the random variable, T_j, is known. At least three distributions were used: Normal, Exponential and Weibull.

In reference to the failure occurrence times, in the failure mode 1, 29 observations were analyzed, over a range of values between 0 and 233 u.t. (units of time). From these observations, 3 were censored from the right, not existing in any case, left censorship.

Adjustments were made in two distributions, Exponential and Normal. The Weibull distribution can not be used, because exists one observation with 0 value.

After the adjustments, the Kolmogorov-Smirnov test was performed. Table 8 are shown P-values obtained. P-values less than 0.05 would suggest that the analyzed data are not from the selected distributions with 95% confidence. As can be seen, these values are in the range that allows us to reject the hypothesis for adjusting the selected distributions.

In reference to the failure occurrence times, in the failure mode 2, 21 observations were analyzed, over a range of values between 1 and 686 u.t. (units of time). From these observations, 3 were censored from the right, not existing in any case, left censorship

Adjustments were made in to three distributions, Exponential, Normal and Weibull. As seen in Table 8, these values are in the range that allows reject the adjustment hypotheses for the exponential distribution. Normal and Weibull distributions have a p-value greater than the rejection value. However, both distributions are so different from the point of view of inference. For this reason, one can conclude that the results do not present the quality needed to say that the general population decreases according to these distributions.

Table 8. P-values obtained using the Kolmogorov-Smirnov test.

Distribution	Exponential	Normal	Weibull
Failure mode I	0,0072243	0,0263075	not applicable
Failure mode II	0,0128391	0,0709869	0,890452
Failure mode III	0,1532750	0,4530200	0,865070

For observations in the switch, identified as failure mode 3. 11 observations were analyzed, of which three were censored. Adjustment was made to the same distributions as in the previous case and similarly was performed Kolmogorov-Smirnov test. As can be seen, the p-values are in the range that does not allow to reject the null hypothesis. We conclude that the observations can be set at three selected distributions. This is because the number of elements of the population can not discern the correct distribution for adjustment, can adjust the three distributions shown large differences between them. In this case, it is equally correct to use non-parametric methods which do not require the choice of the distribution followed by variable analyzed.

Therefore, the results obtained indicate the adequacy of the proposed techniques for both equipment analyzed.

After completing the estimates, the empirical reliability of the transformer has been obtained. The three estimators have been applied for two most common failure modes occurred at different levels of the control parameter used; the temperature of transformer oil (failure mode 1 and 2). The results are presented through graphs for the two types of failure and the three temperature levels defined, see Figures 1 to 6.

First, the results obtained are presented to the power transformer, with reference to the most common failure mode, called failure mode 1, as the first two levels of temperature (Figure 1 and 2). For the most extreme level of temperature, level III, 2 observations were analyzed, and one of these was censored. Therefore, the graphs obtained correspond to a single point, as shown in Figure 3. In Figures 3, 4, 5 and 6 are shown the results for the failure mode 2, in the three temperature levels. Also shown are the results obtained by temperature level for both failure modes together, see Figures 7, 8 and 9.

Figure 10 shows the result obtained for the failure mode analyzed, in the case of the switch. Called as

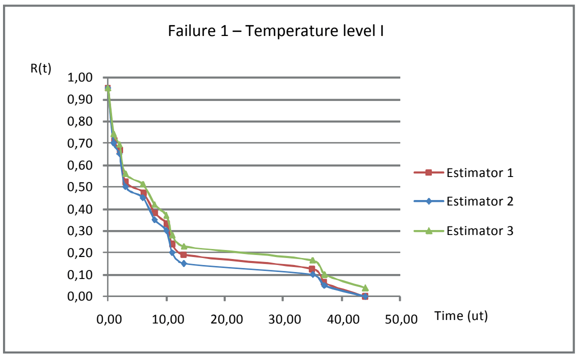


Figure 1. Empirical reliability of transformer failure mode 1. Temperature level I.

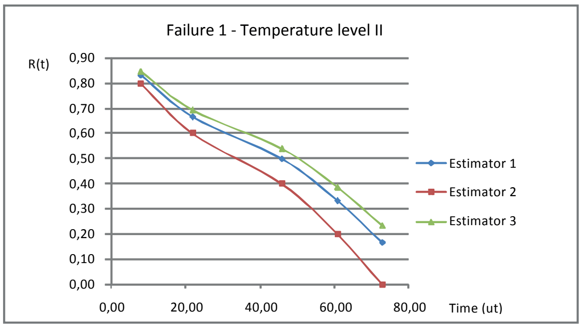


Figure 2. Empirical reliability of transformer failure mode 1. Temperature level II.

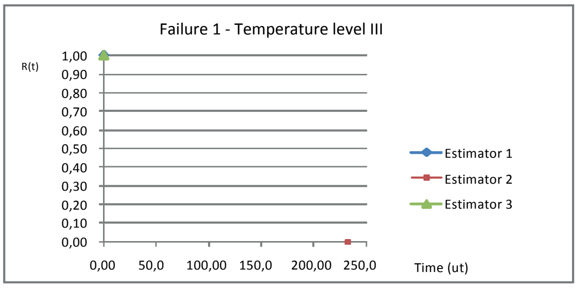


Figure 3. Empirical reliability of transformer failure mode 1. Temperature level III.

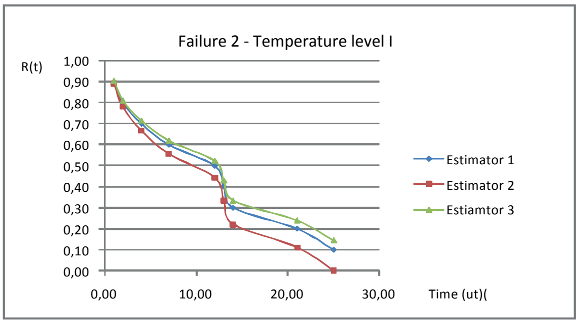


Figure 4. Empirical reliability of transformer failure mode 2. Temperature level I.

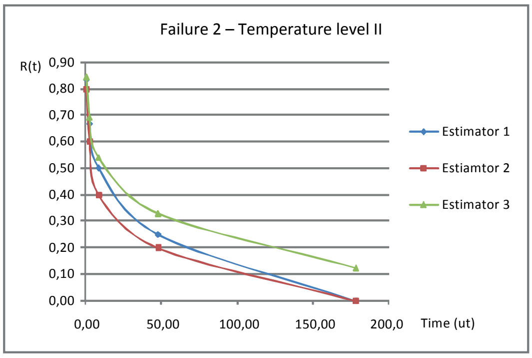


Figure 5. Empirical reliability of transformer failure mode 2. Temperature level II

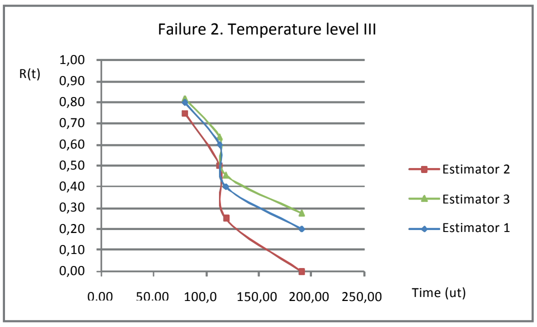


Figure 6. Empirical reliability of transformer failure mode 2. Temperature level III.

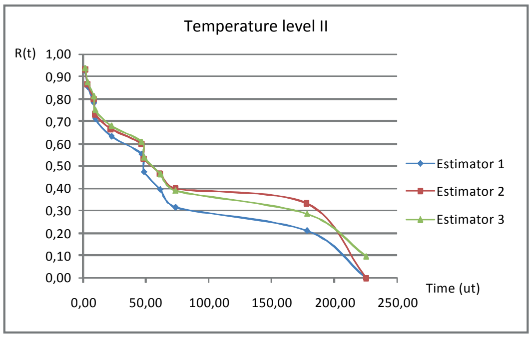


Figure 7. Empirical reliability of transformer. Temperature level I.

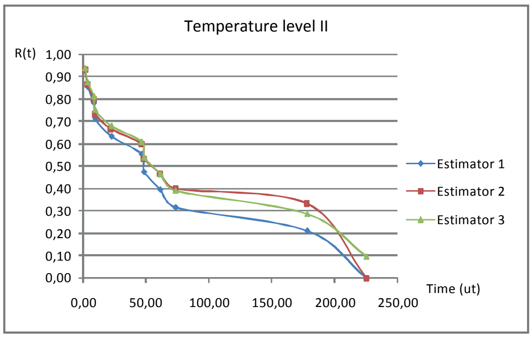


Figure 8. Empirical reliability of transformer. Temperature level II.

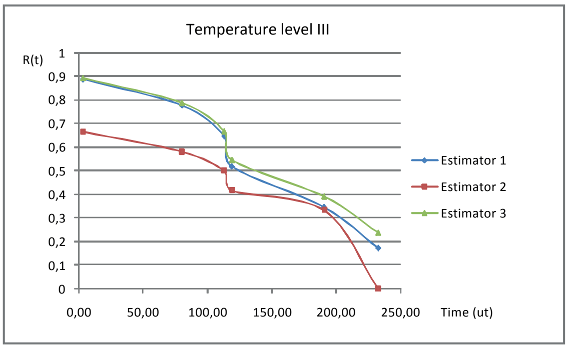


Figure 9. Empirical reliability of transformer. Temperature level III.

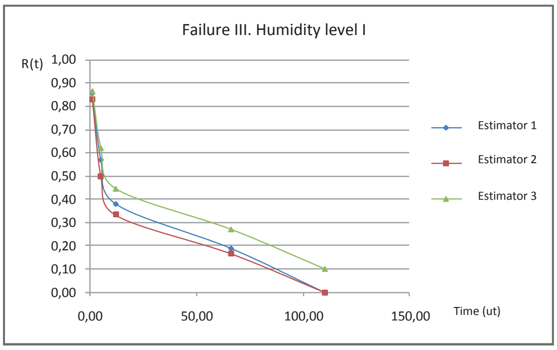


Figure 10. Empirical reliability of switch. Humidity level I.

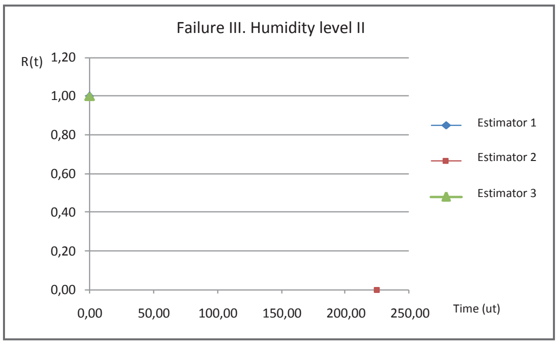


Figure 11. Empirical reliability of switch. Humidity level II.

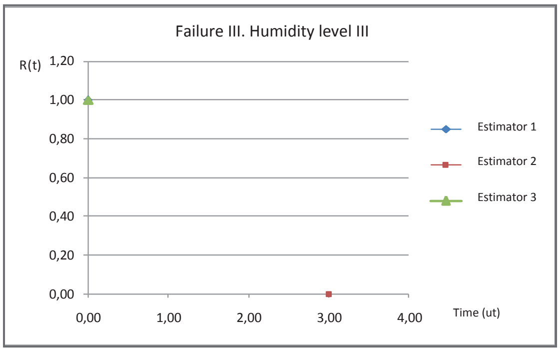


Figure 12. Empirical reliability of switch. Humidity level III.

failure mode 3, for humidity level I. For humidity levels II and III, the number of events uncensored, is unique in both levels, therefore, the graphs correspond to a single point in space to each estimator (Figures 11 and 12).

5. Conclusions

The results obtained for the transformer, under the same duration conditions, have proved that a lower temperature levels, result in a decrease in the reliability. Although a priori can be contradictory, a brief analysis explains this result. When the temperature variable is at level I, above the level of operational performance, it is usual that the failure occurs. It is unusual for the temperature to reach achieve the following levels of temperature, levels II and III, without causing the fault, because in that case the equipment would be operating outside the correct operating range for this parameter. In this way, the time between events occurring at higher temperatures are more distant from each other. This explanation also seems to be in line with the fact that are insignificant the failures occurred at the extreme temperature level, level III. In general, the fault will occur before reaching the temperature extreme level, which would require that the equipment had been operating at a level far from the recommended range. This could also be due to a very sharp increase in temperature occurred for very serious faults, which of course, have a very low frequency of repetition.

Moreover, considering the results obtained with reference to the three estimators used, can be observed other interesting conclusions. it has been observed that at lower temperature levels, the three estimators used, have presented similar results. This coincidence suggests that events do not depend each other, because the estimator can collect dependence of the events, produces the same result that the two estimators with interdependence. In contrast to the higher temperature levels, where each estimator results are quite different from each other, this is shown clearly in the higher temperature level. This suggests that the events at higher temperature level are interdependent. Finally, these results can be concluded that the failures occurred in the equipments are recurring events but interdependent, except when major failures in which there is a relationship of cause and effect, where the concatenated events have eliminated the independence.

The results for the switch has been presented for the first level of humidity, level I. The results obtained for the other two levels, as indicated previously, because it only has been a non-censored observation of each, correspond to points in space. The results obtained at the level I, where the reliability obtained with the third empirical estimator differs substantially from those obtained with the other two estimators. This difference may be caused by the dependence of time in any of the covariates was not considered in the study. Examples of possible covariates would be the outside temperature at the time of the failure, the percentage of humidity or some particular feature on the switch type. It would be interesting to do a new analysis, including or discarding some of the control signals. Regarding the results for levels II and III, note that the expected time to failure with level II is higher than the level I. This can be explained due to faults with this level of humidity are detected and the equipment is subjected to preventive maintenance. This operation aims to recover the operating condition. Consequently, the time between failures increases after this maintenance. For level III shows a sharp decline in survival time. If the humidity levels are reached, could be occurring severe failures related to each other, occurring in an uncontrolled manner in a small interval of time, without the reaction capacity necessary to perform the maintenance operation for the equipment.

It is noteworthy that the results are specific to the equipment studied, because they are directly related to their own operating conditions and environment. This allows improve the design and adaptation of maintenance policies to the operational status of each equipment. If properly combined with policies of condition based maintenance, the reliability expected of assets will increase and decrease the repair costs, replacement or unavailability of equipment.

References

Böhmer, P. E. (1912). *Theorie de unabhängigen Wahrscheinlichkeiten*, Rapports, Mémoires et Procès Verbaux du Septième Congrès International d'Actuaires.

Fuentelsaz, L., Gómez, J., & Polo, Y. (2004). Aplicaciones del análisis de supervivencia a la investigación en economía de la empresa. *Cuadernos de Economía y Dirección de la Empresa*, *19*, 081-114.

González, J. R., & Peña, E. A. (2004). Estimación no paramétrica de la función de supervivencia para datos con eventos recurrentes. *Revista Española de Salud Pública*, *78*(2), 189-199. doi:10.1590/S1135-57272004000200006

Kaplan, A., & Meier, P. (1958). Non-parametric estimation for incomplete estimations. *Journal of the American Statistical Association*, *53*(282), 457-481. doi:10.1080/01621459.1958.10501452

Peña, E., Strawderman, R., & Hollander, M. (2001). Nonparametric estimation with recurrent event data. *Journal of the American Statistical Association, 96*(456), 1299-1315. doi:10.1198/016214501753381922

Wang, M. C., & Chang, S. H. (1999.) Nonparametric estimation of a recurrent survival function. *Journal of the American Statistical Association*, *94*(445), 146-153. doi:10.1080/01621459.1999.10473831