

# Modelling the Leakage Current Behaviour of Polluted Ceramic Insulators by Using Acoustic Emissions and Relative Humidity

Benjamín ÁLVAREZ-NASRALLAH, John Edwin CANDELO BECERRA, José Daniel SOTO ORTIZ,  
Alvin HENAO, Rafael CASTILLO SIERRA

**Abstract:** Understanding the behaviour of leakage current (LC) is not an easy task for engineers and new techniques are needed to identify the outage risk of some electrical circuits. Therefore, this paper presents a Generalized Linear Model (GLM) to characterise the LC variations in 230 kV ceramic insulators located in outdoor electrical substations subject to high pollution levels. The model uses data such as LC, acoustic emission (AE), and environmental variables (EVs). We found that the model represents the LC variation of electrical insulators of high voltage circuits. This model is useful for designing detection systems that represent the contamination levels and predict the behaviour of changes related to EV. With this model, we can determine the risk indicators for failure of electrical insulators in high-pollution areas.

**Keywords:** acoustic emissions; atmospheric variables; ceramic insulators; environmental analysis; generalized linear model; leakage current; pollution monitoring

## 1 INTRODUCTION

The most commonly used insulators in power transmission lines and substations are ceramic, glass and porcelain as their dielectric properties help avoid ground faults. When these materials are exposed to severe environmental conditions, such as high marine pollution, industrial pollution and desert pollution, the dielectric properties are affected [1]. Such conditions can occur in high-pollution areas created by vehicular gas emissions, mineral salts, dust, chemicals and other sources [2, 3].

When ceramic insulators are used in coastal zones, they are continually subjected to pollutants that change the insulators' dielectric properties and frequent washing is required to mitigate the risk of failure. For example, an electrical substation in New Barranquilla, located in the city of Barranquilla, Colombia, which is close to the Caribbean Sea, has a great number of insulators that are subjected to high marine pollution levels. Electrical lines and substations require frequent maintenance while the system is energized, which involves using treated water (resistivity  $>1500 \Omega \cdot \text{cm}$  or  $591 \Omega \cdot \text{inch}$ ) at pressures specified by national and international laws [4, 5]. However, one study has revealed the low effectiveness of washing insulators with treated water [6], which motivates the continuing search for better alternatives.

Washing insulators is an expensive task and the process requires optimising a number of maintenance activities performed during the year. Furthermore, variable monitoring is used to identify the dates when the insulators must be cleaned. For this purpose, some electricity companies measure the LC with a coil around a potential transformer. The LC measurement provides an indication of the pollution levels on the insulators [7]. This method has some problems, such as electromagnetic induction in the measuring instrument and humidity variation, which cause impedance changes in the insulators depending on the time of day. Furthermore, there is no clear reference for determining the contamination of the insulator and that increases the number of cleaning activities [8].

A more reliable solution is to measure the equivalent salt deposit density (ESDD) as described in [9]. This is

the most direct measurement of contaminants on the insulator surface. This method consists of removing manually the impurities from the insulator surface and to measure the conductivity of the sample. The problem with this procedure is that it requires too much time and considerable resources to evaluate continuously the insulator condition.

Another method uses thermal imaging to determine if the insulator is at an acceptable temperature. However, this has weaknesses because it depends on the amount of current that heats parts of the material or where spots can be detected using images. This method requires maintaining a correct distance and angle and specialised software for image analysis [3].

A simple method to determine the need for insulator maintenance is the perception of AE based on the experience of the network operators. Although it is not an accurate method, it is practical and a good indicator of the risk of power system outages. This is also a good indicator for defining the maintenance actions in the search of the optimum time to clean insulators [6].

The AE method is based on measuring the sound intensity emitted by a ceramic insulator chain, where partial electrical discharge is present. When a partial electrical discharge occurs, an instant energy release follows that emits an acoustic wave. Currently, this method is investigated in laboratories where pollution and other agents in the environment are simulated [9]. Measurements of partial discharges are performed based on the standard [10] and according to Annex F 'Non-electrical methods of PD detection'. Despite the monitoring standards, [10] does not specify how to measure partial discharge with acoustic reception. The standard [9] intends to conduct a simulation of pollutants by a saline solution sprayed on the insulator. All previous standards focus on short times and avoid the noise of other external factors in the environment [11-14].

In [15], a correlation study of AE and EV was conducted. This study found a high correlation between the LC and climate variables. The external noise was filtered to identify only the correct sounds of the partial discharges. A first model was obtained with historical

databases, but this model does not represent the variations of  $LC$ .

In this paper, we focus on determining a GLM [16] to obtain the characteristics of the  $LC$  in the electrical insulators of a substation by measuring the  $AE$ , the hour of the day when the measurement was made ( $DS$ ) and  $EV$ s such as temperature ( $TP$ ) and relative humidity ( $RH$ ). This model allows detecting contamination variations under different weather conditions as well as monitoring the non-intrusive  $LC$  of different electrical circuits.

The linear regression model is not appropriate due to the nature of its assumption; thus, we develop a model using GLM. For example, normality is not met by  $LC$  because this variable is only positive and a normal variable includes negative and positive values. Additionally, homoscedasticity is not met even by transforming  $LC$  (this variable is dependent in the model). With this model, we can estimate the  $LC$  considering the changes in the most influential variables to determine the behaviour through the time. The results show how the model fits the different behaviours of the measured  $LC$  over time, which is an important advance for designing different monitoring methods for electrical circuits.

The rest of the document is divided in three sections. Section 2 includes the materials and methods used in this research. Section 3 shows the results and analysis of the data. Finally, Section 4 presents the conclusions and future work.

## 2 MATERIALS AND METHODS

The procedure was conducted in the Nueva Barranquilla substation, located in the city of Barranquilla, Colombia. Fig. 1 shows some of the areas and electrical circuits of the outdoor substation. The most critical area for insulator pollution is the 220 kV circuits, where the voltage level represents a high risk for outages.



Figure 1 220 kV circuit in the Nueva Barranquilla substation. Source: Authors.

### 2.1 Data collection

#### 2.1.1 Leakage current

The data collection is based on [15], which considers measuring  $LC$  with a clamp ETCR 6300, located in a current transformer of the substation as shown in Fig. 2. This clamp has a resolution of 0.001 mA, an accuracy of  $\pm 1.5\%$  and a range of 0-60 A. The clamp ETCR 6300 has a communication port RS-232, which can connect to a

computer.  $LC$  was measured from March to April 2014 with a one-minute sampling period.

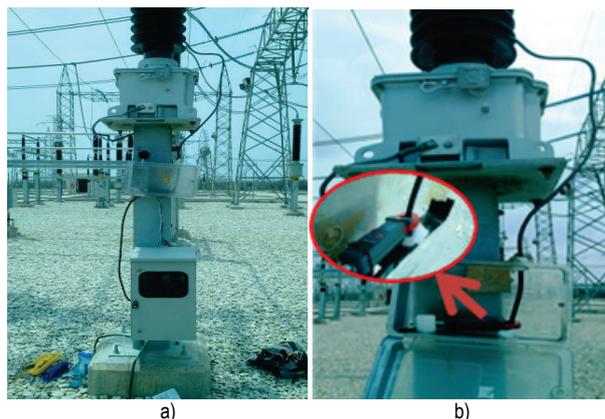


Figure 2 LC measurement (a) current transformer and (b) measuring equipment. Source: Authors.

#### 2.1.2 Acoustic Reception

Sound recording was carried out with a digital sound meter Extech model 407764 and Sound Meter software. This equipment was programmed with a laptop HP to measure the level of noise emission every minute during two washing periods.

An acoustic test was conducted to filter the external noise and maintain only the sounds required for the research. The microphone was located close to the potential transformer where the current clamp was located to capture only the sound of the discharges and avoid the sound of wind with a screen. The configuration of sound measurement is shown in Fig. 3.

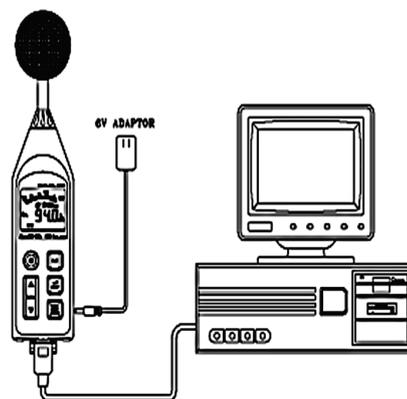


Figure 3 Measurement of sound with sound system digital Integrator<sup>1</sup>.

#### 2.1.3 Climatic Variables

Daily measurements were taken to determine the behaviour of the  $LC$  while considering the  $EV$ s. Additionally, we measured climatic variables with a Davis Vantage Pro2 Plus weather station. This equipment allowed to measure variables such as  $TP$ ,  $RH$ , solar radiation, and wind speed and direction. We used other weather stations located near the electrical substation to compare the results.

<sup>1</sup> Digital Extech sound meter manual.

## 2.2 Methodology of Data Processing

Data were collected sequentially and continuously. Observations with missing data were deleted. Problems of lack of independence were possible because of the way the data were collected. Consequently, random numbers (0, 1), obtained by a Bernoulli's probability distribution ( $p = 0.7$ ) were assigned to initial data to select those observations with 1. The purpose of this procedure was to break the problem of lack of independence of observations (7,045 observations remain after the selection). Next, each variable was analysed to identify outliers [17]; however, no outliers were found.

Observations including data for *LC*, *TP*, *RH*, *AE* and *DS* were considered in the model, where *LC* was treated as the dependent variable. *DS* was processed as a qualitative variable with four levels initially: early morning (0:00–6:00), morning (6:00–12:00), afternoon (12:00–19:00) and night (19:00–24:00). However, there was no statistical difference between the early morning and morning levels; thus, *DS* was handled as a qualitative variable with three levels: morning (0:00–12:00), afternoon (12:00–19:00) and night (19:00–24:00).

Other independent variables, such as date, wind speed and direction, atmospheric pressure, dew point, air density and solar radiation were also measured, but were not significant; thus, they were excluded by the model. However, a GLM was built with the selected observations. The type of statistical distribution for the dependent variable is a key assumption under GLM; therefore, goodness-of-fit tests are important results for this method.

In this article, a gamma distribution for *LC* was assumed because measurement of this variable produces positive values. Consequently, the independent variables were selected by using both Wald's and Pearson's Chi Square tests.

From a procedural point of view, a random group of observations (70 %) was selected to build the model and the rest (30 %) was employed to validate the forecasting capacity of the model using MAPE (mean absolute percentage error). The graph of standardised deviance residuals and predicted value of the linear predictor was also used to validate the model and the assumption of gamma distribution.

### 2.2.1 Generalised Linear Model

The structure of a multiple linear regression model is illustrated by Eq. (1):

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \varepsilon_i, \quad (1)$$

where the dependent variable  $Y_i$ ,  $i = 1, 2, \dots, n$  is modelled as a linear function of independent variables  $X_j$ ,  $j = 1, 2, \dots, k$  plus an error term. In this model,  $\varepsilon_i$  is assumed independent and identically distributed, which implies that  $E\{\varepsilon_i\} = 0$  and  $var\{\varepsilon_i\} = \sigma^2$ . It is frequently assumed that  $\varepsilon_i \sim N(0, \sigma^2)$ ; however, this assumption must be validated to have a valid model [18].

The multiple regression model is a useful framework, but there are cases where the model is not appropriate.

For example, the dependent variable is not normal because its value is restricted (e.g. positive signs) and the constancy of variance is not supported by residuals of the model.

In these cases, GLM is a valid approach to model the dependent variable in terms of independent variables, but the model has the following characteristics [16]:

First, a linear predictor  $\eta_i$  is modelled in terms of independent variables as shown in Eq. (2):

$$\eta_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}. \quad (2)$$

Second, a link function that expresses the relation between the mean of the dependent variable  $E\{Y_i\} = \mu_i$  and the linear predictor as expressed in Eq. (3):

$$\eta_i = f(\mu_i). \quad (3)$$

Third, a function of the variance in terms of the mean as presented in Eq. (4):

$$var\{Y_i\} = Kvar(\mu). \quad (4)$$

The term  $K$  is a constant that represents a dispersion parameter. It is clear now that multiple linear regression is a special case of GLM because  $\varepsilon_i \sim N(0, \sigma^2)$  and  $f(\eta_i) = \mu_i$ . Therefore,  $\mu_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}$  and  $var\{\mu\} = 1$ , so  $var\{Y_i\} = K$ .

On the other hand, in this article it is assumed that the dependent variable  $Y$  follows a gamma distribution. Consequently,  $Y \sim Gamma(\alpha, \beta)$  and  $E\{Y_i\} = \mu_i$ , where the link function is  $\eta_i = 1/\mu_i$ , the variance is  $var\{\mu\} = \mu^2$  and the linear predictor is  $\mu_i = 1/(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki})$ .

## 3 RESULTS

The goodness-of-fit test shows that a GLM assuming a gamma distribution for *LC* and using a reciprocal link function is appropriate (see Tab. 1) because the  $p$ -value does not allow rejecting the null hypothesis of the assumed distribution. The columns named 'value' and 'degree of freedom' ( $DF$ ) allow estimation of the  $p$ -value.

Table 1 Goodness-of-fit results for a Gamma model

Criteria	Reciprocal Link		
	Value	DF	p-value
Pearson's Chi-square test	170.178	4915	1
AIC	-9054.827		
BIC	-9009.318		

On the other hand, the table shows values for the Akaike Information Criterion (AIC) and Bayesian Information Criterion, which allow choosing this model because they present minimum values relative to other explored models. Therefore, according to Eqs. (2) and (3), the structure of the model is as shown in Eq. (5):

$$\mu_i = 1 / (\beta_0 + \beta_1 DS1 + \beta_2 DS2 + \beta_3 \ln AE + \beta_4 \ln RH + \beta_5 \ln AE \ln RH), \quad (5)$$

where the term  $\mu_i$  is the mean of dependent variable  $LC$ , the term  $\ln AE$  represents the logarithm of  $AE$  values and the term  $\ln RH$  refers to the logarithm of  $RH$  values.

**Table 2** Categorical variables of the model

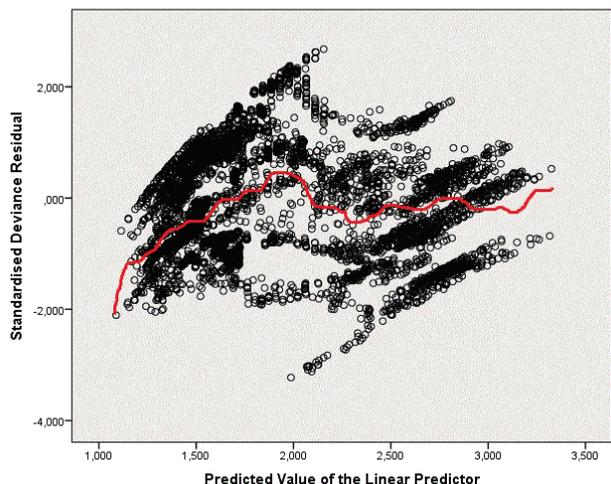
Dummy Variable	Afternoon	Night	Morning
$DS1$	1	0	0
$DS2$	0	1	0

In this model, the hour of the day (or day section,  $DS$ ) was modelled as a qualitative variable with three levels;

**Table 3** Estimation and significance of model coefficients

Parameter	$B$	Std. Error	95% Wald's Confidence Interval		Hypothesis Contrasting		
			Lower	Upper	Wald's Chi-square test	$DF$	$Sig$
Intercept	-105.30	6.00	-117.10	-93.50	307.80	1	0.000
$DS1$	0.200	0.02	0.16	0.22	158.20	1	0.000
$DS2$	-0.03	0.01	-0.05	-0.01	6.20	1	0.012
$\ln AE$	30.46	1.45	27.61	33.31	438.60	1	0.000
$\ln RH$	24.41	1.35	21.75	27.10	324.30	1	0.000
$\ln AE^*$ $\ln RH$	-6.94	0.33	-7.58	-6.29	446.20	1	0.000
Scale	0.034	0.0007	0.033	0.036			

All terms in the model are significant. To check the model, a graph of standardised deviance residual [19] vs predicted value of the linear predictor is shown in Fig. 4. A good model should not show any patron in the graphic obtained; in other words, running means must be approximately straight and flat [20]. To illustrate this, a red curve was added to the graph. The red curve links running means and is slightly flat so we can conclude the model is satisfactory. A curvilinear patron is a symptom of omission of some variables. This is not the case, thus the variables selected are appropriate to model  $LC$ .



**Figure 4** Predicted value vs standardised deviance residuals.

Tab. 4 shows the result of MAPE estimation. The estimation of the mean of  $LC$  provided by the model deviates by 15.89% compared with real values.

In summary, if the model is used to forecast  $LC$  in terms of the  $HR$ ,  $AE$  and  $DS$ , then real values of  $LC$  may be underestimated or overestimated by 16% on average.

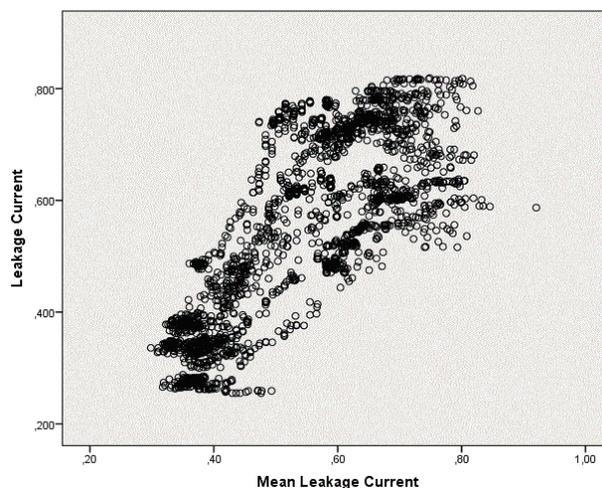
As complementary information, Fig. 5 shows a graph of real values of  $LC$  in terms of predicted mean value of  $LC$ .

as a result, only two dummy variables ( $DS1$  and  $DS2$ ) are required to represent those three levels. Tab. 2 presents the codification used in this research for variables  $DS1$  and  $DS2$ .

According to this table, if the section of the day is the afternoon, then variables  $DS1$  and  $DS2$  will take values of 1 and 0, respectively. If it is the night, then  $DS1$  and  $DS2$  will take the values 0 and 1, respectively. Finally, both variables will take a value of 0 if it is represented the morning. Estimations of coefficients are shown in Tab. 3.

**Table 4** MAPE results

Absolute Percent Error		Statistical Results
Mean		0.1589
95% confidence interval for the mean	Lower limit	0.1544
	Upper limit	0.1634
Standard deviation		0.10481



**Figure 5** Real value of  $LC$  vs predicted mean value of  $LC$  by the model.

Although the positive correlation between the two variables is evident, a larger dispersion is observed as the predicted mean of  $LC$  is greater. This is not an unexpected result because a linear regression model shows the heteroscedastic nature of the modelled data. In addition, at the bottom of Section 2.2.1 is shown how the variance of the mean is a quadratic function of the predicted mean. This behaviour is modelled by GLM.

#### 4 CONCLUSION

This paper presented an  $LC$  model obtained from  $AE$  and  $RH$  that allows identifying circuit outage risk. A GLM represented well the  $LC$  phenomenon, providing good approximation to the actual values measured in the

electrical substation during dry periods. The model allowed prediction of the *LC* to monitor their behaviour during insulator cleaning. The study was conducted for dry periods, which are the critical conditions for ceramic insulators under severe pollution, representing well the changes as repetitive cycles influenced by the atmospheric variables. We can conclude that the model provided good results. Furthermore, the model represented well the pollution behaviour during the critical period, being useful to predict the changes during the different days of the test. An increase in the sound of partial discharge represented an increase in *LC* and the higher probability of a ground fault in the power system produced by the action of pollution and humidity. This research helped test the direct relationship between the *LC* and *AE*. However, the *RH* decreased the intensity of the sound emitted by the insulators and the sound increased with increased humidity. In future, we will be working on detecting *AE* produced by *LC* in other types of insulators, electrical circuits and transmission lines to optimise the maintenance of insulators located in high-pollution environments.

### Acknowledgements

Authors thank the Power System Research Group GISEL of the Universidad del Norte, Barranquilla, Colombia, for their financial support to conduct this research. The authors also thank TRANSELCA S.A. ESP for valuable information provided for this research.

### 5 REFERENCES

- [1] Castillo Sierra, R., Oviedo-Trespalcacios, O., Canelo, J. E., & Soto, J. D. (2014). The influence of atmospheric conditions on the leakage current of ceramic insulators on the Colombian Caribbean coast. *Environmental Science and Pollution Research*, 22, 2526-2536. <https://doi.org/10.1007/s11356-014-3729-3>
- [2] Aponte, G., Castro, J. C., Sanchez, V. H., Espinoza, A., Rosales, N., & Castro, M. (2009). Contamination level evaluation on Colombian north coast. *IEEE Latin America Transactions*, 7, 190-195. <https://doi.org/10.1109/TLA.2009.5256828>
- [3] Diaz, R., Canelo, J. E., & Betancur, J. (2013). Software tool to detect hot spots on electrical insulators. *12<sup>th</sup> International Conference on Environmental and Electrical Engineering* / Wroclaw, Poland, 7-12. <https://doi.org/10.1109/EEEIC.2013.6549596>
- [4] ICONTEC. GTC 96. Guía para la limpieza de aisladores. 2004;3:25.
- [5] Departamento técnico Gamma – Aisladores Corona. Guía para la limpieza de aisladores según Norma IEEE STD 957-1995 - Segunda parte. Bol Técnico Gamma-Corona, 2007.
- [6] Martínez-Darlington, H. E. (2013). *Contaminación de Aisladores de Líneas Aéreas de Alta Tensión*. Madrid, España: Editorial Académica Española.
- [7] Tingtao, Z., Tianyu, L., Ke, C., & Xiaoguang, H. (2011). On-line monitoring system of insulator leakage current based on ARM. *6<sup>th</sup> IEEE Conference on Industrial Electronics and Applications* / Beijing, China, 365-369. <https://doi.org/10.1109/ICIEA.2011.5975610>
- [8] Farzaneh, M. & Chisholm, W. A. (2009). *Insulators for Icing and Polluted Environments*. Estados Unidos: IEEE. <https://doi.org/10.1002/9780470496251>
- [9] IEC 60815. Guide for the selection and dimensioning of high-voltage insulators for polluted conditions, 2008.
- [10] IEC 60270. High-voltage test techniques – Partial discharge measurements, 2000.
- [11] Tsuji, T., Matsumoto, S., Sakata, T., Nakayama, S., Otsubo, M; Honda, C. et al. (2005). Basic study on acoustic noise of polluted insulator and waveform analysis method. *International Symposium on Electrical Insulating Materials* / Kitakyushu, Japan, 694-697. <https://doi.org/10.1109/ISEIM.2005.193465>
- [12] Pei, C. M., Shu, N. Q., Li, L., Li, Z. P., & Peng, H. (2008). On-line monitoring of insulator contamination causing flashover based on acoustic emission. *Third International Conference Electric Utility Deregulation and Restructuring and Power Technologies* / Nanjing, China, 1667-1671. <https://doi.org/10.1109/DRPT.2008.4523673>
- [13] Li, H., Wen, X., Shu, N., & Pei, C. (2009). Application of Acoustic Emission Technology on Monitoring of Polluted Insulator Discharge. *Asia-Pacific Power Energy Engineering Conference* / Wuhan, China, 1-4. <https://doi.org/10.1109/APPEEC.2009.4918180>
- [14] Afendi, M. & Piah, M. (2011). Identification of acoustic signals of surface discharges on glass insulator under different contamination levels. *International Conference on Electrical, Control and Computer Engineering* / Kuantan, Pahang, Malaysia, 511-514. <https://doi.org/10.1109/INECCE.2011.5953936>
- [15] Álvarez-Nasrallah, B., Canelo-Becerra, J. E., Soto-Ortiz, J. D., Pombo-Acevedo, J., & Díaz-Aparicio, O. (2015). Correlation analysis between ceramic insulator pollution and acoustic emissions. *Tecno Lógicas*, 18(34), 101-108. <https://doi.org/10.22430/22565337.212>
- [16] Nelder, J. A. & Wedderburn, R. W. M. (1972). Generalized Linear Models. *Journal of the Royal Statistical Society. Series A (General)*, 13(3), 370-384. <https://doi.org/10.2307/2344614>
- [17] Hoaglin, D. C. & Iglewicz, B. (1987). Fine-Tuning Some Resistant Rules for Outlier Labeling. *Journal of the American Statistical Association*, 82(200), 1147-1149. <https://doi.org/10.2307/2289392>
- [18] Díaz-Rodríguez, M., González-Ariza A. L., Henao-Pérez, A., & Díaz-Mora, M. E. (2013). *Introducción al Análisis Estadístico Multivariado Aplicado: Experiencia y Casos en el Caribe Colombiano*. Barranquilla: Editorial Universidad del Norte.
- [19] Pierce, D. A. & Schafer, D. W. (1986). Residuals in Generalized Linear Models. *Journal of the American Statistical Association*, 81(396), 977-986. <https://doi.org/10.2307/2289071>
- [20] Lee, Y., Nelder, J., & Pawitan, Y. (2006). *Generalized Linear Models with Random Effects: Unified Analysis via H-likelihood*. Chapman and Hall/CRC. <https://doi.org/10.1201/9781420011340>

**Contact information:**

**Benjamín ÁLVAREZ-NASRALLAH**, MSc (Corresponding Author)  
Department of Electrical and Electronics Engineering,  
Universidad del Norte,  
Km 5 vía Puerto Colombia, Barranquilla, Colombia  
Email: bnasrallah@uninorte.edu.co

**John Edwin CANDELO BECERRA**, PhD  
Department of Electrical Energy and Automation,  
Universidad Nacional de Colombia,  
Cra. 80 No 65-223, Facultad de Minas, Medellín, Colombia  
Email: jecandelob@unal.edu.co

**José Daniel SOTO ORTIZ**, MSc  
Department of Electrical and Electronics Engineering,  
Universidad del Norte,  
Km 5 vía Puerto Colombia, Barranquilla, Colombia  
Email: jsoto@uninorte.edu.co

**Alvin Arturo HENAO PÉREZ**, PhD  
Department of Industrial Engineering,  
Universidad del Norte,  
Km 5 vía Puerto Colombia, Barranquilla, Colombia  
Email: henaoa@uninorte.edu.co

**Rafael CASTILLO SIERRA**, MSc  
Department of Electrical and Electronics Engineering,  
Universidad del Norte,  
Km 5 vía Puerto Colombia, Barranquilla, Colombia  
Email: rcastilloj@uninorte.edu.co