

Voltage Sag Estimation for Distribution Systems Using Linear Machine Learning Models

Daniel Galeano-Suárez, César Duarte, Jairo Blanco Solano

Department of Electrical Engineering

Universidad Industrial de Santander

Bucaramanga Santander Colombia

daniel2198611@correo.uis.edu.co, cedagua@uis.edu.co, jablanso@correo.uis.edu.co

Abstract—This paper proposes a voltage sag estimation approach based on linear machine learning models. The proposed approach estimates the voltage sag profile on distributions systems with limited monitored buses. The approach takes advantage of the priory knowledge about the linearity of the problem to train and compare three linear models: Least Squares (LS), Ridge Regression (RR), and Absolute Shrinkage and Selection Operator (LASSO). The approach is tested on the IEEE-34-bus distribution system, and the performance of models is validated through the Mean Square Error (MSE). The results show that the proposed linear machine learning models capture the internal relationships of the problem and estimate the voltage sag with high accuracy on test data.

Index Terms—Voltage sag estimation, Linear models, Machine learning, Distribution systems.

I. INTRODUCTION

Voltage sags are reductions between 0.1 pu and 0.9 pu in Root Mean Square (RMS) voltage supply with a time duration of no more than 1 minute [1]. In practice, sags are usually associated with fault currents on the systems, and its duration depends on the time of clearing the fault event [2].

Voltage sags are considered as one of the most significant components of power quality issues due to disruption of industrial processes, adverse impact on electronic equipment, and financial losses [3]. According to the statistics of the American Academy of Electrical Sciences, the economic losses of the United States due to voltage sags are more than 20 billion US dollars every year [4].

Advanced metering systems appear as a solution to assess voltage sag impact in all end-users [5]. However, the computational requirements necessary to manage big data, costs and distribution network size are reasons that make voltage sag monitoring programs still limited. Thus, a voltage sag monitoring program is designed with only a few monitored buses, and estimation methods are required to determine voltage sag at unmonitored buses.

Various approaches are used to compute voltage sag estimation. In [6], the authors uses compressive sensing theory and ℓ_1 —regularized least square formulation. In [7], presents an approach based on solving an undetermined analytical set of equations using linear programming. The same problem is solved in [8], but using the singular value decomposition

(SVD) technique and the fault position method (FPM). Since faults mainly cause sags in distribution systems, estimation through fault location is extensively explored for sag estimation [9]. In [2], the most probable fault location is estimated, and simulation can determine voltage sag profiles. However, the results can be inaccurate for the multiple fault positions because impedances value can generate similar sets of voltage sags.

All of the previously discussed sag estimation methods are based on the model. This paper presents a different approach using a bunch of learning models that are able to capture and represent complex input/output relationships. In addition, the proposed approach exploits the linearity of the problem to train models with short execution time and generalization capability.

This paper is organized as follows. Section II provides the formulation and solution methodology of the proposed problem. In Section III, the IEEE-34-bus study is presented, and the data generation and training process is described. In addition, the metrics to quantify the errors of models are presented. Finally, conclusions are drawn in section IV.

II. VOLTAGE SAG ESTIMATION USING LINEAR MACHINE LEARNING MODELS

A. Problem Formulation

The problem to be solved consists of estimating the magnitude of voltage sags at unmonitored buses. It is necessary to construct a mathematical model that associates phase voltages at the monitored buses with the sag magnitudes at unmonitored buses. This problem lies within machine learning and is classified as a supervised regression problem.

B. Methodology

In Fig. 1 a general diagram of the proposed methodology for voltage sag estimation method is illustrated. The methodology is divided into three principal parts:

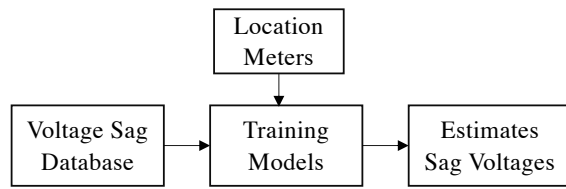


Fig. 1. Methodology for voltage sag estimation.

- Creation of voltage sag database for the distribution systems under study.
- Training linear machine learning models from sag meters and their location on the distribution system.
- Estimation of voltage sags at unmonitored buses.

1) Voltage Sag Database

The first step in the process consists of creating a database of voltage sags. For this objective, three variables are analyzed:

- Fault resistance. This is represented in discrete steps from 0 – 40 $[\Omega]$ by increments of 2 $[\Omega]$.
- Type of fault. This is characterized by generating random single line to ground faults (SLGF), line to line faults (LLF), line to line to ground faults (LLGF), and three-phase faults (LLLGF).
- Location of fault. The fault location is generated performing faults throughout the entire distribution system with a maximum separation distance of 400 meters.

2) Training Process

In this study, linear machine learning models are used for regression. The models are trained through the training set, which contains the samples that map the entire distribution system. In order to select from the database the inputs features X , voltage sag meters must be located. This paper follows the location voltage sag methodology proposed in [10].

a) Location of M meters

The M meters are located through the following procedure:

- Identify main feeder. The main feeder is the three-phase feeder, which starting from the head of the system feeder that has the most extended length L .
- Compute distance D . It is the distance L of the main feeder divided by the number of voltage meters: $D = \frac{L}{M}$.
- Locate the meters on the bus that is closest to the length $L_j = (j - 1) \times D$, where $j = 1, \dots, M$. The distance L_j is measured from the beginning of the main feeder.

b) Input Features

The proposed input features are the voltage magnitude of phases and angles difference. These inputs are linearly independent of each other.

- V_a, V_b, V_c : magnitude of voltage phases.
- Φ_{ab} : difference between angle of phase a and phase b ($\Phi_a - \Phi_b$).
- Φ_{ac} : difference between angle of phase a and phase c ($\Phi_a - \Phi_c$).

c) Targets

The target for the regression problem is the magnitude of voltage sag on the unmonitored buses:

- V_{sag} : voltage sag. It is defined as minimum phase voltage. ($\min\{V_a, V_b, V_c\}$)

Models are repeatedly trained until finding their optimal hyperparameters. After training, models are used to estimate the voltage sag for the unseen testing data. The results are then compared with the expected outputs to validate model performance.

3) Linear Machine Learning Models

Following linear models intended for regression, the target value y is expected to be a linear combination of the input features. The mathematical formulation of each one of the models is as follows:

a) Least Squares

Least squares problem consist in a group of inputs features $X \in \mathcal{R}^{m \times n}$ and targets $y \in \mathcal{R}^m$ and searching a vector $w \in \mathcal{R}^n$ such that Xw is close to y . Closeness is defined as the sum of the squared differences also known as the ℓ_2 -norm squared, $\|Xw - y\|_2^2$. The optimal vector w is founded by solving the follow optimization problem [11]:

$$\min_w \|Xw - y\|_2^2 \quad (1)$$

The coefficient estimates for Least Squares rely on the independence of the features. When features are correlated, and the columns of the features matrix X are linearly dependent, the LS estimate becomes highly sensitive to random errors and produces a significant variance.

b) Ridge Regression

Ridge Regression addresses some of the problems of Least Squares by imposing a penalty on the size of the coefficients. The ridge coefficients minimize a penalized residual sum of squares [12]:

$$\min_w \|Xw - y\|_2^2 + \alpha \|w\|_2^2 \quad (2)$$

The shrinkage parameter $\alpha \geq 0$ controls the amount of variance, prediction accuracy and interpretability of model.

c) Absolute Shrinkage and Selection Operator (LASSO)

LASSO is a linear model that estimates sparse coefficients. For this reason, LASSO is fundamental to the field of compressed sensing. The objective function to minimize is [13]:

$$\min_w \frac{1}{2n_{samples}} \|Xw - y\|_2^2 + \alpha \|w\|_1 \quad (3)$$

The Lasso estimate thus solves the minimization of the least-squares penalty with $\alpha \|w\|_1$ added, where $n_{samples}$ is the number of samples, α is a constant and $\|w\|_1$ is the ℓ_1 -norm of the coefficient vector.

III. CASE STUDY

A. Distribution System Test Case: IEEE-34-bus

The IEEE-34-bus test system is an unbalanced distribution system with three-phase lines represented in black, single-phase lines (a-n) represented in red, and single-phase lines (b-n) represented in green, as illustrated in Fig. 2. In addition, it has two voltage regulators and shunt capacitor banks [14]. This system is simulated by using OpenDSS software.

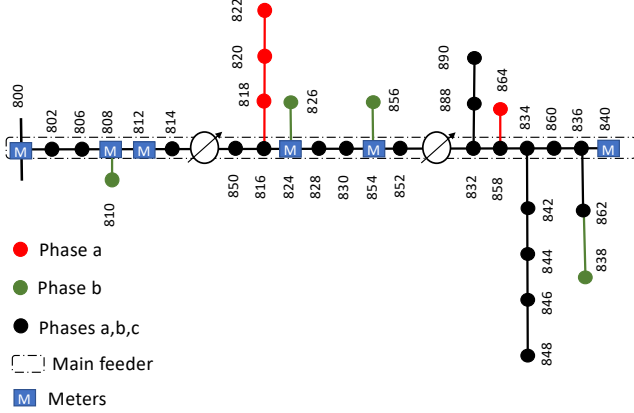


Fig. 2. Illustration of the IEEE-34-bus test system and location of M meters.

B. Training and Validation Data

In OpenDSS via Component Object Model (COM) interface, random faults are simulated to induce voltage sags throughout the entire IEEE-34-bus distribution system. Fault resistance, fault type, and fault location are simulated following the methodology proposed in Section II-B-1. Test data is simulated varying fault resistance from $0.5 - 40.5 \text{ } [\Omega]$ in continuous steps of $2 \text{ } [\Omega]$ and keeping the other variables constant.

C. Linear Models Training Process

Six voltage meters are located in the test system following the methodology proposed in Section II-B-1. Table I shows the identification of the buses where the voltage meters are located.

TABLE I
IDENTIFICATION OF VOLTAGE METERS LOCATION

Voltage Meter	M1	M2	M3	M4	M5	M6
Bus	800	808	812	824	854	840

The input features described in Section II-B-2-a are computed from phase voltages at voltage meters. Having the input features and the respective labels for each sag event, the Least Squares, Ridge, and Lasso models are trained. The implemented models are available in the scikit-learn machine-learning library for Python programming language. In the training of Ridge and Lasso models, a grid-search following and 5-fold cross-validation are performed. Then a fine grid-search is used around the value of α , and in this way, the shrinkage hyperparameter α is found.

D. Linear Models Error

In the training and testing process of each one of the models, the evaluation metric is the mean square error (MSE). If $\hat{y}$ is the predicted value of the i -sample, and y_i is the corresponding true value, then the mean squared error (MSE) estimated over the number of samples ($n_{samples}$) is defined as:

$$MSE(y, \hat{y}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} (y_i - \hat{y}_i)^2 \quad (4)$$

The training mean square error (MSE) is computed using cross-validation. In Fig. 3, the results for Least Squares, Ridge, and Lasso models for training are presented.

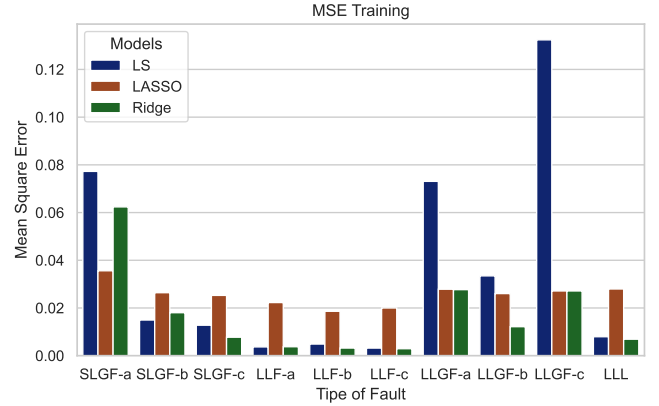


Fig. 3. Mean Square Error for training data.

From the hyperparameters founded in training, the MSE is computed on the test data. In Fig. 4, the results for each type of fault and model are shown.

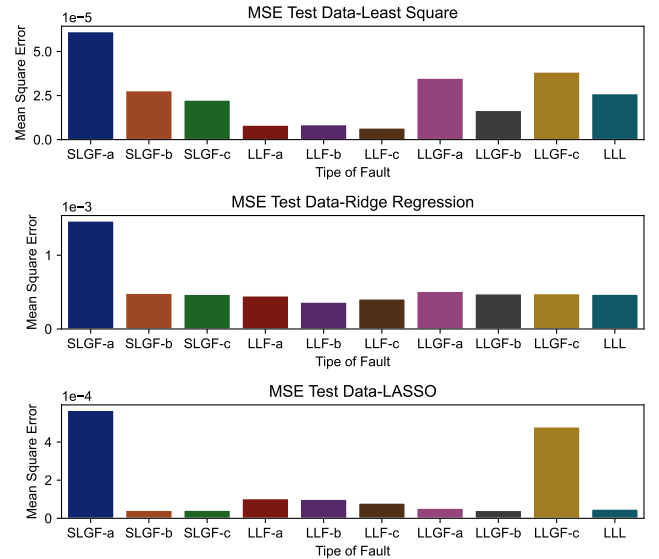


Fig. 4. Mean Square Error for test data.

E. SAG Estimation

This study estimates the voltage sag in unmonitored bus bars using linear models. The profiles of voltage sag estima-

tion in the unmonitored bus bars of IEEE-34 test case systems using the inputs features and the methodology proposed are illustrated in Fig. 5.

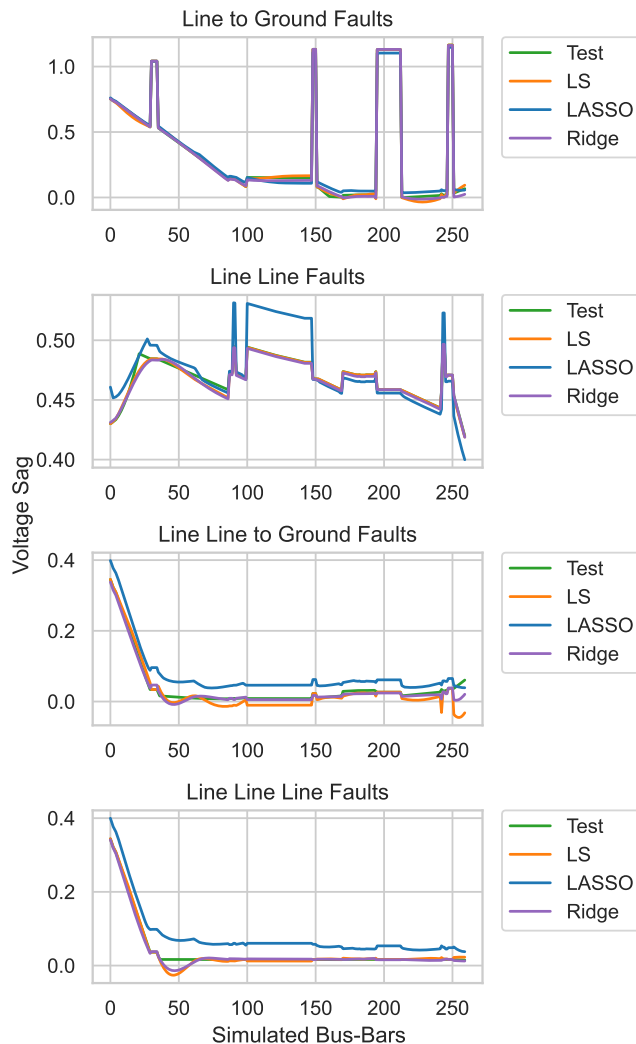


Fig. 5. Voltage Sag Estimation

IV. CONCLUSIONS

This paper proposes a voltage sag estimation approach in distribution systems, using linear models trained from a limited number of monitored buses. The approach proposes a bunch of inputs features, linearly independent, and a methodology for creating the simulation database.

The results demonstrate that linear machine learning model solves the problem with high accuracy on test data. In addition, the results show that the model with the lowest percentage error is Least Square, followed by Ridge, and finally Absolute Shrinkage and Selection Operator (LASSO). Furthermore, the results exhibit that the type of fault with the highest mean square error (MSE) for all models is the single line to ground fault in phase-a.

REFERENCES

- Djokic, S., "Recommended practice for voltage sag and short interruption ride-through testing for end use electrical equipment rated less than 1,000 volts: Ieee p1668/d1p draft standard," 2013.
- Zambrano, X., Hernandez, A., Izzeddine, M., and de Castro, R. M., "Estimation of voltage sags from a limited set of monitors in power systems," *IEEE transactions on power delivery*, vol. 32, no. 2, pp. 656–665, 2016.
- Chan, J. Y., Milanović, J. V., and Delahunty, A., "Risk-based assessment of financial losses due to voltage sag," *IEEE Transactions on Power Delivery*, vol. 26, no. 2, pp. 492–500, 2011.
- Lv, G., Chu, C., Zang, Y., and Chen, G., "Voltage sag source location estimation based on optimized configuration of monitoring points," *CPSS Transactions on Power Electronics and Applications*, vol. 6, no. 3, pp. 242–250, 2021.
- Albu, M. M., Sănduleac, M., and Stănescu, C., "Syncretic use of smart meters for power quality monitoring in emerging networks," *IEEE Transactions on Smart Grid*, vol. 8, no. 1, pp. 485–492, 2016.
- Blanco-Solano, J., Petit, J. F., Ordóñez-Plata, G., Kagan, N., and Almeida, C. F. M., "Voltage sag state estimator based on compressive sensing in distribution systems," *International Journal of Electrical Power & Energy Systems*, vol. 130, p. 106892, 2021.
- Espinosa-Juarez, E. and Hernández, A., "A method for voltage sag state estimation in power systems," *IEEE Transactions on Power Delivery*, vol. 22, no. 4, pp. 2517–2526, 2007.
- Hernández, A., Espinosa-Juárez, E., de Castro, R. M., and Izzeddine, M., "Svd applied to voltage sag state estimation," *IEEE transactions on power delivery*, vol. 28, no. 2, pp. 866–874, 2013.
- Woolley, N., Avendaño-Mora, M., Woolley, A., Preece, R., and Milanović, J., "Probabilistic estimation of voltage sags using erroneous measurement information," *Electric power systems research*, vol. 106, pp. 142–150, 2014.
- Galeano-Suárez, D., Duarte, C., and Blanco, J., "Sarfi-x estimation in unbalanced distribution systems based on synchronized voltage bus measurements and linear support vector machines."
- Boyd, S. and Vandenberghe, L., *Introduction to applied linear algebra: vectors, matrices, and least squares*. Cambridge university press, 2018.
- Rifkin, R. M. and Lippert, R. A., "Notes on regularized least squares," 2007.
- Kim, S.-J., Koh, K., Lustig, M., Boyd, S., and Gorinevsky, D., "An interior-point method for large-scale ℓ_1 -regularized least squares," *IEEE journal of selected topics in signal processing*, vol. 1, no. 4, pp. 606–617, 2007.
- Kersting, W. H., "Radial distribution test feeders," *IEEE Transactions on Power Systems*, vol. 6, no. 3, pp. 975–985, 1991.