Developing a Statistical Risk Assessment and Grid Prediction Tool for Power System Reliability

Chetan Saha | Supervised by Dr. Blake Shaffer, Faculty of Arts, and Joe Calnan, Canadian Global Affairs Institute

Research Question

How can system operators anticipate **Energy Emergency Alerts (EEAs)** and other instability events hours ahead as renewable penetration rises, demand patterns shift, and extreme weather intensifies?

Project Overview

This research project built a Grid Risk Assessment Tool using datasets from AESO, ERCOT, and CAISO (2019–2024). A MySQL database consolidates system frequency, pricing, generation mix, intertie flows, and operator-declared events. We engineered predictive features (ramp rates, reserve adequacy, renewable share, frequency deviations, lagged event indicators) and tested Logistic Regression, Random Forest, and XGBoost against ROC-AUC, precision, and recall.

Research

Research Context

- Reliable electricity = critical for economy & safety
- Renewables, demand shifts, extreme weather stress the grid
- Traditional tools miss early-warning signals

Research Avenues

- Intermittent renewables reduce predictability
- Low reserves increase blackout risk
- Interdependence exposes cross-border fragility
- Statistical and machine-learning models trained on operational and market data can provide early-warning signals of grid instability.

Research Objectives

- Predictive risk tool using ISO data (AESO, ERCOT, CAISO)
- Machine learning: Logistic Regression, Random Forest, XGBoost
- Early-warning alerts to guide operators & policy

CANADIAN GLOBAL AFFAIRS INSTITUTE
INSTITUT CANADIEN DES AFFAIRES MONDIALES

mitacs

Data and Methodology

Data Sources

- ISOs: AESO (2019–2024), ERCOT (2021–2023), CAISO (2022–2024)
- Variables: frequency (2-sec upcycled), pool/SMP prices, generation mix, intertie flows, demand, reserves, EEA/alerts
- Storage: Structured MySQL database for clean, reproducible feature extraction

Feature Engineering

- Supply-side: renewable share, dispatchable capacity, intertie support, reserve margins
- Operational flexibility: ramp rates, dispatchable ratio, regulation headroom
- Demand-side: forecast errors, price spikes, pool–SMP spreads
- Stress indicators: frequency <59.95 Hz, net load gaps, lagged event flags

Models

- **Logistic Regression**: interpretable baseline
- **Random Forest**: balanced accuracy + interpretability → *selected model*
- **XGBoost**: tested for precision trade-offs

Validation

- Metrics: ROC-AUC, precision, recall, F1 score
- Time-aware cross-validation to avoid data leakage
- Case studies: cold snaps, high renewable volatility, reserve shortages

Discussion and Operator Insights

- Predictive alerts improve situational awareness and reduce reliance on reactive interventions.
- Integration into ISO dashboards enables risk scoring, tiered alerts, and visual explanations.
- Trade-offs:
  - High recall = capture most events, but more false alarms.
  - High precision = fewer false alarms, but risk of missing events.
- Opportunities for system operator training: custom alerts, scenario exercises, and cross-functional learning.

Policy and Future Impacts

- Aligns with NERC reliability standards; complements contingency analysis and adequacy planning.
- Supports market design: dynamic ancillary service sizing, flexibility markets, and demand response triggers.
- Recommendations: encourage adoption in AESO, ERCOT, CAISO, expand to PJM, IESO, and SPP.
- Future work:
  - Integrate PMU data for sub-second insights.
  - Develop dashboards & APIs for operators.
  - Explore Bayesian/ensemble models for rare-event prediction.
  - Pathway toward a commercial SaaS platform for ISOs, utilities, and regulators.

Results and Findings

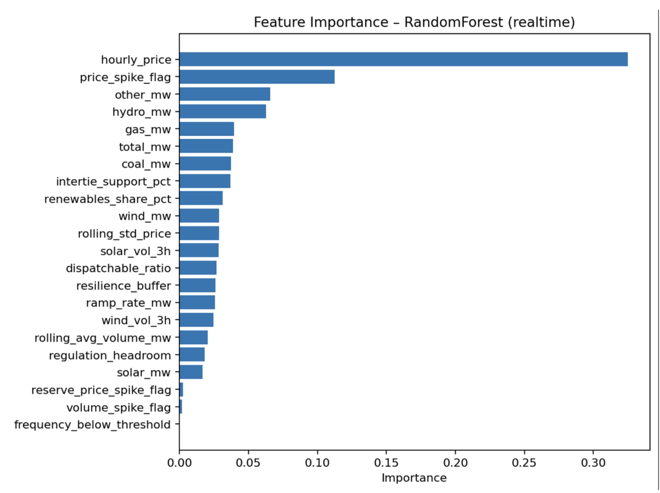

Figure 1: Feature Importance graph of Realtime Random Forest Classifier Model

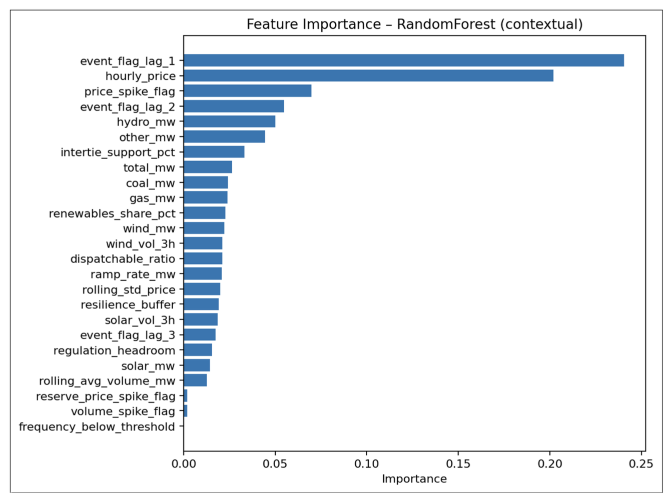

Figure 2: Feature Importance graph of Contextual Random Forest Classifier Model

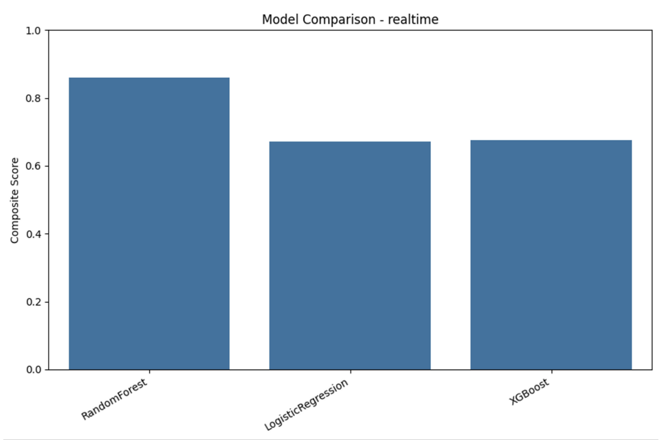

Figure 3: Comparison graph of the Statistical Models for Real-time use

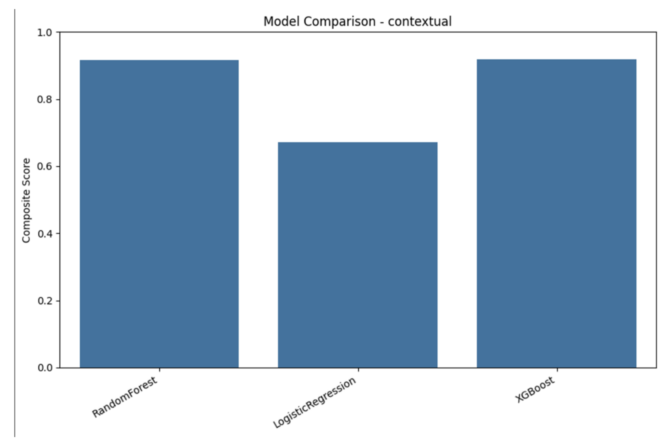

Figure 4: Comparison graph of the Statistical Models for Contextual use

- Random Forest (real-time) composite = 0.86; Contextual = 0.92
- Contextual features outperform real-time-only sets.
- Random Forest offers high discrimination with balanced precision/recall suitable for operational use.
- Drivers of risk align with grid physics and market conditions: tight reserves + steep ramps + high renewable share with weak intertie support.

Conclusion

- Predictive modeling adds a statistical early-warning layer to traditional contingency analysis. The tool surfaces high-risk periods for proactive actions (reserves, intertie scheduling, demand response) and informs policy/market design(ancillary services, flexibility procurement).
- Next Steps: integrate PMU and weather signals; expand ISO coverage; deploy a dashboard/API for operators and regulators; evaluate–benefit of alert thresholds.

Works Cited

- Team, B. (2021, January 21). *Ukraine-Baltic Energy Cooperation in the context of Russian hybrid aggression*. Baltic Security Foundation. https://balticsecurity.eu/ukraine_baltic_energy_cooperation/
- Pérez, E., Gómez-Expósito, A., & Ramos, A. (2020). Frequency-based early warning systems for power grid instability. Electric Power Systems Research, 180, 106141. https://doi.org/10.1016/j.epsr.2019.106141
- Oh, S., Kong, J., Choi, M., & Jung, J. (2020). Data-driven prediction method for power grid state subjected to heavy-rain hazards. Applied Sciences, 10(14), 4693. https://doi.org/10.3390/app10144693