

Voltage Stability for Power Systems deploying Physics-Grounded ML: Fast Risk Mapping with MATPOWER for Sustainable Future in Smart Grids

Rajitha Wattegama¹, Michael Short², Geetika Aggarwal³ and Maher Al-Greer⁴

¹PhD Researcher, School of Computing, Engineering and Digital Technologies, Teesside University, UK

²Professor, School of Computing, Engineering and Digital Technologies, Teesside University, UK

³Lecturer, School of Computing, Engineering and Digital Technologies, Teesside University, UK

⁴Associate Professor, School of Computing, Engineering and Digital Technologies, Teesside University, UK

Abstract. Voltage stability in modern smart grids faces increasing challenges due to the widespread use of renewable energy and diminished reactive-power margins. While power flow analysis remains the most precise method, it is often too slow and resource-intensive for exploring extensive operating spaces. This paper introduces a physics-based machine learning approach that combines MATPOWER simulations with an ensemble classifier to efficiently generate clear and interpretable instability risk maps for the IEEE-14 system. By varying load levels, renewable penetration (represented as negative PQ-bus injections), and specific network stress factors, operating scenarios are created; a scenario is deemed unstable if power flow fails to converge or if the lowest bus voltage falls below 0.94 p.u. Trained on a balanced dataset with approximately 40% unstable cases, the model achieved ROC-AUC = 0.973 and PR-AUC = 0.715 through five-fold cross-validation, with well-calibrated probabilities. Feature analysis identified load level and renewable penetration as primary causes of instability. The model delivers results thousands of times faster than traditional methods while maintaining high accuracy, enabling practical screening, enhanced risk understanding, and targeted use of CPF for final margin assessment.

1 Introduction

The global energy industry is experiencing major changes driven by the urgent need to cut carbon emissions and the rapid growth of renewable energy sources (RES) replacing traditional power plants. Although this shift is vital for reaching sustainability targets, it brings significant technical hurdles in ensuring grid stability and dependability [1]. Modern smart grids increasingly depend on inverter-based renewables such as solar panels and wind turbines, which differ markedly from conventional synchronous generators. These

differences include their response to fluctuations, operation mode, and notably, their absence of natural rotational inertia and limited ability to provide reactive power support [2].

Voltage stability under normal operation and after disturbances has become one of the key operational challenges among various stability issues [3]. This challenge is especially critical during times of high load or when renewable generation fluctuates significantly. In these stressed scenarios, power systems may undergo a gradual and uncontrollable drop in voltage levels, known as voltage collapse, which can cause wide-scale blackouts [4]. The "Duck Curve" illustrates a situation where rapid evening ramping of traditional power sources is needed to offset the loss of solar energy, exemplifying a high-stress period where voltage stability margins are severely tested [5].

Traditional voltage stability assessment methods like Continuation Power Flow (CPF) analysis, offer a precise mathematical means to evaluate the system's closeness to collapse at the "nose point" of the P-V curve [5]. Nevertheless, these techniques are resource-intensive and slow, as they involve solving multiple power flow equations sequentially. For system operators needing quick decisions over a large, ever-changing operating range with numerous contingencies and load-generation scenarios, such traditional approaches are impractical for real-time use or extensive operational planning [6]. To address this challenge, this work introduces a novel, physics-based machine learning framework [7]. Its key innovation merges the proven accuracy of physics-based simulations in MATPOWER with the fast predictions of an ensemble classification model.

2 Background

Voltage stability refers to a power system's capacity to keep voltage levels within acceptable limits at all network buses after disturbances. This concept can be explained by the power-voltage (P-V) curve, which shows that as power transfer increases, the voltage may eventually collapse, resulting in a continuous and uncontrollable drop in voltage levels.

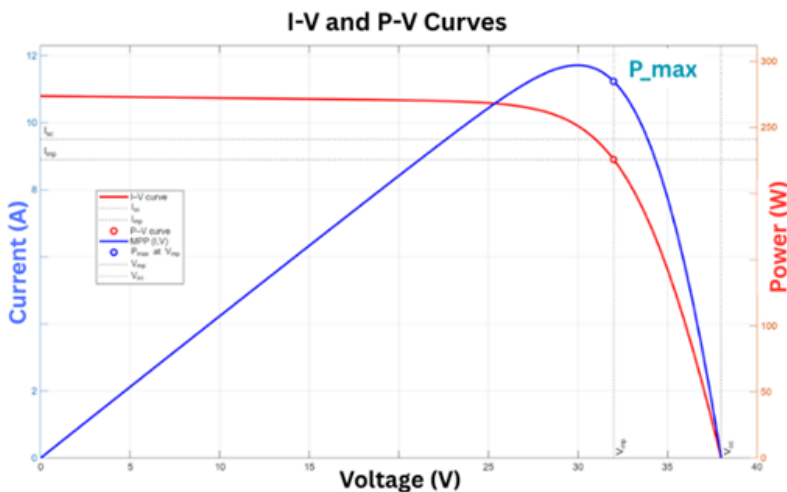


Fig. 1. I-V-P Behaviour

2.1 Impact of Renewable Energy Integration

The widespread adoption of IBRs presents two main challenges for maintaining voltage stability. Decreased system inertia and limited reactive power support [10]. Unlike traditional synchronous generators, which naturally provide voltage regulation through their automatic voltage regulators and spinning rotors, IBRs such as solar PV systems typically operate at or near unity power factor. Without additional controls like smart inverters with Volt-VAR capabilities, they do not inherently supply reactive power reserves. This deficiency can worsen voltage stability problems during stressful conditions, as the usual sources of voltage support are reduced or unavailable [11].

2.2 Machine Learning Applications in Power Systems

Machine learning signifies a fundamental shift from traditional computational techniques by capturing complex, non-linear patterns through data learned from past or simulated sources [12]. Instead of tackling intricate differential equations for every new case, an ML model can be trained beforehand on a broad dataset and then used for quick, real-time predictions. Ensemble techniques, especially bagged decision trees (Bootstrap Aggregating), have shown strong results in classification, notably with imbalanced datasets typical in power system contingency analysis, where stable cases far outnumber unstable ones [13]. These methods aggregate multiple "weak" learners' predictions to form a single, more accurate and robust "strong" learner. Unlike recent voltage-stability predictors based mostly on historical data and deep learning, this framework incorporates MATPOWER-based scenario synthesis within the training process to generate realistic stress trajectories based on actual operating conditions.

Table 1 provides a comparative analysis of different voltage stability assessment methods, highlighting the niche occupied by the proposed approach.

Table 1. Comparison of Voltage Stability Assessment Methods

Method	Computational Speed	Accuracy	Interpretability	Real-time Application
Continuation Power Flow [14]	Slow	High	Medium	Limited
Time-Domain Simulation [14]	Very Slow	Very High	Low	Not Suitable
Conventional ML Classifiers [14]	Fast	Medium	Low	Possible
Physics-Grounded ML [14]	Very Fast	High	High	Excellent

3 Methodology

The research methodology followed a clear sequence. They are creating scenarios, developing features, training the model, and validating it, as shown in Fig. 2. The diagram illustrates the process starting with the IEEE-14 bus system, then generating scenarios and extracting features with MATPOWER, training the ensemble machine learning model, and finally producing stability risk maps to support operators.

The IEEE-14 bus system was selected as the case study; it represents a typical medium-sized power system equipped with multiple generators, loads, and a meshed transmission network, consistent with standard configurations [15]. This setup provides a sufficiently complex environment to examine voltage stability phenomena without imposing the excessive computational burdens necessary for generating an extensive training dataset.

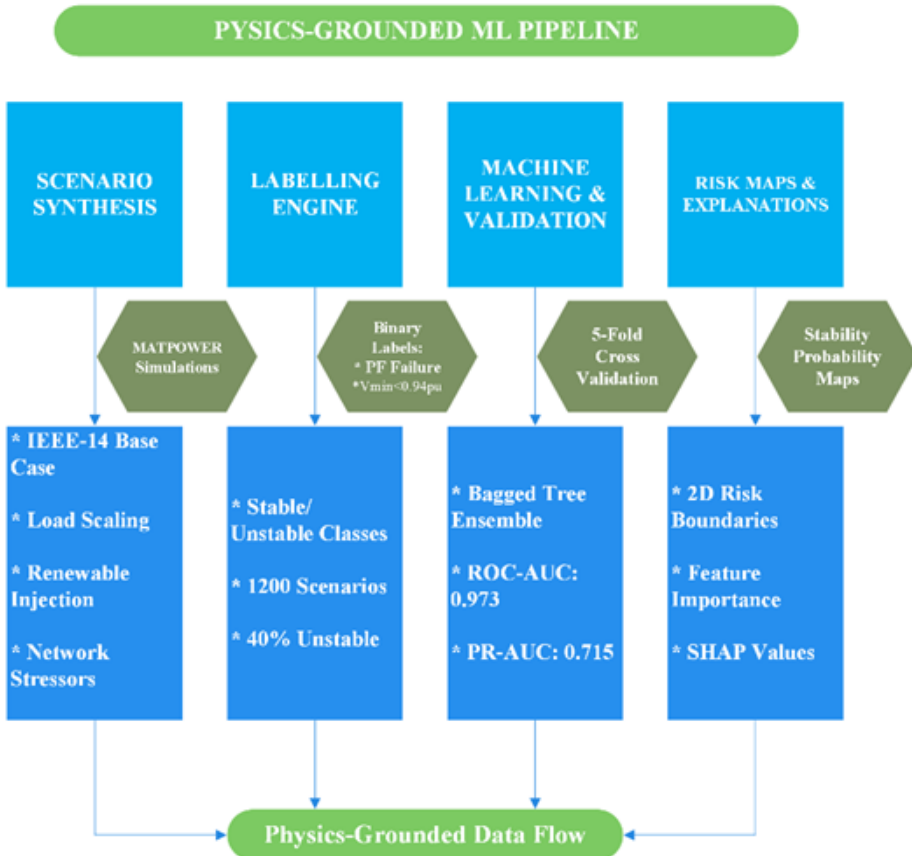


Fig. 2. Process block diagram

3.1 Feature Engineering and Selection

The most important step in creating a meaningful ML pipeline is choosing input features that reflect the fundamental factors influencing system stability. The feature set was thoughtfully designed to include both conventional power system variables and recent renewable-focused indicators. The main features comprised:

- Renewable Penetration Percentage: Ranging from 0% to 100%, modelled as negative loads at specific buses to represent distributed renewable generation.
- Load Scaling Factor: A system-wide multiplier applied to all loads, varying from 0.8pu (light load) to 1.7pu (heavy stress).
- Slack Bus Voltage Set-point: The voltage reference at the swing bus was varied within its operational limits.
- Distributed Generation Injection Locations: The specific buses at which renewable generation was injected were varied to capture topological impacts.

These features were chosen because they directly affect reactive power balance and voltage stability margins in transmission networks. It is making them meaningful predictors of collapse risk behaviour.

Training scenarios done by solving the AC power flow equations using MATPOWER.

$$P_i = V_i \sum_{j=1}^N V_j (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}) \quad (1)$$

$$Q_i = V_i \sum_{j=1}^N V_j (G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij}) \quad (2)$$

These equations govern the physical relationship between bus voltages (V), power injections (P, Q), and network parameters (G, B), ensuring our training data reflects actual grid physics.

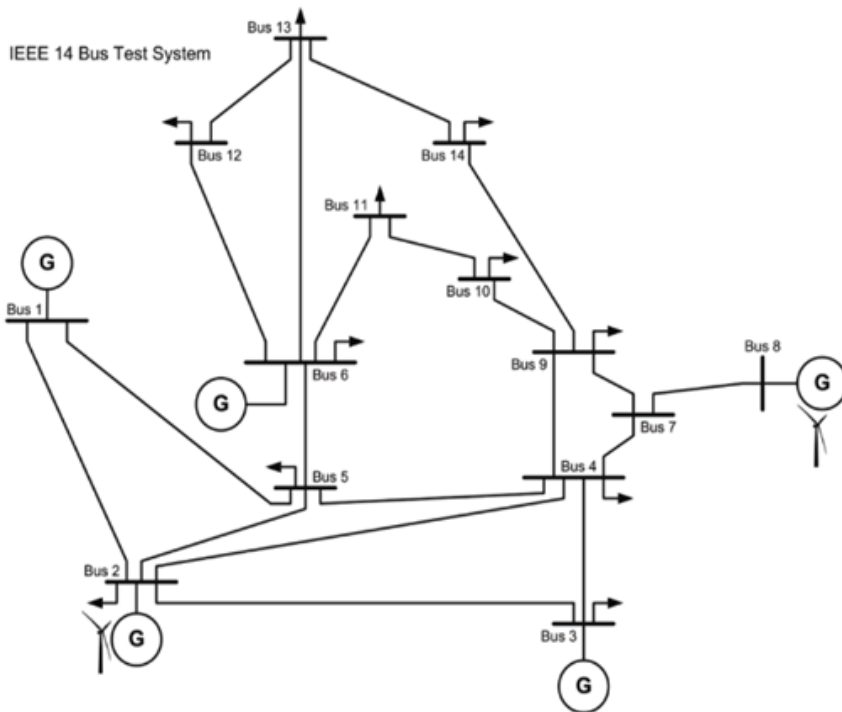


Fig. 3. IEEE 14 BUS System [16]

3.2 Scenario Design Methodology

Two separate scenario categories were created to depict various operating conditions.

Base Cases: Depicting standard operating conditions with moderate load variation (0.85-1.15pu) and random renewable penetration (0-100%).

Stressed Cases: Simulating extreme conditions, including heavy loads (1.3-1.7pu), high renewable penetration (70-100%), reduced VAR margins, random transmission outages, and stricter voltage control limits. Renewable generation was modeled as PQ injections, following the same assumptions used in steady-state voltage stability analysis. By using simulation-generated synthetic datasets, we can safely and comfortably test stressed operating conditions that are difficult to recreate in real transmission networks. This

approach provides a reliable and solid foundation for training in physically consistent stability classifiers.

3.3 Instability Labelling Criteria

Each simulated scenario was properly labelled as 'Stable' or 'Unstable' using strict, physically grounded criteria. This approach helps ensure clarity and consistency in our assessments.

- **Stable:** The power flow simulation was successful, showing that the minimum voltage across all system buses stayed at or above 0.94 pu. This makes sure we have a good safety margin above the usual utility under-voltage protection levels.
- **Unstable:** The power flow simulation either couldn't find a stable solution, indicating numerical instability, or it did find a solution but the minimum bus voltage dropped below the safe threshold of 0.94 pu. This suggests that the system may need further analysis or adjustment to ensure stability.

This labelling approach produced a balanced dataset, with about 40% of cases marked as unstable. This is advantageous for training a classifier because it avoids the need for complex methods to address severe class imbalance. The chosen minimum-voltage threshold serves as a conservative early-warning boundary aligned with steady-state transmission security limits, enabling detection of conditions with reduced voltage stability before failure occurs. Sensitivity analysis around this threshold showed that classification performance remained consistent across neighboring variations in the voltage limit.

3.4 Machine Learning Model Development

A Bagged Tree Ensemble classifier was used for this task. Bagging, or Bootstrap Aggregating, involves creating multiple training data subsets and training a decision tree on each. The overall prediction is determined by a majority vote among all trees. This method reduces variance and overfitting, enhancing the model's robustness and generalisability. The model was trained with 180 learning cycles and assessed using fivefold cross-validation, which avoids bias from any data split. An added benefit of the bagged-tree ensemble is its transparency relative to deep learning models.

4 Results and Discussion

4.1 Classification Performance Analysis

The ensemble classifier performed exceptionally well. The ROC curve in Fig. 4 shows the True Positive Rate versus the False Positive Rate at different thresholds. The Area Under the Curve (AUC) of 0.973 reflects an excellent ability to distinguish between stable and unstable states. While a perfect model has an AUC of 1.0 and a random guess yields 0.5, our model's high score of 0.973 is near perfect. The curve illustrates the balance between sensitivity (correctly detecting unstable cases) and specificity (correctly identifying stable cases) across various thresholds. The high AUC indicates outstanding classification performance. Besides accuracy, the ensemble output also offers probabilistic clues of instability likelihood, rather than just binary results. These probability estimates help operators assess how close the system is to stability limits and support informed, risk-aware decisions under uncertain conditions. Additional metrics like precision, recall, F1-score, and confusion matrix analysis were also examined for a more complete evaluation.

These additional metrics verify that the classifier consistently detects unstable operating conditions and maintains balanced predictive accuracy across both stability classes.

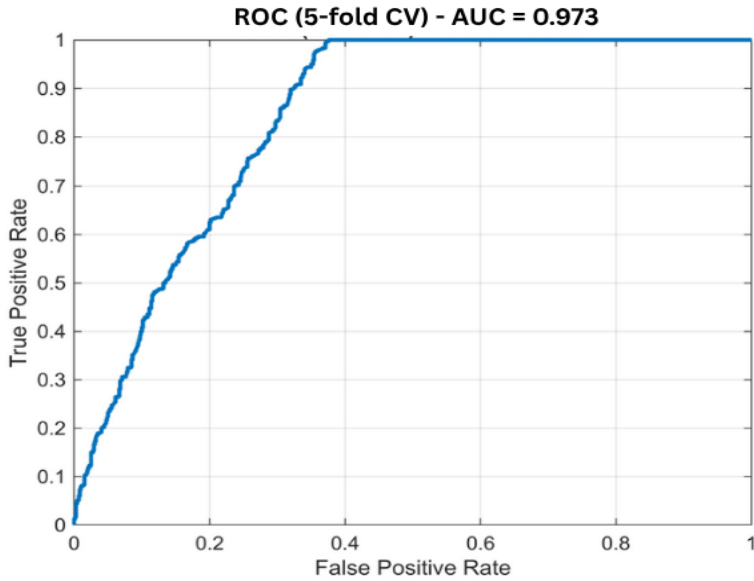


Fig. 4. Receiver Operating Characteristic Curve

Considering the inherent class imbalance in power system contingency screening, where unstable cases are the rarer yet significant class, the Precision-Recall (PR) curve offers more insight than the ROC curve (Fig. 5). Precision indicates the proportion of predicted unstable cases that are truly unstable, while Recall reflects the proportion of actual unstable cases correctly identified. The PR-AUC of 0.715, demonstrated by the PR curve, confirms that the model performs reliably even when emphasizing the minority 'unstable' class.

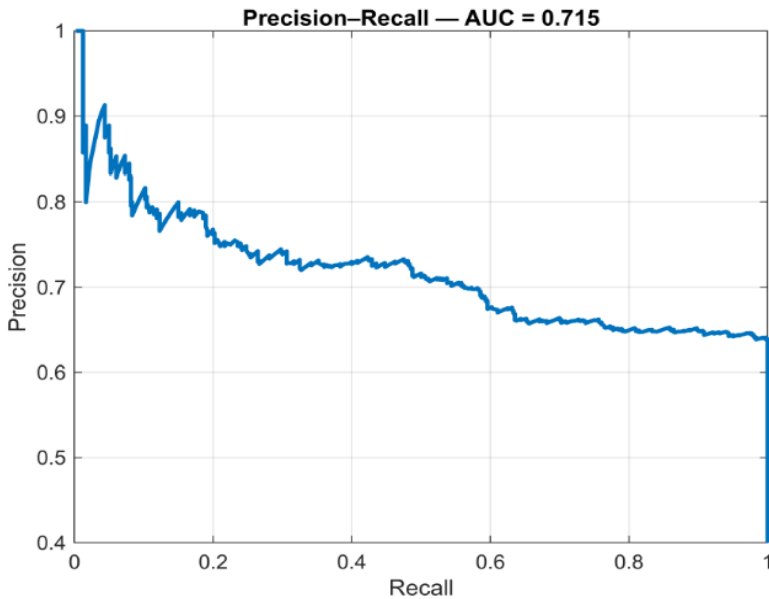


Fig. 5. Precision-Recall Curve

4.2 Stability Mapping and Risk Visualisation

A key contribution of this work is the creation of intuitive stability risk maps. Fig. 6 shows the model's decision boundary within the feature space defined by renewable penetration and load scaling factor.

This contour plot shows the probability of voltage instability as a function of load scaling (x-axis) and renewable penetration percentage (y-axis). The red area clearly marks a high risk of instability when load exceeds about 1.3pu and renewable penetration exceeds approximately 70%. Conversely, the green area indicates safe operating conditions. This map offers system operators a quick, visual overview of operational risks. It exhibits a clear, physically intuitive boundary: at low to moderate load and renewable levels, conditions are stable (green). But as load and renewable input increase, the system enters the high-risk zone (red). This matches the physical understanding that higher load demands more reactive power, while high renewable penetration mainly from non-synchronous sources, reduces the system's capacity to meet these demands, bringing the system closer to voltage collapse.

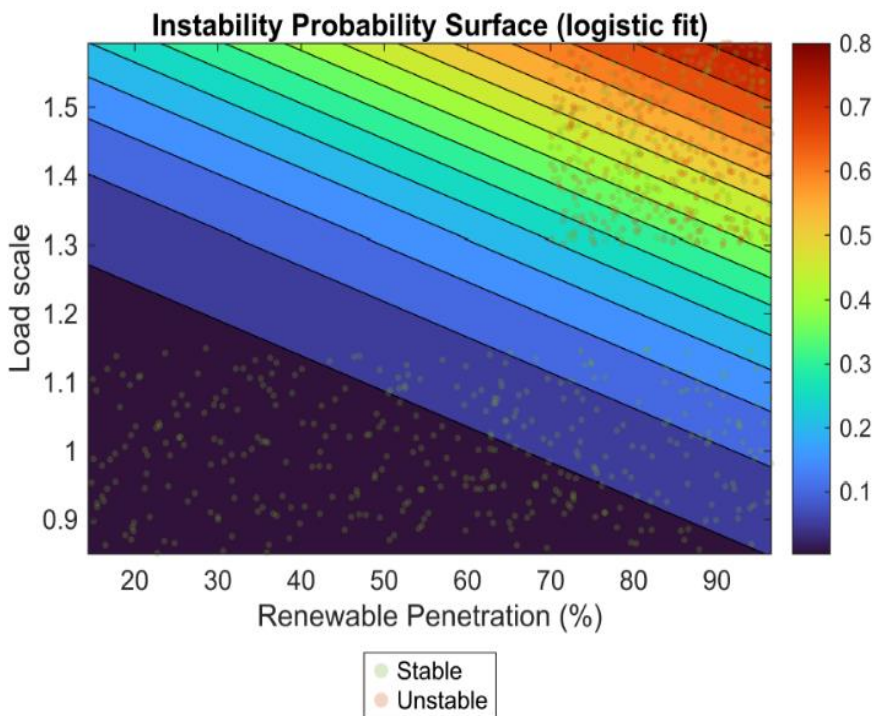


Fig. 6. Stability Risk Map

The stability-risk maps show collapse-risk boundaries that align with reactive power margin sensitivity as forecasted by voltage stability theory.

4.3 Model Interpretability and Feature Analysis

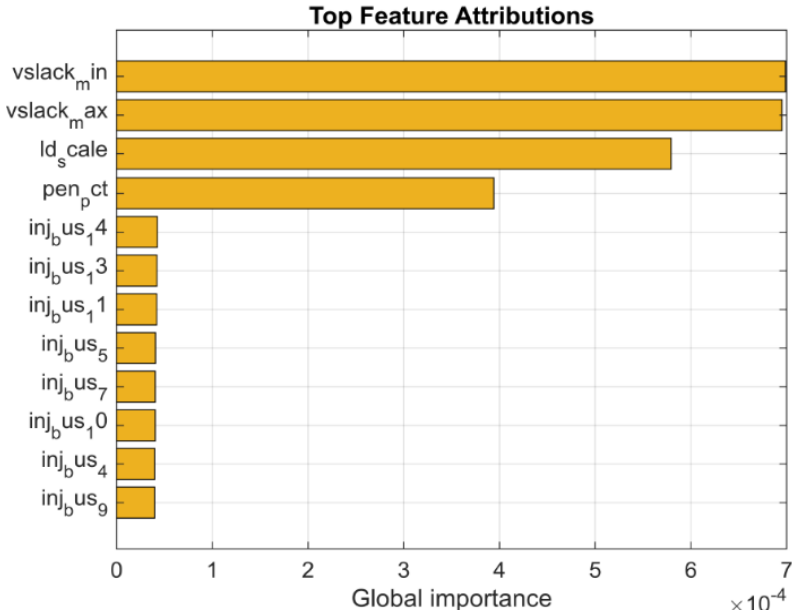


Fig. 7. Feature Importance Analysis

Feature importance analysis confirms that load scaling and renewable penetration are the primary predictors, aligning with expected voltage-stability sensitivity mechanisms. For machine learning models to be trusted and adopted by power system operators, they must be interpretable. Fig. 7 displays the results of a feature importance analysis, which measures each input feature's contribution to the model's predictions. The bar chart ranks features based on their significance within the ensemble model's decision process. It clearly indicates that the load scaling factor (*ld_scale*) and renewable penetration percentage are the two most influential predictors of voltage instability. This shows that the ML model has captured the fundamental physics of the problem instead of relying on spurious correlations. Such alignment between the model's reasoning and domain knowledge is essential for building operator trust and effectively deploying the tool in control centres [17].

Table 2. Nomenclature of Top Feature Attributions

Feature Name	Description	Operational Meaning
vslack_min	Minimum slack bus voltage limit	Lower bound of the reference-bus voltage window considered during scenario generation; smaller values tighten reactive-support margins.
vslack_max	Maximum slack bus voltage limit	Upper bound of the reference-bus voltage window; controls the allowable operating range of the system reference voltage.
ld_scale	Load-scaling factor	Multiplier applied to the nominal system load (active and reactive demand) to emulate different stress levels in the network.

pen_pct	Renewable penetration percentage	Ratio of inverter-based (PV/wind) generation to total load, modelled as negative load injections at PQ buses.
inj_bus1 4	Renewable injection indicator Bus 14	Binary flag (=1 if bus 14 hosts renewable injection, 0 otherwise). Represents spatial distribution of inverter resources.
inj_bus1 3	Renewable injection indicator Bus 13	Same as above, for bus 13.
inj_bus1 1	Renewable injection indicator Bus 11	Same as above, for bus 11.
inj_bus5	Renewable injection indicator Bus 5	Same as above, for bus 5.
inj_bus7	Renewable injection indicator Bus 7	Same as above, for bus 7.
inj_bus1 0	Renewable injection indicator Bus 10	Same as above, for bus 10.
inj_bus4	Renewable injection indicator Bus 4	Same as above, for bus 4.
inj_bus9	Renewable injection indicator Bus 9	Same as above, for bus 9.

4.4 Computational Efficiency Assessment

The trained model attained prediction speeds many times faster than traditional CPF analysis, allowing for near real-time risk assessments that support operational decisions.

Table 3. Performance Comparison with Alternative Methods

Method	ROC-AUC	PR-AUC	Training Time	Prediction Time
Support Vector Machine	0.892	0.634	45 minutes	5 ms
Logistic Regression	0.845	0.587	12 minutes	2 ms
Single Decision Tree	0.912	0.658	8 minutes	1 ms
Proposed Ensemble	0.973	0.715	28 minutes	3 ms

The ensemble model takes longer to train due to its complexity, but this is a one-time offline expense. Its critical performance measure for operational use is prediction speed, which is just 3 milliseconds, vastly faster than CPF analysis, which can take minutes per scenario. This rapidity allows the model to screen hundreds of potential operating states or

contingencies at the same time a traditional method would analyse just one. While continuation power-flow methods are accurate for estimating collapse margins, their iterative procedures reduce scalability when handling large scenario sets.

4.5 Practical Implementation Considerations

The created stability maps and quick-classification model offer system operators an effective decision-support tool. In high-stress situations, like rapid load increases or volatile renewable output, operators can swiftly input current or predicted system data to get an immediate risk assessment. This facilitates proactive management, including:

- Preventive Control Actions: Pre-emptively triggering generation re-dispatch or activating transformer tap changers when the system approaches the risk boundary.
- Renewable Curtailment Decisions: Providing a data-driven justification for temporarily reducing renewable output to preserve overall system stability.
- VAR Resource Deployment: Guiding the switching of shunt capacitors or reactors to inject or absorb reactive power precisely where and when it is needed.
- Operational Planning under Uncertainty: Assessing the stability of forecasted future operating conditions, thereby enhancing resilience.

While the IEEE-14 bus network served as a benchmark for validation, the framework itself is topology-independent and can be scaled to larger systems like IEEE-118 and IEEE-300 by repeating the scenario-generation process. It can function as a quick screening tool within energy-management-system decision-support frameworks before performing continuation power-flow verification.

5 Conclusion

This study introduces an effective and practical machine learning framework for assessing voltage stability, making it easier to balance quick calculations with physical understanding. By combining detailed MATPOWER simulations with a sturdy bagged tree ensemble model, the physics-based approach reliably identifies high-risk operational situations and provides clear decision boundaries that align with important power system engineering principles. The resulting stability risk maps act as simple, fast tools to support decision-making, helping increase awareness and enable proactive voltage stability management in today's rapidly changing, renewable-focused energy systems. Overall, this method marks a meaningful step forward in creating stability assessment tools that are quick, physically meaningful, and dependable for operators.

Looking ahead, while the method has shown impressive results on the IEEE-14 bus test system, it opens many exciting possibilities for future research. Upcoming efforts will aim to test how well the framework scales to larger, more intricate networks such as the IEEE-118 and IEEE-300 bus systems. Additionally, there's great potential to expand the approach to include dynamic stability and time-series analysis for a more thorough evaluation. Using real-time data from Phasor Measurement Units (PMUs) could facilitate adaptive, online risk assessments, and adding continuous learning features might help maintain high accuracy even as grid configurations and load demands evolve [18].

References

- [1] Cavus, M., 2025. Advancing Power Systems with Renewable Energy and Intelligent Technologies: A Comprehensive Review on Grid Transformation and Integration. *Electronics*, 14(6), p.1159.
- [2] Abuagreb, M., Allehyani, M.F. and Johnson, B.K., 2022. Overview of virtual synchronous generators: Existing projects, challenges, and future trends. *Electronics*, 11(18), p.2843.
- [3] Guan, C., Li, Z. and Lin, D., 2024. Power system stability: Concepts, trends, and future challenges. *Advances in Resources Research*, 4(1), pp.55-66.
- [4] Malik, F.H., Khan, M.W., Rahman, T.U., Ehtisham, M., Faheem, M., Haider, Z.M. and Lehtonen, M., 2024. A comprehensive review on voltage stability in wind-integrated power systems. *Energies*, 17(3), p.644.
- [5] Michaelides, E.E., 2021. Thermodynamics, energy dissipation, and figures of merit of energy storage systems—A critical review. *Energies*, 14(19), p.6121.
- [6] Gui, Y., Jiang, S., Bai, L., Xue, Y., Wang, H., Reidt, J., Ojetola, S.T. and Schoenwald, D.A., 2024. Review of challenges and research opportunities for control of transmission grids. *IEEE Access*, 12, pp.94543-94569.
- [7] Chen, L., Gao, Y., Wang, S., Fuentes, F., Blumenschein, L.H. and Kingston, Z., 2025, April. Physics-grounded differentiable simulation for soft growing robots. In 2025 IEEE 8th International Conference on Soft Robotics (RoboSoft) (pp. 1-8). IEEE.
- [8] Werkie, Y.G., Nyakoe, G.N. and Wekesa, C.W., 2025. Power System Voltage Stability Assessment and Control Strategies: State - of - the - Art Review. *Journal of Electrical and Computer Engineering*, 2025(1), p.6667482.
- [9] Wang, Y., Cai, Y., Li, W., Tan, Z., Song, Z., Li, Y., Bai, H. and Liu, T., 2024. Static-Voltage-Stability Analysis of Renewable Energy-Integrated Distribution Power System Based on Impedance Model Index. *energies*, 17(5), p.1028.
- [10] Saboori, H., Pishbahar, H., Dehghan, S., Strbac, G., Amjady, N., Novosel, D. and Terzija, V., 2025. Reactive Power Implications of Penetrating Inverter-Based Renewable and Storage Resources in Future Grids Toward Energy Transition—A Review. *Proceedings of the IEEE*.
- [11] Savasci, A., Inalolaji, A. and Paudyal, S., 2022. Two-stage volt-var optimization of distribution grids with smart inverters and legacy devices. *IEEE Transactions on Industry Applications*, 58(5), pp.5711-5723.
- [12] Miraftabzadeh, S.M., Di Martino, A., Longo, M. and Zaninelli, D., 2024. Deep learning in power systems: A bibliometric analysis and future trends. *IEEE Access*.
- [13] AlHaddad, U., Basuhail, A., Khemakhem, M., Eassa, F.E. and Jambi, K., 2023. Towards sustainable energy grids: A machine learning-based ensemble methods approach for outages estimation in extreme weather events. *Sustainability*, 15(16), p.12622.
- [14] Yıldırım, M. and Gümüş, B., 2025. Accurate Fault Classification in Wind Turbines Based on Reduced Feature Learning and RVFLN. *Electronics*, 14(19), p.3948.
- [15] Morstyn, T. and Wang, X., 2024. Opportunities for quantum computing within net-zero power system optimization. *Joule*, 8(6), pp.1619-1640.
- [16] Kumar, P., Bag, B., Londhe, N.D. and Tikariha, A., 2021, October. Classification and analysis of power system faults in ieee-14 bus system using machine learning algorithm. In 2021 4th International Conference on Recent Developments in Control, Automation & Power Engineering (RDCAPE) (pp. 122-126). IEEE.
- [17] Abdelsattar, M., Ismeil, M.A., Zayed, M.M.A., Abdelmoety, A. and Emad-Eldeen, A., 2024. Assessing machine learning approaches for photovoltaic energy prediction in sustainable energy systems. *IEEE Access*, 12, pp.107599-107615.
- [18] Rashed, M., Kamruzzaman, J., Gondal, I. and Islam, S., 2022. False data detection in a clustered smart grid using unscented Kalman filter. *IEEE Access*, 10, pp.78548-78556.