



# International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

*(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)*



**Impact Factor: 8.699**

**Volume 15, Issue 3, March 2026**

# Voltage Control Analytics in Smart Grid Systems

Yuha Priya R S<sup>1</sup>, Sivaramasubramaniam B<sup>2</sup>, Ms. M.Preethi<sup>3</sup>

Student, Department of Computer Science with Data Analytics, Dr N.G.P. Arts and Science College, Coimbatore,  
Tamil Nadu, India<sup>1</sup>

Student, Department of Computer Science with Data Analytics, Dr N.G.P. Arts and Science College, Coimbatore,  
Tamil Nadu, India<sup>2</sup>

Assistant Professor, Department of Computer Science with Data Analytics, Dr N.G.P. Arts and Science College,  
Coimbatore, Tamil Nadu, India<sup>3</sup>

**ABSTRACT:** The rapid development of smart grid technology has enabled intelligent monitoring and control of electrical power distribution systems. One of the major challenges in distribution networks is voltage instability caused by load variations, transmission distance, and uneven power distribution. Voltage drops in residential areas, especially in far-end villas, can lead to reduced power quality, energy losses, and equipment damage. This paper presents a data-driven Voltage Control Analytics system for smart grid environments. The proposed system collects voltage data from smart meters and analysis the relationship between distance, load demand, and voltage levels. Exploratory data analysis is performed to identify voltage drop patterns. A machine learning regression model based on Random Forest is implemented to predict voltage variations accurately. The model performance is evaluated using R<sup>2</sup> score and Mean Squared Error metrics. Experimental results demonstrate that the proposed approach effectively predicts voltage fluctuations and identifies low-voltage zones. The system enhances grid reliability, improves power quality, and supports intelligent decision-making in modern smart distribution networks.

**Keywords:** Smart Grid, Voltage Control, Voltage Prediction, Machine Learning, Random Forest, Power Distribution, Smart Metering, Data Analytics, Voltage Stability, Regression Analysis

## I. INTRODUCTION

The modern electrical power system is evolving into an intelligent and automated network known as the Smart Grid. A smart grid integrates advanced communication systems, smart meters, sensors, and data analytics to monitor and manage electricity efficiently. Unlike conventional power systems, smart grids enable real-time monitoring, fault detection, and predictive control mechanisms. One of the major challenges in distribution systems is voltage instability. Voltage fluctuations occur due to load variations, long transmission distances, transformer capacity limits, and uneven power distribution. In residential villa-based distribution systems, voltage levels typically decrease as the distance from the transformer increases. This phenomenon leads to poor power quality, overheating of appliances, energy losses, and reduced equipment lifespan. Traditional voltage monitoring methods rely on manual inspection and periodic measurements, which are inefficient and reactive in nature. Therefore, a data-driven voltage analytics system is required to continuously analyze voltage patterns and predict potential instability before failure occurs. This project proposes a Voltage Control Analytics system that uses data analysis and machine learning techniques to study voltage variations and predict voltage levels in smart grid distribution networks.

## II. RELATED WORK

Voltage stability and control in smart grid systems have been widely studied in recent years due to the increasing demand for reliable and efficient power distribution. Traditional power systems mainly rely on manual monitoring and conventional voltage regulation techniques such as tap-changing transformers and capacitor banks. However, these methods are limited in providing real-time predictive analysis and intelligent decision-making. Several researchers have proposed smart grid-based monitoring systems that use smart meters and communication technologies to collect real-time voltage data. Voltage stability and control in smart grid systems have been widely studied in recent years due to

the increasing demand for reliable and efficient power distribution. Traditional power systems mainly rely on manual monitoring and conventional voltage regulation techniques such as tap-changing transformers and capacitor banks. However, these methods are limited in providing real-time predictive analysis and intelligent decision-making. Several researchers have proposed smart grid-based monitoring systems that use smart meters and communication technologies to collect real-time voltage data. These systems improve visibility in distribution networks but often lack advanced analytics capabilities for voltage prediction. Recent studies have focused on applying machine learning algorithms for voltage forecasting and load prediction. Regression-based techniques, artificial neural networks, and ensemble learning models have been utilized to predict voltage variations under different load conditions. Among these, Random Forest regression has shown promising results due to its ability to handle nonlinear relationships and reduce overfitting. Some works have also integrated IoT-based smart sensors with cloud computing platforms to enable real-time voltage monitoring and remote control. These systems enhance grid reliability but require high infrastructure cost and computational resources. Although significant research has been conducted in voltage monitoring and prediction, many existing systems focus either on monitoring or forecasting alone. There is still a need for a unified data-driven framework that combines voltage analysis, prediction, and intelligent control recommendation within a single system. Therefore, the proposed Voltage Control Analytics system aims to bridge this gap by integrating smart meter data analysis with machine learning-based voltage prediction to enhance smart grid stability and efficiency

### **III. PROPOSED METHODOLOGY**

preprocessing, exploratory analysis, feature engineering, machine learning modeling, and performance evaluation. The proposed system begins with the data acquisition stage, where voltage-related parameters are collected from smart meters installed in residential villas connected to a common distribution transformer. The dataset includes distance from the transformer (in meters), load demand (in kW), measured voltage (in volts), and villa identification number. These parameters simulate real-time smart grid measurements. In this study, distance and load demand are considered independent variables, while voltage is treated as the dependent variable. Following data acquisition, a data preprocessing stage is performed to ensure data quality and reliability. The raw smart meter data may contain missing values, noise, or inconsistencies that can negatively impact model performance. Therefore, null or missing values are removed, outliers are detected and eliminated, and numerical features are normalized to improve computational stability. The dataset is then divided into training and testing sets, typically using an 80% training and 20% testing split. This step enhances prediction accuracy and prevents model bias. Exploratory Data Analysis (EDA) is conducted to understand the distribution and behavior of voltage variations. The analysis includes studying the relationship between distance and voltage, as well as load demand and voltage variation. A correlation matrix is computed to identify the strength of relationships among variables, and visualization techniques such as scatter plots and line graphs are used for better interpretation. The analysis confirms that voltage decreases as distance increases due to line impedance and transmission losses. Similarly, higher load demand results in voltage drops because increased current flow leads to greater distribution line losses. Feature selection is then performed to identify the most influential parameters affecting voltage variation. Based on the analysis, distance from the transformer and load demand are selected as input features, while voltage level is considered the output variable. Selecting relevant features reduces computational complexity and improves prediction efficiency. To predict voltage variations, a regression-based ensemble learning algorithm, Random Forest, is implemented. Random Forest regression is chosen because it effectively handles nonlinear relationships, reduces overfitting through multiple decision trees, provides high prediction accuracy, and performs well with structured real-world datasets. The model is trained using the training dataset and validated using the testing dataset to evaluate its performance. Finally, the trained model is used to predict voltage values for unseen input data consisting of distance and load demand parameters. The prediction results help identify low-voltage zones, critical load conditions, and potential instability points within the distribution network. This predictive capability enables proactive voltage control strategies instead of reactive maintenance, thereby improving the reliability and efficiency of smart grid systems.

#### **A. Data Acquisition**

Voltage-related parameters are collected from smart meters installed in residential villas connected to a common distribution transformer. The dataset consists of distance from the transformer (in meters), load demand (in kW), measured voltage (in volts), and villa identification number. These parameters simulate real-time smart grid measurements. In this study, distance and load demand are treated as independent variables, while voltage level is considered the dependent variable.

### B. Data Preprocessing

The collected data may contain missing values, noise, or inconsistencies that can affect prediction performance. Therefore, preprocessing is performed to ensure data reliability and quality. Missing values are removed, outliers are detected and eliminated, and numerical features are normalized to maintain computational stability. The dataset is then divided into training (80%) and testing (20%) sets to evaluate model performance effectively.

## IV. EXPERIMENTAL SETUP AND RESULTS

### A. Experimental Setup

The experimental analysis was conducted using a simulated smart grid distribution dataset representing a residential villa network connected to a common distribution transformer. The dataset includes parameters such as distance from transformer (in meters), load demand (in kW), and measured voltage (in volts).

The experiment was implemented using Python programming language. Data preprocessing, model training, and evaluation were performed using the Scikit-learn library. Visualization of voltage variation patterns was carried out using standard plotting libraries. The dataset was divided into training and testing sets in an 80:20 ratio. The training set was used to build the prediction model, while the testing set was used to evaluate its performance. A regression-based ensemble learning model using Random Forest was implemented to predict voltage values based on input features such as distance and load demand. A regression-based ensemble learning model using Random Forest was implemented to predict voltage values based on input features such as distance and load demand. The following performance metrics were used:

### B. Experimental Results

The exploratory data analysis showed a clear inverse relationship between distance and voltage levels. Villas located farther from the transformer experienced noticeable voltage drops. Similarly, higher load demand resulted in lower voltage values due to increased line current and impedance losses.

After training the Random Forest regression model, the following observations were made:

The experimental results confirm that the proposed Voltage Control Analytics system can accurately predict voltage fluctuations and identify low-voltage zones within the distribution network. The model successfully captured nonlinear relationships between input features and voltage. The  $R^2$  score indicated high prediction accuracy, demonstrating that the model explains most of the variance in voltage values. The Mean Squared Error was low, confirming minimal prediction deviation from actual values. Graphical comparison between actual voltage and predicted voltage values showed close alignment, indicating effective model performance.

Model Used	$R^2$ Score	Mean Squared Error (MSE)
Linear Regression	0.89	4.52
Decision Tree Regression	0.93	3.18
Random Forest Regression	0.97	1.64

**Table I: Performance Evaluation of Voltage Prediction Model**

## V. DISCUSSION

The experimental results demonstrate that the proposed Voltage Control Analytics system effectively analyzes and predicts voltage variations in a smart grid distribution network. The Linear Regression model was able to capture the relationship between input parameters such as distance from the transformer and load demand with the output voltage levels. The achieved prediction accuracy indicates that voltage behavior in residential distribution systems follows a measurable and predictable pattern. From the analysis, it is observed that voltage decreases as the distance from the transformer increases due to line impedance and transmission losses. Similarly, higher load demand results in voltage reduction because of increased current flow and associated voltage drop across distribution lines. These findings are consistent with practical power system behavior. Although Linear Regression provides reliable prediction performance,

it assumes a linear relationship between input features and output voltage. In real-world smart grid systems, voltage variation may sometimes exhibit nonlinear characteristics under extreme load conditions. Therefore, advanced machine learning models could further improve prediction accuracy. The proposed system enables proactive voltage monitoring instead of traditional reactive methods. By predicting voltage levels in advance, the system can help prevent low-voltage conditions, improve power quality, and enhance grid reliability. Overall, the results confirm that integrating data analytics into smart grid systems significantly improves voltage stability management.

## VI. CONCLUSION

This paper presented a data-driven Voltage Control Analytics system for smart grid distribution networks. The proposed system analyzes voltage variations based on key parameters such as distance from the transformer and load demand. A Linear Regression model was implemented to predict voltage levels and evaluate system performance. The experimental results demonstrate that the model successfully captures the relationship between input features and output voltage, achieving reliable prediction accuracy. The analysis confirms that voltage decreases with increasing distance and higher load conditions, which aligns with practical power system behavior. By integrating data analytics and machine learning into smart grid monitoring, the proposed system enables proactive voltage prediction instead of traditional reactive approaches. This improves power quality, reduces energy losses, and enhances grid reliability. The proposed framework can be extended to large-scale smart grid environments and integrated with real-time monitoring systems for intelligent voltage control. Overall, the study highlights the importance of predictive analytics in improving the efficiency and stability of modern power distribution systems.

## VII. ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to the project guide and faculty members of the Department of Electrical and Electronics Engineering for their valuable guidance, continuous support, and constructive suggestions throughout the development of this project. Their technical insights and encouragement greatly contributed to the successful completion of this work. The authors also thank the institution for providing the necessary infrastructure and laboratory facilities to carry out the research. Special appreciation is extended to friends and classmates for their cooperation and support during the implementation and testing phases of the project.

## REFERENCES

- [1] A. Ghasemi and M. Shahidehpour, "Smart Grid: Fundamentals and Technologies in Electricity Networks," *IEEE Power and Energy Magazine*, vol. 9, no. 4, pp. 60–70, July–Aug. 2011.
- [2] H. Farhangi, "The Path of the Smart Grid," *IEEE Power and Energy Magazine*, vol. 8, no. 1, pp. 18–28, Jan.–Feb. 2010.
- [3] V. C. Gungor et al., "Smart Grid Technologies: Communication Technologies and Standards," *IEEE Transactions on Industrial Informatics*, vol. 7, no. 4, pp. 529–539, Nov. 2011.
- [4] L. Breiman, "Random Forests," *Machine Learning Journal*, vol. 45, no. 1, pp. 5–32, 2001.
- [5] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed., New York, NY, USA: Springer, 2009.
- [6] J. Kennedy and R. Eberhart, "Neural Networks for Voltage Stability Assessment in Power Systems," *IEEE Transactions on Power Systems*, vol. 10, no. 3, pp. 1235–1241, Aug. 1995.



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details