An analysis of distribution transformer failure using the statistical package for the social sciences (SPSS) software

Análisis de fallas en transformadores de distribución utilizando el software estadístico SPSS (Statistical Package for the Social Sciences)

M. G. Mago[1], L.Valles[2], J. J.Olaya[3]

ABSTRACT

A methodology was developed for analysing faults in distribution transformers using the statistical package for social sciences (SPSS); it consisted of organising and creating of database regarding failed equipment, incorporating such data into the processing programme and converting all the information into numerical variables to be processed, thereby obtaining descriptive statistics and enabling factor and discriminant analysis. The research was based on information provided by companies in areas served by Corpoelec (Valencia, Venezuela) and Codensa (Bogotá, Colombia).

Keywords: single-phase transformer, descriptive statistics, factor analysis, discriminant analysis.

RESUMEN

Se ha desarrollado una metodología que comprende análisis de fallas en los transformadores de distribución, utilizando el programa estadístico para ciencias sociales (SPSS), para lo cual debe llevarse a cabo lo siguiente: organización y conformación de base de datos de equipos fallados, incorporación de estos datos al programa para su procesamiento y conversión de toda la información adquirida en variables numéricas para que sean procesadas y se obtengan así los resultados de la estadística descriptiva, del análisis factorial y del análisis discriminante. El trabajo se apoyó en información suministrada por empresas cuyas áreas servidas pertenecen a Corpoelec (Valencia-Venezuela) y Codensa (Bogotá).

Palabras clave: transformador monofásico, estadística descriptiva, análisis factorial, análisis discriminante.

Received: June 30th 2011
Accepted: May 11th 2012

Introduction

This paper provides an objective analysis of single-phase transformer (single-phase oil-immersed distribution transformers) failure using statistical package for social sciences (SPSS) statistical software to sample failed equipment (Tapia 2003) and assess the causes of events originating in industrial sector operations and maintenance areas. Diagnostic tools were used for matching experience and numerical calculations (simplifying complexity to simplicity), representing continuous improvement thereby enabling progress in other areas to develop technology delaying probable failure. This research contributes to estimating failed equipment in Venezuela and Colombia, using data-collection instruments used in electricity-gernerating companies to validate this study (Hernández Sampieri *et al*., 1999).

These types of fault suffered by these transformers can be classified as shown below (Normas para Transformadores de Distribuciòn, CADAFE, 1979):

Thermal occuring when the operating temperature exceeds that set by the manufacturer, causing gradual dielectric oil degradation resulting in medium-term damage to equipment due to an overload;

Arcing occurs when transformer protection fails, resulting in an external short-circuit internally damaging equipment, thereby leaving part of an electric grid unexpectedly out of service; and

Partial discharge involving small electric shocks occuring within gas cavities in a solid or liquid insulating medium. In distribution transformers they are associated with damaging surges in equipment isolation.

Statistical concepts

Statistics are not an end but a tool for analysing data; the present study adopted a statistical approach involving, "frequency calculations, fashion, means, standard deviation, variance, maximum and minimum values for each variable" (Mago and Monagas 2004). Statistical analysis of data comprised factor analysis, "to construct hypothetical scenarios or dimensions," and discriminant analysis, "to investigate the differences between individual groups" (Mago and Monagas 2004).

[1] María Gabriela Mago Ramos. Affiliation: Universidad de Carabobo, Venezuela. Electrical Engineer. MSc Industrial Engineering, MSc Electrical Engineering, E-mail: sremaria@gmail.com

[2] Luis Valles Defendine. Affiliation: Universidad de Carabobo, Venezuela. Mechanical engineer. MSc Mechanical Engineering, PhD Engineering, E-mail: le.vallesd@gmail.com

[3] Jhon Jairo Olaya Florez. Affiliation: Universidad Nacional de Colombia. Metallurgical Engineer. PhD Engineering, Universidad Nacional Autónoma de México. E-mail: jjolaya@unal.edu.co

How to cite: Mago M. G., Valles L., Olaya J. J. (2012). An analysis of distribution transformer failure using the statistical package for the social sciences (SPSS) software. Ingeniería e Investigación. Vol. 32, No. 2, August 2012, pp. 40-45.

These techniques led to establishing whether a distribution transformer's geographical location and the climactic conditions in a particular area served would affect probable failure. This would support developing stochastic models reliably endorsing diagnostic expertise (Hoel *et al*., 1972). The following were thus considered:

Factor analysis for identifying the lowest number of possible factors, a set of variables used to represent the relationship between intercorrelated variables (Mago 2011) as follows:

- the relationship between fault distribution and equipment capability;
- the relationship between fault distribution and year of manufacture, and
- the relationship between failures and the environment or geographical location housing a transformer.

Discriminant analysis for determining which variables best explain group membership (Mago 2011), identifying the purpose of analysis, criteria and independent variables:

- differentiating characteristics concerning types of distribution transformer failure and peak month of occurrence; and
- features differentiating types of distribution transformer failure and manufacturers.

Determining failure indices in single-phase distribution transformers

Data-collection instrument

A data-collection instrument was designed for a Venezuelan company that sought general information about failed units such as brand, capacity, voltage level, date of manufacture, reason for replacement, external appearance, description of possible causes of failures, etc. (Mendenhall and Sincich 1997). A Colombian company which provided the information used a similar instrument in its standards and quality procedures which served to collect data in this investigation. Table 1 shows the companies providing information for this study.

Table 1. Sample of failed companies and equipment

Location	Company	No. of failed pieces of equipment
Venezuela	Corpoelec	111
Colombia	FYR Ingenieros C.A	129

Organising and processing information and incorporating data into the statistical program (Venezuela)

When the database of failed transformers was incorporated into SPSS (version 11.0) for processing, all the information so gathered was converted into numerical variables to be processed by the software. Table 2 gives descriptive statistics regarding failed transformer data, showing the number of items of data analysed, valid minimum, maximum, mean, standard deviation, variance, mean, median and mode range of values for each item of information. The items shown were date of failure occurrence, non-specified equipment failure, capacity, municipality and feed (Miller *et al*., 1992).

Organising and processing Colombian information and incorporating data into SPSS

As with Venezuela, once data regarding failed processors had been incorporated, descriptive statistics were used to present the results shown in Table 3 (Mage and Monagas 2004), i.e. the number of items of data analysed, valid, minimums, maximums, averages, standard deviation, variance, mean, median, and mode range of values for each item of information relating to month or date the fault occurred, failed transformer capacity, year of manufacture and type of failure.

Table 2. Descriptive statistics

	N	Range	Min	Max	Standard deviation	Varian
Month	129	4.00	1.00	5.00	1.47249	2.168
Brand	129	3.00	1.00	4.00	.97074	.942
Year	129	9.00	2000.00	2009.00	2.34436	5.496
Failure	129	9.00	2.00	11.00	2.63821	6.960

	Month	Failure	KVA	Municipality	Power supply
N	111	111	111	111	111
	0	0	0	0	0
Media	4.0360	1.0000	37.4099	2.4955	18.6216
Median	4.0000	1.0000	37.5000	1.0000	18.0000
Mode	3.00	1.00	25.00	1.00	1.00
Standard deviation	1.78340	.00000	22.76343	1.97196	12.90036
Varian	3.18051	.00000	518.17363	3.88862	166.41916
Range	6.00	.00	157.50	7.00	47.00
Min	1.00	1.00	10.00	1.00	1.00
Max	7.00	1.00	167.50	8.00	48.00

Table 3. Descriptive statistics

		Month	KVA	Failure	Year
N	Valid	129	129	129	129
	Missing	0	0	0	0
Mean		2.8140	9.2636	5.1783	2006.1395
Median		2.0000	10.0000	4.0000	2007.0000
Mode		2.00	5.00	4.00	2007.00

Results and Discussion

Venezuela

5.1.1. Table 4 shows that the peak month for failure was March (24.3%) followed by May in 2012 (23.4%).

Table 4. Transformer failure rate by month

Month	Frequency	Percentage	Valid percentage
January 1	10	9.0	9.0
February 2	13	11.7	11.7
March 3	27	24.3	24.3
April 4	10	9.0	9.0
May 5	26	23.4	23.4
June 6	14	12.6	12.6
July 7	11	9.9	9.9
Total	111	100.0	100.0

Figure 1 presents a frequency histogram showing that there was a higher failure rate in March and May.

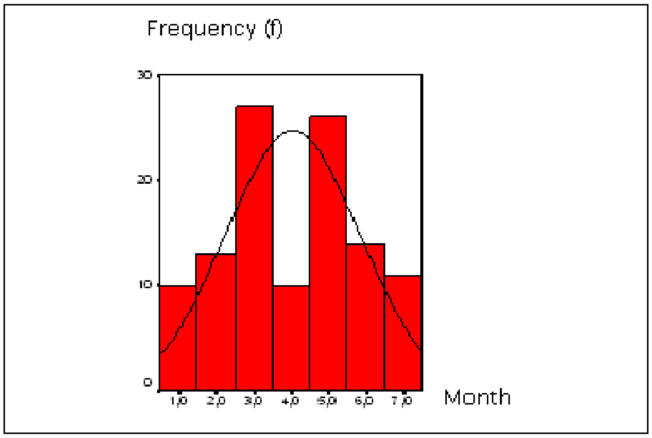

Figure 1. Frequency histogram showing transformer failure by month

The data collection instruments provided only an indication that a transformer had burned, without specifying any cause of failure.

Table 5 shows the frequencies and percentages reflecting a higher failed transformer rate regarding 25 KVA (27%) and 50 KVA transformers (26.1%).

Table 5. Failed transformer capacity

	Frequency in KVA	Percentage	Valid percentage
10,00	6	5.4	5.4
15,00	15	13.5	13.5
25.00	30	27.0	27.0
37.50	22	19.8	19.8
50.00	29	26.1	26.1
75.00	5	4.5	4.5
100.00	3	2.7	2.7
167.50	1	0.9	0.9
Total	111	100.0	100.0

Figure 2 shows checking the frequency histogram and validating the above.

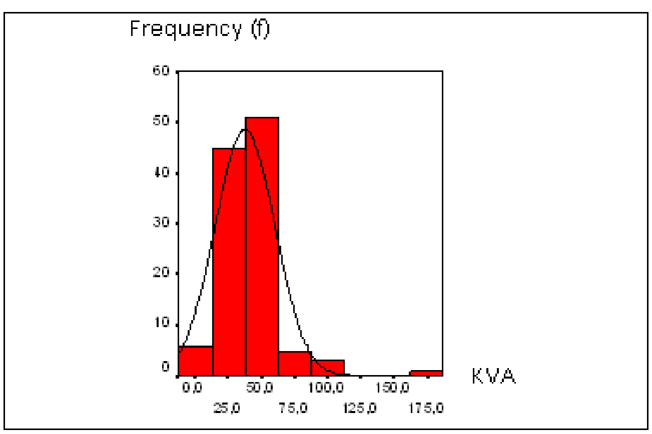

Figure 2. Frequency histogram showing failed transformer capacity

Table 6 shows that Valencia (54.1%) and San Diego (11.7%) had the highest transformer failure rates.

Table 6. Location of municipality where transformer failed

Town	Frequency	Percentage	Valid percentage
Valencia 1	60	54.1	54.1
Los Guayos 2	8	7.2	7.2
Naguanagua 3	11	9.9	9.9
San Diego 4	13	11.7	11.7
Candelaria 5	4	3.6	3.6
Miguel Peña 6	10	9.0	9.0
Guacara 7	4	3.6	3.6
No data	1	0.9	0.9
Total	111	100.0	100.0

Figure 3 gives a frequency histogram showing greater failure rate in Valencia (54.1%), followed by San Diego (11.7%).

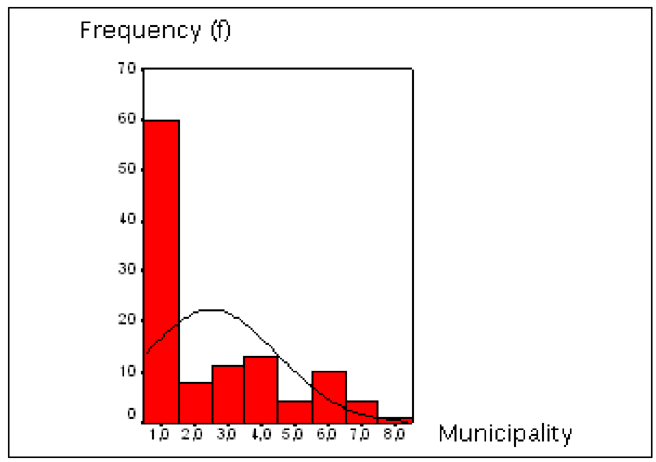

Figure 3. Frequency histogram for failed transformer location by municipality

According to the frequency histogram (Figure 4), Cardenera 1 had the highest power supply failure rate (18.1%), followed by Cardenera II (16% not 6.3%).

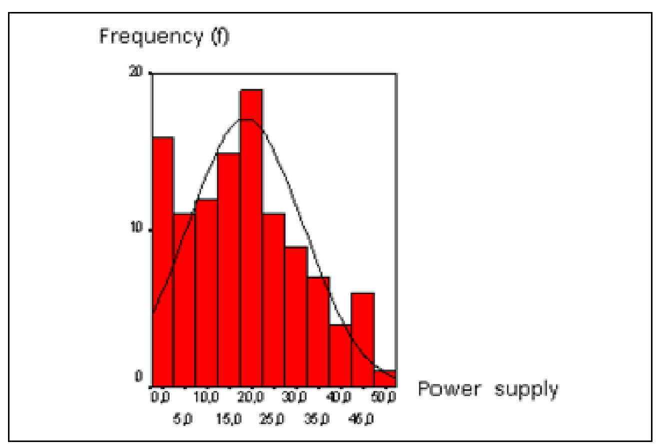

Figure 4. Frequency histogram showing faulty circuit feeder transformers

Factor and discriminant analysis

Table 7 gives a correlation matrix for all variables (Miller *et al.*, 1992). Only "municipality" had a greater than 0.70 correlation coefficient (using SPSS for validation) using the extraction method with principal components analysis (other important features could not be correlated).

Table 7. Correlation matrix for all factors

Correlation	Month	KVA	Municipality	Cause
Month	1.000	-.061	-.065	-.040
KVA	-.061	1.000	-.146	.005
Municipality	-.065	-.146	1.000	.038
Cause	-.040	.005	.038	1.000
Month		.263	.250	.338
KVA	.263		.063	.480
Municipality	.250	.063		.347
Cause	.338	.480	.347	

Only transformer failure (burning out) was considered.

Table 8. Component matrix

	1	2
Month	-9.953E-02	-.774
KVA	-.706	.411
Municipality	.777	.139
Cause	.198	.534

This obtained a function based on all information for predicting the cause of transformer failure according to geographical location. Table 9 shows three discriminant functions identifying the failed transformers and Table 10 shows the functions by i group centroid (Montgomery and Runger 1996, Spiegel *et al.*, 2001).

Table 9. Coefficients for canonical discriminant functions

	1	2	3
KVA	.508	.844	.207
Power supply	-.766	.186	.769
Month	.919	-.492	.371

Table 10. Functions regarding group centroids

Municipality	1	2	3
Valencia	6.428E-02	.107	-3.230E-02
Los Guayos	.971	-.141	-.303
Naguanagua	.123	2.610E-03	.330
San Diego	-3.923E-02	-.370	.117
Candelaria	-1.180	-.456	-.239
Miguel Peña	-.840	.110	-.100
Guacara	.320	-.156	.223
No data	-.633	.847	.275

Canonical discriminant functions not typified were evaluated in group means. "Location": month*(0.919) + power supply*(-0.766) + capacity*(0.508)

Colombia

Table 11 shows that May 2010 was the peak month for failures (35.7%).

Table 11. Transformer failure rate by month

	Frequency	Percentege	Valid percentege
April 1	27	20.9	20.9
May 2	46	35.7	35.7
June 3	6	4.7	4.7
July 4	24	18.6	18.6
August 5	26	20.2	20.2
Total	129	100.0	100.0

Figure 5 gives a frequency histogram showing greater failure rate in May and quite similar rates for April and August.

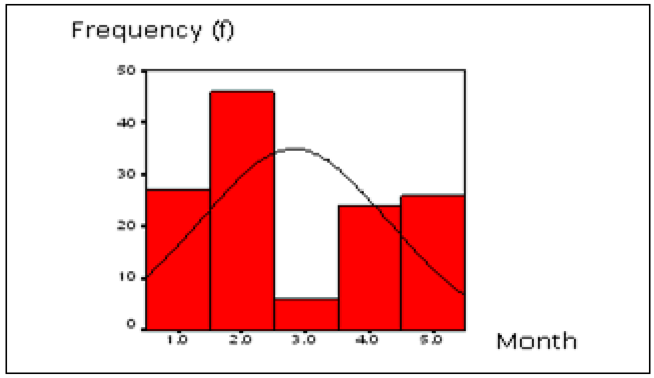

Figure 5. Failed transformer frequency histogram by month

Table 12 shows the highest fault rate occurred in 5 KVA (49.6%) and 10 KVA transformers (26.4%).

Table 12. Frequency rate regarding failed transformer capacity

KVA	Frequency	Percentage	Valid percentage
5.00	64	49.6	49.6
10.00	34	26.4	26.4
15.00	24	18.6	18.6
25.00	7	5.4	5.4
Total	129	100.0	100.0

Figure 6 gives a frequency histogram showing a high 5 kVA transformer failure rate (49.6%), followed by 10 KVA (26.4%) and 15 kVA (18.6%).

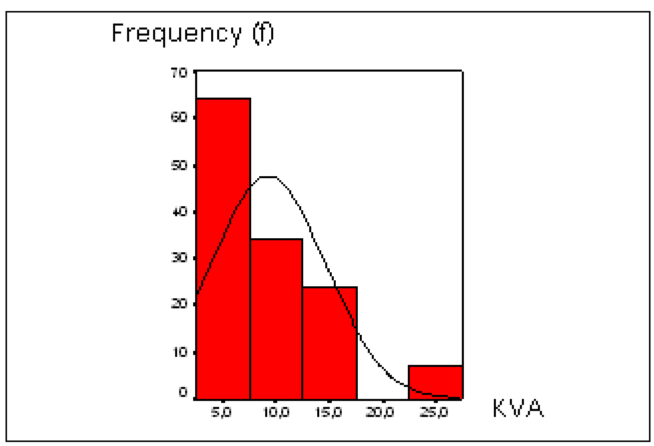

Figure 6. Frequency histogram showing failed transformer capacity

Table 13 shows that the highest failure rate concerned external circuit shorting (72.1%).

Table 13. Failure frequency rate

	Frequency	Percentage	Valid percentage
Surge	5	3.9	3.9
Overload	4	3.1	3.1
External short 5	93	72.1	72.1
Not known	11	8.5	8.5
Warranty	1	.8	.8
Maintenance	15	11.6	11.6
Total	129	100.0	100.0

Figure 7 shows a histogram describing failure frequency, external short circuits (72.1%) accounting for most failures.

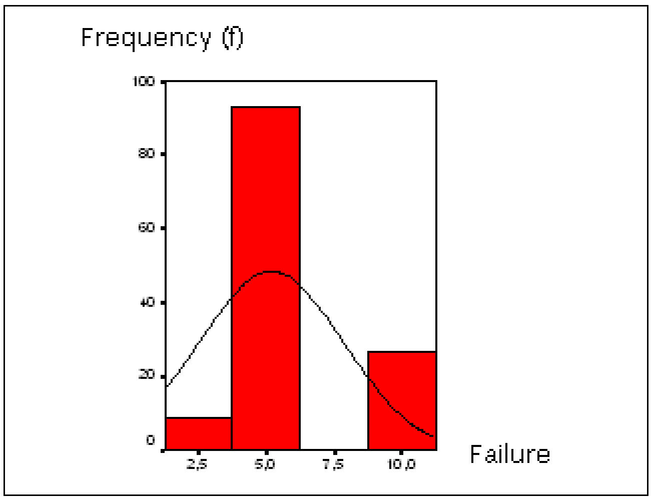

Figure 7. Histogram showing transformer failure rate

Table 14 shows that 2007 was the year of manufacture having the highest failure rate (36.4%), followed by 2008 (20.2%) and 2006 (10.9%).

Table 14. Year failed transformers were manufactured

	Frequency	Percentage	Valid percentage
2000	6	4.7	4.7
2001	1	.8	.8
2002	9	7.0	7.0
2003	5	3.9	3.9
2004	8	6.2	6.2
2005	3	2.3	2.3
2006	14	10.9	10.9
2007	47	36.4	36.4
2008	26	20.2	20.2
2009	10	7.8	7.8
Total	129	100.0	100.0

Figure 8 gives a frequency histogram showing that most failures were related to equipment made in 2007.

Factor and discriminant analysis

Table 15 shows the correlation matrix for all variables having coefficients greater than 0.70, using SPSS for validation; principal components analysis extraction method was used (i.e. for "failure"). Table 16 shows this matrix's components.

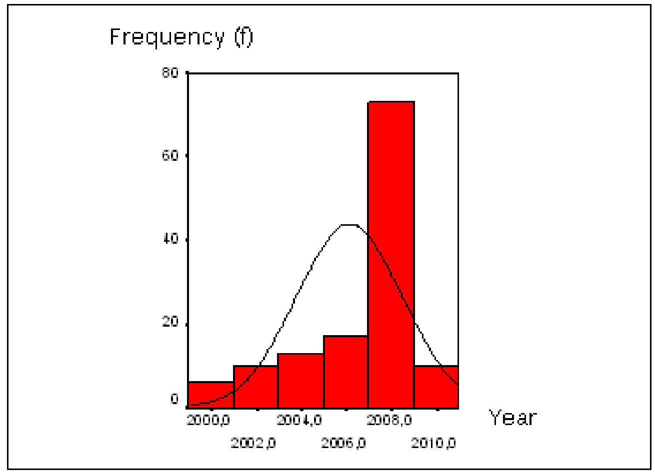

Figure 8. Frequency histogram showing failed transformer year of manufacture

Table 15. Correlation matrix

Correlations	Month	KVA	Year	Failure
Month	1.000	-.017	-.445	-.277
KVA	-.017	1.000	-.224	.271
Year	-.445	-.224	1.000	.148
Failure	-.277	.271	.148	1.000
Brand	.545	.162	-.587	-.101

Table 16. Component matrix

	Components	
	1	**2**
Month	.792	-.215
KVA	.190	.821
Year	-.816	-.175
Failure	-.325	.746
Extraction method: principal component analysis regarding 2 components extracted		

Table 17 shows that three discriminant functions identified a failed transformer.

Table 17. Extraction method: principal components analysis regarding 2 components extracted

	Function		
	1	**2**	**3**
Brand	,973	,458	,117
Year	-,191	,566	,916
Month	-,719	,834	-,240
"Failure": month(-0.719) + year*(-0.191) + brand*(0.973)*			

Conclusions

Venezuela

a) Regarding factor analysis: no correlations were statistically significant, except that most burned-out transformers were in

Valencia (even though it was not possible to identify transformer capacity, estimated cause, brand or other conditions enabling a statistical evaluation of other elements and correlating variables).

b) Regarding discriminant analysis: three discriminant functions identified failed transformers; however, function 1 provided better variance and correlation function (see Table 9), also predicting fault type according to the equipment's geographical location.

Colombia

a) Regarding factor analysis:

Peak month for occurrence and year of manufacture for failed transformer correlated

Transformer capacity and type of failure correlated

b) Regarding discriminant analysis: Three discriminant functions identified failed transformers. However, the feature giving the best variance and correlation function was No. 1 (see Table 17), which also led to predicting equipment failure.

If the variables had had greater discriminating power, one could have accurately determined distribution transformer geographical area and this would have enabled assessing environmental conditions affecting their operation and increasing the probability of failure.

References

Bausela, E. SPSS Un Instrumento de Análisis de Datos Cuantitativas". Universidad de León, España. Revista de Informática Educativa y Medios Audiovisuales. Vol 2 (4). 2005, pp 62-69.

Castañeda, M; Cabrera, A; Navarro, Y; De Vries, W. "Procesamiento de Datos y Análisis Estadísticos Utilizando SPSS". Pontificia Universidade Catolica do Rio Grande do Sul. Brasil. 2010, pp 81-95

Díaz, C; Batanero, C; Cobo B. "Fiabilidad y Generalizabilidad. Aplicaciones en Evaluación Educativa". Universidad de Granada, Revista del Instituto de Enseñanza Secuandaria de Huétor Vega. Granada, España, 2002, Nro. 54, pp 3-21.

Hernandez Sampieri; Fernández Baptista. "Metodología de la Investigación". 1ª edición. Editorial Mc. Graw- Hills. DF, Mexico, 1999, pp 52

Hoel, P; Port, S; Stone, C. "Introduction to Stochastic Processes". Universidad de California, Los Angeles. USA,1972, pp 152

Mago, M. "Análisis de Fallas en Transformadores de Distribución Utilizando Métodos No Convencionales". Area de Estudios de Postgrado. Facultad de Ingeniería. Universidad de Carabobo. Valencia, Venezuela, 2011, pp 126

Mago, M; MONAGAS E."Fallas Mas Comunes en Los Motores de Inducción de Empresas del Sector Industrial del Edo. Carabobo". Trabajo de ascenso. Escuela de Eléctrica. Facultad de Ingeniería. Universidad de Carabobo. Valencia, Venezuela, 2004, pp 44

Mago, M. "Casos de Aplicación de Métodos Numéricos en Conveyor de una Planta Ensambladora de Vehículos". VI Encuentro Colombia Venezuela de Estadística. Facultad Experimental de Ciencia y Tecnología. Universidad de Carabobo. Valencia, Venezuela, 2010,

Mago, M. "Aplicación del Software SPSS para una Caso de Regresión Lineal Simple y Múltiple en una Planta Ensambladora de Vehículos". VII Jornadas de Aplicaciones Matemáticas. Facultad Experimental de Ciencia y Tecnología. Universidad de Carabobo. Valencia, Venezuela, 2008,

Mendenhall W; Sincich T."Probabilidad y Estadística para Ingeniería y Ciencias". 4ta edición. Editorial Prentice Hall. México, 1997, pp 337

Millar; I- Freud; J-Johnson, R. "Probabilidad y Estadistica para Ingenieros". 4ta edición. Editorial Prentice-Hall Hispanoamericana, S.A. Naucaplan de Juarez, México, 1992, pp 87

Montgomery; D-Runger, G. "Probabilidad y Estadistica Aplicadas a la Ingenieria". 1ª edición. Editorial Mc Graw-Hill, Interamericana. Mexico, 1996, pp 16

Normas Para Transformadores de Distribuciòn. CADAFE. Caracas, Venezuela, 1979, pp 12

Spiegel, M; Schiller,J; Srinivasan, R. "Probabilidad y Estadistica". 2da edición. Editorial Mc Graw-Hill.Bogotá, Colombia, 2001, pp 45

Tapia, J. "Introducción al Analisis de datos con SPSS para Windows. Colección Docencia Universitaria. Ediciones de la Universidad Ezequiel Zamora, Barinas, Venezuela, 2003, pp 127

Tinoco, O. "Una aplicación de la prueba chi cuadrado con SPSS". Ind. data. [online]. ene./jun. 2008, vol.11, no.1 [citado 29 Mayo 2012], p.73-77. Disponible en la World Wide Web: <http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S1810-99932008000100011&lng=es&nrm=iso>.ISSN 1810-9993., 2000.