Automating Detection and Diagnosis of Faults, Failures, and Underperformance in PV Plants

Applications of AI/ML

Scott Sheppard, Turbine Logic
Lead Engineer – Renewables Monitoring

Daniel Fregosi, Wayne Li, EPRI
Kamran Paynabar, Georgia Tech

Solar Applications of Artificial Intelligence and Machine Learning
Oct 31, 2023
The Problem…Using Available Data to Full Potential

- Large plant may have “10s” of inverters
- Limited sensor data available to detect DC side faults

- Each Inverter contains 10s to 100s of Combiner Boxes
- Each CB may have current and voltage measurements
- Can be used for diagnostics – not typically used today

Can we couple physics-based modeling and AI to better detect and localize string level faults?
Motivation

- **Goal:** Provide a better diagnostic solutions with fewer false alarms and actionable M&D insight

- **SUBTLE FAILURES ACROSS THE DC COLLECTOR FIELD OFTEN GO UNDETECTED FOR LARGE AMOUNTS OF TIME**
  - Determining the source of the failure is time consuming
  - Aerial inspections
    - Performed to detect small-scale faults across the DC collector field
    - Are typically performed infrequently
    - Current gold standard

- **MODEL-DRIVEN APPROACH USES BIG DATA AVAILABLE AT PV SITES ENABLES REAL-TIME DETECTION**
  - Improves ability to detect faults while reducing presence of false alarms
  - Improves ability to locate faults to more specific hardware components
  - Provides further aid in diagnosing cause of underperformance

M&D centers have abundant data available, how can it better be used for detection of subtle faults?
Modeling Approach – Fault Detection

Physics-based models coupled with AI to identify failed string and tracker outages

- ML/AI used to determine repeated hardware configurations through plant
  - Enables automatic setup of plant layout, reducing time spent by 90%
- Fault detection driven by feature extraction of measured and modeled signals

<table>
<thead>
<tr>
<th>Sensitivity</th>
<th>TPR</th>
<th>FPR</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>65%</td>
<td>19%</td>
<td>0.57</td>
</tr>
<tr>
<td>Med</td>
<td>49%</td>
<td>10%</td>
<td>0.54</td>
</tr>
<tr>
<td>Low</td>
<td>36%</td>
<td>7%</td>
<td>0.46</td>
</tr>
</tbody>
</table>

Tunable detection algorithm allows user to tailor results to personal needs.
Modeling Approach – Fault Diagnosis

Decomposition-based approach integrates optimization and dictionary learning to decompose signals

Functional PCA, Xgboost, and Random Forest were used to diagnose the fault type.

Methodology enables accurate classification of DC-side faults typically only detected through aerial scans.
OTHER APPLICATIONS OF ML/AI IN POWER GENERATION
RESEARCH VS INDUSTRIAL APPLICATIONS

RESEARCH
- Time spent on testing
- High Data Quality
- Lots of backtesting
- Ability to try many methods

INDUSTRY
- Off the shelf
- Robust
- Data is flawed
- Time spent diagnosing, not maintaining AI

SETO funding and industry partners have allowed us to focus on solutions that work for industry.
<table>
<thead>
<tr>
<th>Which Tool?</th>
<th>Performance Benchmarking</th>
<th>Anomaly Detection</th>
</tr>
</thead>
<tbody>
<tr>
<td>What are the problem features?</td>
<td>Recurrent Neural Networks • Autoencoders • Bayesian Regression • Probabilistic Forecasting</td>
<td>Dynamic Linear Models • Auto–associative Neural Networks • Bayesian Hypothesis Testing</td>
</tr>
<tr>
<td>Time dependent?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Must run in real-time?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Must update automatically?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fault Detection and Diagnosis</th>
<th>Forecasting</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Bayesian Networks • Xgboost • Random Forest</td>
</tr>
<tr>
<td></td>
<td>• Artificial Neural Networks • Uncertainty Quantification • Uncertainty Propagation</td>
</tr>
</tbody>
</table>
ANOMALY DETECTION

PROBLEM
Identify faulted equipment

TECHNIQUE
Dynamic Linear Model
Digital Twin

CHALLENGES
Noisy data
Incomplete data
Every site has different data

LESSONS
Automatic trending
Automatic thresholding
Flag items to look at

Model is constantly evaluating new data, checking for significant changes

After an outage, the model quickly recognizes that new data belongs to a new trend and automatically retrains itself
OUTAGE PERFORMANCE BENCHMARKING

**PROBLEM**
Quantify outage impact
Benchmark future outages

**TECHNIQUE**
Recurrent Neural Network
Autoencoder
Digital Twin

**CHALLENGES**
Ambient conditions are cyclic
Some changes are expected

**LESSONS**
Flag items to look at
Automatic training
VIRTUAL EXPERT

PROBLEM
Identify faulted hardware
Knowledge retention

TECHNIQUE
Bayesian Belief Network

CHALLENGES
Non-centralized failure logs
Sparse data

LESSONS
Observations can be overlooked
SME’s implicitly know relationships
Engineers hate data entry

Intelligent prompts for additional observations refines diagnoses.
PERFORMANCE FORECASTING (OPTORA)

Prediction uncertainty considers model error and weather forecast uncertainty

Coupling with demand forecast allows optimization of unit dispatch

**PROBLEM**
Accurately Forecasting Production
Optimizing Unit Dispatch

**TECHNIQUE**
Artificial Neural Network
Digital Twin

**CHALLENGES**
Inconstant weather forecasts
Varying operational costs
Varying unit performance

**LESSONS**
Need “Invisible” AI
Need to correct input data
CLOSING THOUGHTS

▪ Ignoring physics in AI/ML?
  • You're missing its true potential

▪ Complex models and architectures designs in solutions?
  • You're setting up for failure

▪ If your AI can't handle messy data, is it even AI?

▪ Static models are just trending
  • Without automated trend analysis and retraining there’s no learning

▪ Without domain expertise, AI/ML is just a novice playing expert