Degradation Science and Lifetime Extension for Complex Materials Systems:

Applying Spatiotemporal-graph Learning at Scale with Time-series, Spectral and Image Datasets

Roger H. French

SDLE Research Center
Materials Science & Engineering Dept.,
Case Western Reserve University, Cleveland OH 44106 USA

A selection of Lifetime Extension Articles
roger.french@case.edu
http://dmseg5.case.edu/people/faculty.php?id=rf131

Funded by Ohio Third Frontier under award Tech 11-060 & 12-004
SDLE Research Center: Acknowledgements

CWRU Faculty
- Roger French, Laura Bruckman, Jeffrey Yarus, Jennifer Braid, Mehmet Koyutürk, Yinghui Wu

Post-doctoral Research Associates
- Two openings: PV Degradation, Statistics & Data Science

Graduate Students
- Menghong Wang, Donghui Li, Arash Khalilnejad, Ahmad Karimi, Alan Curran, JiQi Liu, Arafath Nihar
- Raymond Wieser, Kunal Rath, Sameera Nalin Venkat, Tian Wang

Undergraduates
- Ben Pierce, Tyler Burleyson, Carolina Whitaker, Minh Luu, Asher Baer, Daniel Arnholt, Hein Aung
- Deepa Bhuvanagiri, Jakob Wegmueller, Leonie Kemper, Medha Nayak, Yageen Hassan, Cora Lutes
- Alejandra Ramos, Patrick Upton, Reeve Woodward, Rounak Chawla, Will Oltjen

High School: Brandon Jackson

SDLE Staff: Jonathan Steirer Rich Tomazin
Lifetime & Degradation Science: Data Science & Analytics

Long-lived, Real-world Systems
- Focus on Degradation Rates, not Failures
  - For Lifetime Performance
- Cross-correlate Mechanism & Rates
- Develop Mechanistic Network Models

Peta-byte with Peta-flop Computing
- Distributed Computing: Hadoop/Hbase
- In-place Distributed R Analytics
- Automated Acquisition, ETL & Data Ingestion
- Built DevOps Infrastructure, In HPC Env.

Develop Population-based Studies
- Engineering Epidemiology of
  - Real-world Power Plants
  - Accelerated Laboratory Exposures

Data Science Approach Using
- Data-driven, Unbiased Analysis
- With Mechanistic Chemistry & Physics
- Inferential Statistics
- Statistical & Machine Learning

Develop Domain Science Guided, Network Models
- Energy-CRADLE: Integrated Real-world and Lab-based studies
- sgSEM: Integrated Physical & Statistical sub-models
- Graph-based Network Models: Across Populations and Through Time
Stressor & Response Framework: Deconstructionist Approach

Stressor & Response Framework
- Materials/Components/Systems
- Track degradation -> time
- Parameter space of stress conditions

Cross-correlation of Stressors
- Of Lab & Real World Exposures
- Response is comparable

Informs Design and Mitigation
- Failures & success
  - Provide insights
- Informs testing and standards

Develop System Technology Models

Cross Correlate Stressors & Responses

Also can be used to make new qualification standards
Longer Lasting Components: Development & Qualification Schema

Learnings from Degradation Pathways
- Define multiple areas to mitigate
- So as to extend lifetime
- In newly fabricated components

Qualification Testing
- Typically Pass/Fail
- No scientific learnings
- Adapt Study Protocol for Data Analysis

Accelerated & HAST Tests
- Can mislead design process
- May not be relevant to real-world lifetime

Field Failures Inform Future Designs
- And identify critical degradation modes

No scientific insights from success of qualification testing

Design → Accelerated Testing → Development → Product Qualification → Manufacture

Failures not seen in field may keep new technologies out of the market

By the time issues arise, usually too late
Lifetime & Degradation Science Framework and Thrusts

Exposures & Evaluations of Fielded Materials & Components

Forensic Studies on Field-Retrieved Components

PV Module Cracked Backsheets

(a) China, 4 years

Italy, 5 years

Lab-based Accelerated Studies

CNN Classification of Accelerated Exposed Backsheet Cracks

Network models: <S|M|R> Mechanistic Degradation Pathway

PET: Surface Erosion

PET Degradation
Under accelerated exposures with water spray (top) and humidity (bottom)

Chemical degradation mechanisms identified along pathways
Surface erosion with spray.

Deep Learning with spatiotemporal-Graph Modeling

Deep Learning with spatiotemporal-Graph Modeling

Improved Power Prediction by st-Graph Neural Network Model

Figure 8: Six examples of crack inspection task performed on the test images using the trained Model O. The different colors in the left and right column images indicate different crack classes shown in the color bar.


Outline

Lifetime Performance & Degradation of c-Si Photovoltaic Modules

Distributed & High Performance Computing
• AI/ML for Energy Science
• Spatiotemporal Modeling: Real-world Field Surveys
• Lab-based Exposures & Evaluations: Accelerated Testing

Image Processing & Machine Learning
• CNN Machine Learning to Quantify Corrosion using Electroluminescent Imaging
• Predictive and Inferential Models of Photovoltaic Lifetime Performance

Time-series Analysis of Photovoltaic Power Plant Systems
• IEA-PVPS: Assessment of Performance Loss Rate
• Remote Diagnostics of PV Plants: PLR, Degradation Mechanisms, Shading
• St-Graph Learning: High Accuracy Power Forecasting & PLR Prediction

Degradation Pathway Network Graph Modeling
• \(<S|M|R>\) network Structural Equation Modeling (netSEM)
• Towards 50 Year PV Modules: Module Architecture, Encapsulants, Backsheets
Lifetime Performance & Degradation of Crystalline Silicon Photovoltaic Modules

Corrosion
Encapsulant yellowing
Delamination
Polyamide backsheet cracking
Structure of c-Si Photovoltaic module

Polymer components in solar module

- **Front & back encapsulant**
  - Front: UV transparent
  - Back: UV cutoff
  - High optical transmittance
  - Low modulus, high adhesion

- **Backsheet**
  - Opaque, multilayer laminate
  - Electrical insulation
  - Prevent moisture ingress

Determine the lifetime of the module

Polymer encapsulant “aggravates or creates a number of failure mechanisms”

- R. G. Ross, “PV reliability development lessons from JPL’s Flat Plate solar array project”

---

DOE-SETO PV Reliability Research Projects @ SDLE

Start 2015
1. Module Level Exposure and Evaluation Test for Real-world and Lab-based PV Modules: Common Data and Analytics for Quantitative Cross-correlation and Validation
2. Backsheets: Correlation of Long-Term Field Reliability with Accelerated Laboratory Testing

Start 2018

Start 2019
4. Towards 50 Year Lifetime PV Modules: Double Glass (DB) vs. Glass/Backsheet (GB)
5. A Data-driven Approach to Real-world Degradation of Backsheets

Start 2021
6. Robust PV Performance Loss Rate Determination: Using Spatiotemporal Graph Neural Network Models in a Reliable System-Topology-Aware Learning Framework
7. Gaining Fundamental Understanding of Critical Failure Modes and Degradation Mechanisms in Fielded Photovoltaic Modules via Multiscale Characterization
Distributed and High Performance Computing: AI/ML for Energy Science

Ahmad Karimi, Arash Khalilnejad, JiQi Liu, Philip Hwang

Real-world Data Source: CWRU SDLE Global SunFarm Network

SDLE PV Data Covers ~3.4 GW
- PV Fleet Geospatial Distribution
- PV System Field Surveys for Backsheet Degradation

Encompasses ~1 % of Global PV Power Production
- 787 PV Project Sites
- 5638 PV Systems (Inv. & Modules)
- 60 PV Module Brands/Models
- 38 PV Inverter Brands/Models
- Across 13 Köppen-Geiger Climatic Zones
- Single Modules to 265 MW plants
- Going Back Up To 15 years

Epidemiological PV Populations
- Of Time-series data streams
- Real-world power production
- Real World Exposure Conditions
- Operating Over Real Time-scales

18 Different Companies Have Signed On
- To our Data Use Agreement
Measuring Degradation Mechanisms: Time-Series Exposure & Evaluation

Accelerated & Real-World Exposures
- Lab-based exposures to mimic
- Real-world degradation

Multiple Datatypes
- “Point” values
- Spectra
- Images
- Hyper-spectral Images

Statistically Informed Study
- Large Volume of Samples
- Diverse Exposures
  - Real-world & Lab Base
  - Accelerated & Real-time
- Many Evaluations
  - Mechanistic & Performance
CRADLE v2.2 Architecture: Petabyte and Petaflop Computing

Hadoop Distributed Computing
With Hbase & Spark
Using R & Python For Analytics
In-place Analytics
Write-back All Results Into Hbase

CRADLE – DISTRIBUTED / HIGH PERFORMANCE COMPUTING

Data Science & Machine Learning Platform

rcradletools package
pycradletools package
Apache Spark 2.2

HBase
weather meta energy result
power material

HDFS
1 Edge Node, 2 Master Node, 12 Data Node,
1 Thrift Server, 180 Cores, 96 TB Disk, 2 TB RAM

Hadoop & Spark Cluster

Secure server
Proprietary data

Global Commercial Sun Farms
SDLE Research Lab
CWRU Research sites
Research Collaborators
Academic partners
Data Sources

4,000 Compute Cores
108,000 GPU Cores

March 29, 2021, VuGraph 18
**FAIRification of Datasets and Models, Enables AI learning**

**Making Datasets & Models FAIR**
- By “FAIRification”

**Enables Models to find Data**
- And Data to find Models

**So that they can advance**
- Without human intervention

**This is an aspect of the Semantic Web**
- And Resource Description Framework
- Hbase triples are an example of RDF

**We just received a DOE SETO AI award**
- For st-GNN, that involves FAIRification

---

**Enabling this in Hadoop/Hbase Environment**
- Can enable automation of analysis

---

NoSQL DB Abstraction of Hadoop/Hbase

Combines Lab data (Spectra, Images etc.) With Time-series Data (PV Power Plant Data)

High Performance PV Data Analytics: Petabyte Data Warehouse In A Petaflop HPC Environment

- In-place Analytics: Distributed R-analytics in Hadoop/HDFS
- In-memory Data Extraction: To Separate HPC Compute Nodes

A non-relational data warehouse for the analysis of field and laboratory data from multiple heterogeneous photovoltaic test sites

IEEE JPV

Yang Hu, Member, IEEE, Venkat Yashwanth Gunapati, Pei Zhao, Devin Gordon, Nicholas R. Wheeler, Mohammad A. Hossain, Member, IEEE, Timothy J. Peshek, Member, IEEE, Laura S. Bruckman, Guo-Qiang Zhang, Member, IEEE, and Roger H. French, Member, IEEE
Outline

Lifetime Performance & Degradation of c-Si Photovoltaic Modules

Distributed & High Performance Computing
  • AI/ML for Energy Science
  • Spatiotemporal Modeling: Real-world Field Surveys
  • Lab-based Exposures & Evaluations: Accelerated Testing

Image Processing & Machine Learning
  • CNN Machine Learning to Quantify Corrosion using Electroluminescent Imaging
  • Predictive and Inferential Models of Photovoltaic Lifetime Performance

Time-series Analysis of Photovoltaic Power Plant Systems
  • IEA-PVPS: Assessment of Performance Loss Rate
  • Remote Diagnostics of PV Plants: PLR, Degradation Mechanisms, Shading
  • St-Graph Learning: High Accuracy Power Forecasting & PLR Prediction

Degradation Pathway Network Graph Modeling
  • $<S|M|R>$ network Structural Equation Modeling (netSEM)
  • Towards 50 Year PV Modules: Module Architecture, Encapsulants, Backsheets
CNN Machine Learning to Quantify Corrosion using Electroluminescent Imaging

Automated Pipeline for Photovoltaic Module Electroluminescence Image Processing and Degradation Feature Classification

Ahmad Maroof Karimi, Justin S. Fada, Mohammad Akram Hossain, Shuying Yang, Timothy J. Peshek, Jennifer L. Braid, Member, IEEE, and Roger H. French, Member, IEEE

Generalized and Mechanistic PV Module Performance Prediction from Computer Vision and Machine Learning on Electroluminescence Images

Ahmad Maroof Karimi, Justin S. Fada, Nicholas A. Parrilla, Benjamin G. Pierce, Mehmet Koyutürk, Roger H. French, Member, IEEE, Jennifer L. Braid, Member, IEEE

*SDLE Research Center, Case Western Reserve University, 10900 Euclid Ave., Cleveland, Ohio 44106, USA
1 Department of Computer and Data Sciences, Case Western Reserve University
2 Department of Materials Science and Engineering, Case Western Reserve University

“SunEdison” Test to failure EL & I-V Dataset

5 Brands, 30 commercial modules, 60 PV cells each
- A, B, C, D, and E
- 3 samples/replicates of each

2 Accelerated exposures
- Standard IEC 61215 damp-heat (85 oC/ 85 % RH)
- Thermal cycle test conditions

I-V and EL captured as step wise
For damp heat, 15 modules
- 500 hours intervals step from 0-3000 hours,
- 4 modules continued to 4200 hours with 300 hours steps

For thermal cycling, 15 modules
- 200 cycle intervals from 0-600 cycles
- 4 modules continued to 1000 cycles with 100 cycles step

I-V curves were obtained with a Spire
- 4600 SLP flash solar simulator.

EL images
- Sensovation coolSamBa HR-830 8.3 megapixel camera.

Electroluminescence Image of a PV
Time-series module degradation: EL Images & I-V Measurements

5 brands of PV modules
- 3000 h Damp Heat
  - 85 °C/85% RH
- In 500 h steps

EL Images show degradation
- different modes
- and rates

And Power Output
- Different rates
Planar Indexed Module → Individual Cell Images

Cell Extraction

- Starts with planar index module
- Simple matrix slicing used to extract cells
  - Further refined image processing would result in lost information
- Results in single cell images
  - Resembles face recognition problem
Supervised PV Cell Classification by Degree of Corrosion

Cell-level images receive corrosion score
- Busbar corrosion = 0 to 4

Convolutional Neural Network
- Trained to assign corrosion score
- Based on manual classification

Module-level images scored as
- Average of cell-level corrosion

Confusion matrix for cell classification into 5 corrosion levels in order of increasing severity from 0-4
Data Assembly & Integration for Predictive and Inferential Models of Photovoltaic Lifetime Performance
## EL – I-V Feature Correlation

### Generalized EL features:
- EL median intensity ($F_{med}$)
- Fraction dark (“inactive”) pixels ($F_{FDP}$)

### Degradation-specific EL features:
- Normalized busbar width ($F_{NBBW}$)
- Module feature ratio ($F_{MFR}$)

### Correlation map of I-V and EL features
- Identify related features, such as
  - EL median and series resistance ($R_s$)
  - Series resistance and power ($P_{mp}$)

### Develop predictive models for
- Overall module performance ($P_{mp}$)
- Mechanistic degradation ($R_s$ etc.)
Generalized Power Prediction: Across Brands and Exposure Types

Power prediction models

\[ P_{mp-n} = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3 \]

<table>
<thead>
<tr>
<th>Model</th>
<th>(X)</th>
<th>(\beta_0)</th>
<th>(\beta_1)</th>
<th>(\beta_2)</th>
<th>(\beta_3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td>(1 - F_{med-n})</td>
<td>0.936</td>
<td>-1.396</td>
<td>-0.756</td>
<td>-0.039</td>
</tr>
<tr>
<td>Inactive Region</td>
<td>(F_{FDP-n})</td>
<td>0.936</td>
<td>-1.502</td>
<td>-0.478</td>
<td>-0.128</td>
</tr>
<tr>
<td>NBBW</td>
<td>(F_{NBBW-n})</td>
<td>0.935</td>
<td>-0.576</td>
<td>-0.064</td>
<td>-0.005</td>
</tr>
<tr>
<td>MFR</td>
<td>(F_{MFR})</td>
<td>0.932</td>
<td>-0.650</td>
<td>-0.204</td>
<td>0.014</td>
</tr>
</tbody>
</table>

Metrics to measure the performance of the models

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Mean</th>
<th>Standard dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjusted (R^2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>0.88</td>
<td>0.025</td>
</tr>
<tr>
<td>Inactive Region</td>
<td>0.87</td>
<td>0.016</td>
</tr>
<tr>
<td>NBBW</td>
<td>0.70</td>
<td>0.03</td>
</tr>
<tr>
<td>MFR</td>
<td>0.81</td>
<td>0.017</td>
</tr>
<tr>
<td>RMSE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>11.87</td>
<td>5.57</td>
</tr>
<tr>
<td>Inactive Region</td>
<td>12.44</td>
<td>4.96</td>
</tr>
<tr>
<td>NBBW</td>
<td>13.35</td>
<td>10.53</td>
</tr>
<tr>
<td>MFR</td>
<td>9.53</td>
<td>4.75</td>
</tr>
<tr>
<td>MAPE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>3.46</td>
<td>1.58</td>
</tr>
<tr>
<td>Inactive Region</td>
<td>3.94</td>
<td>2.08</td>
</tr>
<tr>
<td>NBBW</td>
<td>4.34</td>
<td>3.08</td>
</tr>
<tr>
<td>MFR</td>
<td>2.83</td>
<td>1.45</td>
</tr>
</tbody>
</table>
Outline

Lifetime Performance & Degradation of c-Si Photovoltaic Modules

Distributed & High Performance Computing

- AI/ML for Energy Science
- Spatiotemporal Modeling: Real-world Field Surveys
- Lab-based Exposures & Evaluations: Accelerated Testing

Image Processing & Machine Learning

- CNN Machine Learning to Quantify Corrosion using Electroluminescent Imaging
- Predictive and Inferential Models of Photovoltaic Lifetime Performance

Time-series Analysis of Photovoltaic Power Plant Systems

- IEA-PVPS: Assessment of Performance Loss Rate
- Remote Diagnostics of PV Plants: PLR, Degradation Mechanisms, Shading
- St-Graph Learning: High Accuracy Power Forecasting & PLR Prediction

Degradation Pathway Network Graph Modeling

- <S|M|R> network Structural Equation Modeling (netSEM)
- Towards 50 Year PV Modules: Module Architecture, Encapsulants, Backsheets
Times Series Analysis of Photovoltaic Power Plant Systems

Alan Curran
Köppen-Geiger Climatic Zones & PV System Fleets

**Type = Humidity/Wetness**  
**Subtype = Temp./Temperature Range**

---

### Köppen-Geiger Climatic Zones

#### A-Tropical
- **T_{min} ≥ +18°C**
- **Type = Rainforest**
- **Criterion:** \( P_{min} ≥ 60 \text{ mm} \)

#### B-Arid
- **\( P_{ann} < 10 \text{ Pa} \)**
- **Type:**
  - s: Desert \( P_{max} ≤ 5 \text{ Pth} \)
  - h: Hot \( T_{ann} ≥ +18°C \)
  - k: Cold \( T_{ann} < +18°C \)

#### C-Temperate
- **\(-3°C < T_{min} < +18°C\)**
- **Type:**
  - s: Dry Summer \( P_{max} > 10 \text{ Pmin} \)
  - w: Dry Winter \( P_{max} < P_{min} \)
  - f: Without dry season Not Cs or Cw
  - a: Hot Summer \( T_{max} ≥ +22°C, 4 T_{min} ≥ +10°C \)
  - b: Warm Summer \( T_{max} < +22°C, 4 T_{min} ≥ +10°C, T_{max} > +38°C \)
  - c: Cold Summer \( T_{max} < +22°C, 4 T_{min} ≥ +10°C, T_{max} < +38°C \)
  - s: Dry Summer \( P_{min} < P_{min}, P_{max} < 10 \text{ Pa} \)
  - w: Dry Winter \( P_{max} > 10 \text{ Pmin} \)
  - f: Without dry season Not Cs or Cw

#### D-Cold (Continental)
- **\( T_{min} ≤ -3°C \)**
- **Type:**
  - f: Without dry season Not Ds or Dw
  - a: Hot Summer \( T_{max} ≥ +22°C \)
  - b: Warm Summer \( T_{max} < +22°C, 4 T_{min} ≥ +10°C \)
  - c: Cold Summer \( T_{max} < +22°C, 4 T_{min} ≥ +10°C, T_{max} > +38°C \)
  - d: Very cold Winter \( T_{max} < +22°C, 4 T_{min} ≥ +10°C, T_{max} ≤ +38°C \)

#### E-Polar
- **\( T_{max} < +10°C \)**
- **Type:**
  - T: Tundra \( T_{max} ≤ 0°C \)
  - F: Frost/Ice cap \( T_{max} < 0°C \)

---

**Generated based on precipitation and temperature**
- Begun in 1884, further classified 1954
- Consistent and comprehensive climatic zones

**29 total K-G Climatic Zones defined**
- Understand environmental stressors

**Use Geospatial tools for geospatiotemporal datasets**
Assessment of Performance Loss Rate of PV Power Systems

From IEA-PVPS Task 13 on Performance & Reliability
SubTask 2.5: PLR Determination
Report In Press
Performance Loss Rate (PLR) of PV Systems

Performance loss rates (%/a)
- Long term prediction of power/energy production of systems
- Crucial in Levelized Cost Of Electricity calculations

Reliable modules are key to PV expansions into the grid
- Current systems monitored to validate warranties
- PLR yield information on power loss mechanisms
- Identify needed performance improvements

PLR calculations have distinct steps
- Filtering out unwanted data
  - Nightime, cloud cover, extreme values etc.
- Correcting for weather influences
  - Irradiance, temperature, wind, etc.
- Identifying outliers
- PLR modeled using
  - Year-on-year, Timeseries Decomposition, Regression

0. Exploratory data analysis for data quality & grading

1. Input data cleaning & filtering

2. Performance metric selection, corrections & data aggregation

3. Timeseries feature corrections

4. Statistical modeling of PLR


PVplr Package Capabilities

**Data cleaning**
- Irradiance and temperature filters
- Inverter clipping
- Shading influences

**PLR evaluation**
- Linear regression
- Year-on-year regression
- Non-linear PLR

**Dataset Grading**
- Missing, Gaps, Anomalies

**Weather correction models**
- PVUSA
- 6K
- XbX
- XbX + UTC
  - Most similar to temp.corrected perf. ratio

**PLR uncertainty**
- Model coefficient variance
- Bootstrap uncertainty

**Time-series corrections**
- Outlier detection
- STL decomposition

\[ P_{\text{pred.}} = \beta_0 + \beta_1 G + \beta_2 T + \epsilon \]

\[ P = G' = G_{POA}/G_{STC} \]

\[ T' = T_{mod} - T_{STC} \]

\[ P = G'(P_{NP} + k_1 \ln(G') + k_2 \ln(G')^2 + k_3 T' + k_4 T' \ln(G') + k_5 T' \ln(G')^2 + k_6 T'^2) \]

\[ P_{\text{cor}} = \frac{P_{\text{obs}}}{1 + \gamma_T(T_{\text{obs}} - T_{\text{rep}})(G_{\text{obs}} / G_{\text{rep}})} \]

\[ P_{\text{cor}} = \beta_0 + \beta_1 G + \epsilon \]

\[ P = G_{POA}(\beta_0 + \beta_1 G_{POA} + \beta_2 T_{amb} + \beta_3 W) \]
PLR Prediction: Interlab Benchmarking by “Voting” or Preference Aggregation

This is 1 PV System at EURAC in Italy
- We studied 19 real PV systems

Use 27 Filter/Metric/Model combinations for PLR
- Rank order all the PLR Results
- Determine difference from mean PLR

PLR determined by an “Ensemble” model of methods
- Standard Error determined by # of methods

Comparison of two mean PLRs
- Use 83.4% Cis for 5% significance level
Remote Diagnostics of PV Plants: PLR, Degradation Mechanisms, Shading

Determined from 8 Years of $I-V, P_{mp}$ Time-series Datastreams

Jiqi Liu, Menghong Wang, Alan J. Curran, Ahmad Maroof Karimi, Weiheng Huang, Erdmut Schnabel, Michael Köhl, Jennifer L. Braid, Roger H. French

SDLE Research Center, Case Western Reserve University, Cleveland OH
Fraunhofer-ISE, Freiburg Germany

http://sdle.case.edu
“Fraunhofer-ISE” Real-world Exposures with $I-V$, $P_{mp}$ Time-series Datastreams

With the Fraunhofer ISE data set,

- $I-V$ curves are taken every 10 minutes
  - With $P_{mp}$ measured each minute
- 8 modules total
- 3 locations ($BSh$, $BWh$, $ET$ climate zones)
- Glass-Glass samples F
- Glass-Backsheet samples G
- 3-6 years of data available

<table>
<thead>
<tr>
<th>Location</th>
<th>Sample Type</th>
<th>System Age (Months)</th>
<th>Climate Zone</th>
</tr>
</thead>
<tbody>
<tr>
<td>GC</td>
<td>F</td>
<td>57</td>
<td>BWh</td>
</tr>
<tr>
<td>GC</td>
<td>G</td>
<td>57</td>
<td>BWh</td>
</tr>
<tr>
<td>GC</td>
<td>F</td>
<td>57</td>
<td>BWh</td>
</tr>
<tr>
<td>UFS</td>
<td>F</td>
<td>32</td>
<td>ET</td>
</tr>
<tr>
<td>UFS</td>
<td>G</td>
<td>57</td>
<td>ET</td>
</tr>
<tr>
<td>NEG</td>
<td>F</td>
<td>61</td>
<td>BSh</td>
</tr>
<tr>
<td>NEG</td>
<td>G</td>
<td>36</td>
<td>BSh</td>
</tr>
<tr>
<td>NEG</td>
<td>G</td>
<td>61</td>
<td>BSh</td>
</tr>
</tbody>
</table>

The Zugspitze, Germany (alpine climate)

Gran Canaria, Spain (maritime climate)

Negev Desert, Israel (arid climate)
PV Modules at Real World Sites

Two Module Architectures:
- Glass/Backsheet (GB)
- Double Glass (DG)

Gran Canarias

Zugspitze

Negev
Data-driven I-V Feature Extraction Algorithm (ddiv)

**ddiv**\textsuperscript{[1]} analysis of all time-series I-V curves
- 3.2 million curves from 8 modules

To extract both I-V features and I-V steps
- “Normal” I-V curve single step I-V curve
- I-V features include
  - $P_{mp}$, $I_{sc}$, $I_{mp}$, $V_{oc}$, $V_{mp}$, $R_s$, $R_{sh}$, FF

For multi-step (MS) I-V curves
- Extract I-V features for each stepped curve.
- And save the step’s ‘Cutoff’ voltage (V)

<table>
<thead>
<tr>
<th>Steps #</th>
<th>$I_{sc}$</th>
<th>$R_{sh}$</th>
<th>$V_{oc}$</th>
<th>$R_s$</th>
<th>$P_{mp}$</th>
<th>$I_{mp}$</th>
<th>$V_{mp}$</th>
<th>FF</th>
<th>Cutoff</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.11</td>
<td>852.74</td>
<td>47.08</td>
<td>1.24</td>
<td>111.42</td>
<td>2.85</td>
<td>39.06</td>
<td>76.13</td>
<td>NA</td>
</tr>
<tr>
<td>3</td>
<td>2.92</td>
<td>477.96</td>
<td>43.92</td>
<td>10.39</td>
<td>50.18</td>
<td>2.67</td>
<td>18.77</td>
<td>39.2</td>
<td>21.11</td>
</tr>
<tr>
<td></td>
<td>3.64</td>
<td>14.66</td>
<td>81.76</td>
<td>27.06</td>
<td>55.82</td>
<td>1.98</td>
<td>28.21</td>
<td>18.78</td>
<td>28.27</td>
</tr>
<tr>
<td></td>
<td>2.87</td>
<td>31.70</td>
<td>46.90</td>
<td>1.41</td>
<td>82.13</td>
<td>1.96</td>
<td>41.90</td>
<td>61.08</td>
<td>NA</td>
</tr>
</tbody>
</table>

Power Loss Modes: Current Mismatch, Recombination, Series Resistance, Uniform Current

Power loss factors for module 2 (BWh: DG)
- obtained from outdoor Isc -Voc analysis

Rate of change of each power loss mode
- from outdoor $I_{sc}$ -$V_{oc}$ analysis

Climate zone dependency: Common mechanistic loss factors in a KG-CZ
- BWh, Gran Canarias: Current Mismatch
- BSh, Negev: Cell shunting
- ET, Zugspitze: Series resistance increase
Shading Profile

Group data into each year
- Years should have more than 100 days I-V
- Calculate Multiple Steps (MS) for each time point
- Spline model + find local peak

\[ M_{St} = \frac{N_{ms,t}}{N_{total,t}} \]

<table>
<thead>
<tr>
<th>Time in a Day</th>
<th>MS(%)</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:10:00</td>
<td>35.1</td>
<td>2011</td>
</tr>
<tr>
<td>15:30:00</td>
<td>65.1</td>
<td>2011</td>
</tr>
<tr>
<td>07:10:00</td>
<td>46.9</td>
<td>2012</td>
</tr>
<tr>
<td>14:46:00</td>
<td>57.1</td>
<td>2012</td>
</tr>
<tr>
<td>08:10:00</td>
<td>55.7</td>
<td>2013</td>
</tr>
<tr>
<td>15:20:00</td>
<td>61.9</td>
<td>2013</td>
</tr>
<tr>
<td>08:40:00</td>
<td>58.9</td>
<td>2014</td>
</tr>
<tr>
<td>15:45:00</td>
<td>78.0</td>
<td>2014</td>
</tr>
<tr>
<td>07:55:00</td>
<td>32.1</td>
<td>2015</td>
</tr>
<tr>
<td>16:10:00</td>
<td>52.1</td>
<td>2015</td>
</tr>
<tr>
<td>07:50:00</td>
<td>40.0</td>
<td>2016</td>
</tr>
<tr>
<td>15:15:00</td>
<td>64.0</td>
<td>2016</td>
</tr>
</tbody>
</table>
st-Graph Learning:
High Accuracy Power Forecasting
& PLR Prediction

Ahmad M. Karimi,
Yinhui Wu,
Mehmet Koyuturk,
Roger H. French
**Spatiotemporal Graph Neural Network (st-GNN)**

**Interest**
- Information from neighboring nodes undergoing similar exposure

**Sequence of**
- Graph convolution layer
- Temporal convolutional layer
  - 1-D convolution

**Coherence**
- Spatial dependencies
- Temporal dependencies
Dataset

- SS1 + SS2 dataset: 316 power plants
- 2 years of data (730 days)
- 5 minutes interval
- 288 points makes up a day
- 210,240 points for a system
- Data partition
  - 690 days training, 20 days validation, 20 days testing
- Input Features for modeling
  - Power timeseries\(P_{mp}\)

Power forecasting models

- Power \(P_{mp}\)
PV Network Representation

Calculate distance between two nodes

\[ d_{\text{lon}} = \text{lon}_2 - \text{lon}_1, \quad d_{\text{lat}} = \text{lat}_2 - \text{lat}_1 \]
\[ a = (\sin(d_{\text{lat}}/2))^2 + \cos(\text{lat}_1) \times \cos(\text{lat}_2) \times (\sin(d_{\text{lon}}/2))^2 \]
\[ d = 2 \times R \times \arcsin(\sqrt{a}) \]

where, \( R \) is radius of the earth

- Equation to convert element of distance matrix to weight matrix
- \( \varepsilon_c = 0.5 \)

\[ w_{ij} = \begin{cases} 
\exp(-\frac{d_{ij}^2}{\sigma^2}), & \text{if } i \neq j \text{ and } \exp(-\frac{d_{ij}^2}{\sigma^2}) \geq \epsilon \\
0, & \text{otherwise.} 
\end{cases} \]

\( d_{ij} \) = distance between node i and node j
\( \sigma \) is normalizing constant
\( \epsilon \) is constant which control graph sparsity

H. B. Goodwin, The haversine in nautical astronomy, Naval Institute Proceedings, vol. 36, no. 3 (1910)
Spatiotemporal Graph Neural Network (st-GNN) Representation

Two spatio-temporal block
- Two temporal convolution layer
- One spatial convolution layer

Output Layer Block
- Two temporal
- Fully connected layer

\[ H = 24 \text{ number of previous time points, } N = 316 \text{ PV Systems} \]

\[ H = 24 \text{ number of time lag points} \]
\[ N = 316 \text{ PV Systems} \]
Single System PV Power Forecast

PV power forecast one day

- Fluctuation in the curve due to cloud cover
- Forecast for spatiotemporal convolution ($\epsilon_c = 0.25$)
- Forecast for temporal (1-D) convolution ($\epsilon_c = 1.0$)
- Spatiotemporal curve follows observed values trend closely
PV Power Forecast Error

Mean absolute percentage error
- Function of $\varepsilon_c$

Find optimal $\varepsilon_c$

Violin plots
- Each violin plot represents 316 systems
- Range of box plots show 1st and 3rd quantile

MAPE values for the PV power forecast for 316 systems.

![Diagram showing violin plots with threshold cutoff $\varepsilon_c$.](image)
Power Forecasting ($P$)

Forecasting duration (in minutes)
- 15
- 60
- 90
- 120

Threshold ($\epsilon_c$)
- 0 (All nodes connected)
- 0.125
- 0.25
- 0.375 (Optimal Value)
- 0.5
- 1.0 (Temporal convolution)

Key points:
- Optimal value is consistent across all future prediction time
- Optimal value occurs at $\epsilon_c = 0.375$
Outline

Lifetime Performance & Degradation of c-Si Photovoltaic Modules

Distributed & High Performance Computing
- AI/ML for Energy Science
- Spatiotemporal Modeling: Real-world Field Surveys
- Lab-based Exposures & Evaluations: Accelerated Testing

Image Processing & Machine Learning
- CNN Machine Learning to Quantify Corrosion using Electroluminescent Imaging
- Predictive and Inferential Models of Photovoltaic Lifetime Performance

Time-series Analysis of Photovoltaic Power Plant Systems
- IEA-PVPS: Assessment of Performance Loss Rate
- Remote Diagnostics of PV Plants: PLR, Degradation Mechanisms, Shading
- St-Graph Learning: High Accuracy Power Forecasting & PLR Prediction

Degradation Pathway Network Graph Modeling
- <S|M|R> network Structural Equation Modeling (netSEM)
- Towards 50 Year PV Modules: Module Architecture, Encapsulants, Backsheets
Network Structural Equation Modeling:
Identifying Mechanistic Degradation Pathways

<table>
<thead>
<tr>
<th>Stressor</th>
<th>Mechanism / Mode</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irrad. (S1)</td>
<td>EVA Hydrolysis (M4)</td>
<td>Pmax (R8)</td>
</tr>
<tr>
<td>Temp. (S2)</td>
<td>Metalization Corrosion (M6)</td>
<td>FF (R9)</td>
</tr>
<tr>
<td>RH (S3)</td>
<td>PET Hydrolysis (M5)</td>
<td>Wet Ins. Resistance (R10)</td>
</tr>
</tbody>
</table>

\[ \beta_{4,3}, \beta_{4,2}, \beta_{5,2}, \beta_{5,3}, \beta_{6,4}, \beta_{7,5}, \beta_{8,6}, \beta_{9,6}, \beta_{10,7} \]
netSEM: Network Structural Equation Modeling

- Network structural equation modeling
- Degradation pathways with mechanisms
- R package developed by SDLE Center\(^1\), v0.6.0

Graph modeling also being considered

**Stressor:** time in dy (decimal years)

**Mechanistic variables:** (normalized to 0-1 range)
- \(nR_{s,IV}\): Corrosion
- \(nI_{sc,IV}\): Optical transmission loss
- \(nV_{mp,PIV}\): Recombination and shunting

**Response:** \(nP_{mp,IV}\): power output

\(<S|R>\): \(nP_{mp,IV}\) changing with time

\(<S|M>\): mechanistic variables changing with time

\(<S|M|R>\) Paths: From I-V, EL/PL, & \(Suns-V_{oc}\) features
- That track degradation mechanisms
- Have functional forms


Stressor

Variable

Tracking Variable 1

Degr. Mech. 1

Tracking Variable 2

Response Variable

Degr. Mech. 2

Tracking Variable 3

Degr. Mech. 3

Mechanistic variables:
- \(nRs,IV\): Corrosion
- \(nI_{sc,IV}\): Optical transmission loss
- \(nV_{mp,PIV}\): Recombination and shunting

Response:
- \(nP_{mp,IV}\): power output
Towards 50 Year PV Modules:
Module Architecture, Encapsulants, Backsheets:
Step-wise Study to Identify the Mechanistic Degradation Pathways

Sameera Nalin Venkat
JiQi Liu
Alan Curran
Laura Bruckman
Roger French
Stepwise Study Protocol: 8 MM Variants, 2 Accel. Exposures, 3 Evaluations, Network Models

Fabrication of 8 Variants for 2 Exposures = 32 Minimodules (MMs)

4 x monofacial multicrystalline PERC

Backsheet or glass
UV-cutoff encapsulant
Solar cells
Transparent encapsulant
Glass
Laminator platten

Brand
Architecture

CSI
CWR
U
GB
DG
Encapsulant

EVA
POE

mDH: 80°C+85% RH
1 step: 21 days

mDH (14 days) + FSL (7 days, 420 Wm⁻² light);
total: 21 days

2 Accelerated Exposures
(5 Exposure Steps + Baseline)

3 Stepwise Evaluations
Electrical & Imaging

I-V
Suns-\text{V}_{\text{oc}}

EL/PL

Data-Driven Modeling

mDH: 80°C+85% RH
1 step: 21 days
Model (Principle 1): GB with EVA by CWRU (mDH Exposure)

**<Stressor|Mechanism|Response> model:** pairwise relationships between variables

**Principle 1 (Markovian) for two variables:** remaining variables don’t influence them

2 minimodules (8 cells) and 6 time points: 48 observations
For statistical significance

**Pairwise relationships:** fitted using one of the 7 functional forms

\[ R_{adj}^2 \] is also calculated: insight into how well the variables are correlated

<table>
<thead>
<tr>
<th>Model</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple linear (SL)</td>
<td>[ y = \beta_0 + \beta_1 x ]</td>
</tr>
<tr>
<td>Quadratic (Quad)</td>
<td>[ y = \beta_0 + \beta_1 x + \beta_2 x^2 ]</td>
</tr>
<tr>
<td>Simple quadratic (SQuad)</td>
<td>[ y = \beta_0 + \beta_2 x^2 ]</td>
</tr>
<tr>
<td>Nonlinearizable exponential (nls)</td>
<td>[ y = \beta_0 + \beta_3 (1 \pm \exp(\beta_6 (x - \beta_7)) ]</td>
</tr>
<tr>
<td>Exponential (Exp)</td>
<td>[ y = \beta_0 + \beta_3 e^{x} ]</td>
</tr>
<tr>
<td>Logarithmic (Log)</td>
<td>[ y = \beta_0 + \beta_4 \log x ]</td>
</tr>
<tr>
<td>Change point (CP)</td>
<td>[ y = \beta_0 + \beta_1 x + \beta_2 (x - c) ]</td>
</tr>
</tbody>
</table>
Model (Principle 1): GB with EVA by CWRU (mDH Exposure)

**Stressor (S)**
- trans: optical transmission
- corros: corrosion
- recomb_shunt: recombination & shunting

**Degradation modes**
- n_Vmp_PIV
- n_Pmp_IV
- trans
- n_Lsc_IV
- n_Rs_IV

**Response (R)**
- Model: Quad adj-R-Sqr: 0.46
- Model: Quad adj-R-Sqr: 0.374
- Model: Squa adj-R-Sqr: 0.364
- Model: Quad adj-R-Sqr: 0.477
- Model: Quad adj-R-Sqr: 0.477
- Model: Quad adj-R-Sqr: 0.195
- Model: Quad adj-R-Sqr: 0.305
- Model: Quad adj-R-Sqr: 0.093
- Model: Log adj-R-Sqr: 0.963
- Model: Log adj-R-Sqr: 0.271
- Model: Log adj-R-Sqr: 0.402
- Model: Log adj-R-Sqr: 0.269
- Model: Exp adj-R-Sqr: 0.477

**Tracking variables (M)**
In order to understand degradation: we need to study not only \(<S|R>\) but also \(<S|M>, <S|R>, |M|R>\).
<S|R> Pathway: 83.4% CI Results at the End of Exposure

83.4% CIs at exposure step (es) 5

DG and GB with EVA made by CSI: better performing
- Both exposures

GB with EVA made by CWRU
- Experiences power drop
- In mDH

<S|R>: direct pathway connecting dy and \( nP_{mp, IV} \)

83.4% confidence intervals at exposure step = 5
- Colored by encapsulant and architecture

Wider confidence interval: higher uncertainty in \( nP_{mp, IV} \)
Variation of $nP_{mp,IV}$ with Exposure Time

Combinations that exhibit signs of stability (blue)

- DG by CSI in both the exposures
- DG by CWRU in mDH+FSL
- GB with EVA by CSI in mDH
- GB with POE by CSI in mDH+FSL
Variation of $nP_{mp, IV}$ with Exposure Time

Combinations that exhibit power loss (black)
- GB by CWRU
- GB with POE by CSI in mDH

Wider confidence interval: higher uncertainty in trend
Model (Principle 2)

Principle 2: non-Markovian
- dy and other variables simultaneously impact \( nP_{mp, IV} \)

Multiple regression fitting

\[
nP_{mp, IV} = 1.26 + 0.11dy^2 + 1.68\log nI_{sc, IV} - 0.2nR_{s, IV} - 0.11(nR_{s, IV} - 1)_c
\]
Degradation Science with spatiotemporal-Graph models

Graph Representation Learning:
message passing with spatiotemporal attention in graph convolution layers, incorporating non-local spatiotemporal dependencies

Material data storage: CRADLE

Metadata Field & Lab Data Manufacturing Data
Longer Lasting Components: Development & Qualification Schema

Learnings from Degradation Pathways
- Define multiple areas to mitigate
- So as to extend lifetime
- In newly fabricated components

Qualification Testing
- Typically Pass/Fail
- No scientific learnings
- Adapt Study Protocol for Data Analysis

Accelerated & HAST Tests
- Can mislead design process
- May not be relevant to real-world lifetime

Field Failures Inform Future Designs
- And identify critical degradation modes

No scientific insights from success of qualification testing

Design

Accelerated Testing

Development

Early Failure

Exhaustive Accelerated Testing (HAST)

Product Qualification

Quality of Product

Successful Commercialization

Wearout & Degradation

Failures not seen in field may keep new technologies out of the market

Safety & Performance

By the time issues arise, usually too late

Typically Pass/Fail

Accelerated & HAST Tests

Quality of Product

Successful Commercialization

Wearout & Degradation

Field Failures Inform Future Designs
- And identify critical degradation modes
We think we can help! And, propose to hold a workshop (3/10/2021) to identify & seed collaborations to benefit missions across the NNSA Complex

Virtual, unclassified workshop: Weds. March 10, 2021

What are your compelling problem sets
• You’d like help with?

What skills sets
• Could we help feed your pipelines with?

Lifetime Extension Workshop goals
• Identify key technical areas for
  • Aging, Reliability, & Material Lifetimes,
• Including cross-cutting needs for data analysis algorithms
  • (e.g., sparse/unreliable data)
• Propose topical areas for student projects
• Recommend paths for forming stronger partnerships
  • to benefit NNSA programs and missions

Rapidly assessing risk in COTS components?

Assessing risk of degradation in fielded systems?

Extracting insight from sparse, noisy data sets?

Linking models, accelerated aging, and field experiments?

We request help in identifying appropriate attendees from each site