DuraMAT Data Hub and Analytics Thrust: Collaborative tools for investigating PV Degradation

Ben Ellis, Robert White, Mike Deceglie, Anubhav Jain

Open data movement

- Open-sourcing data is becoming more common
- Open data empowers collaborative research
- Provides transparency in research methods and findings
- Wide variety of data available on may platforms

AWS Public Datasets

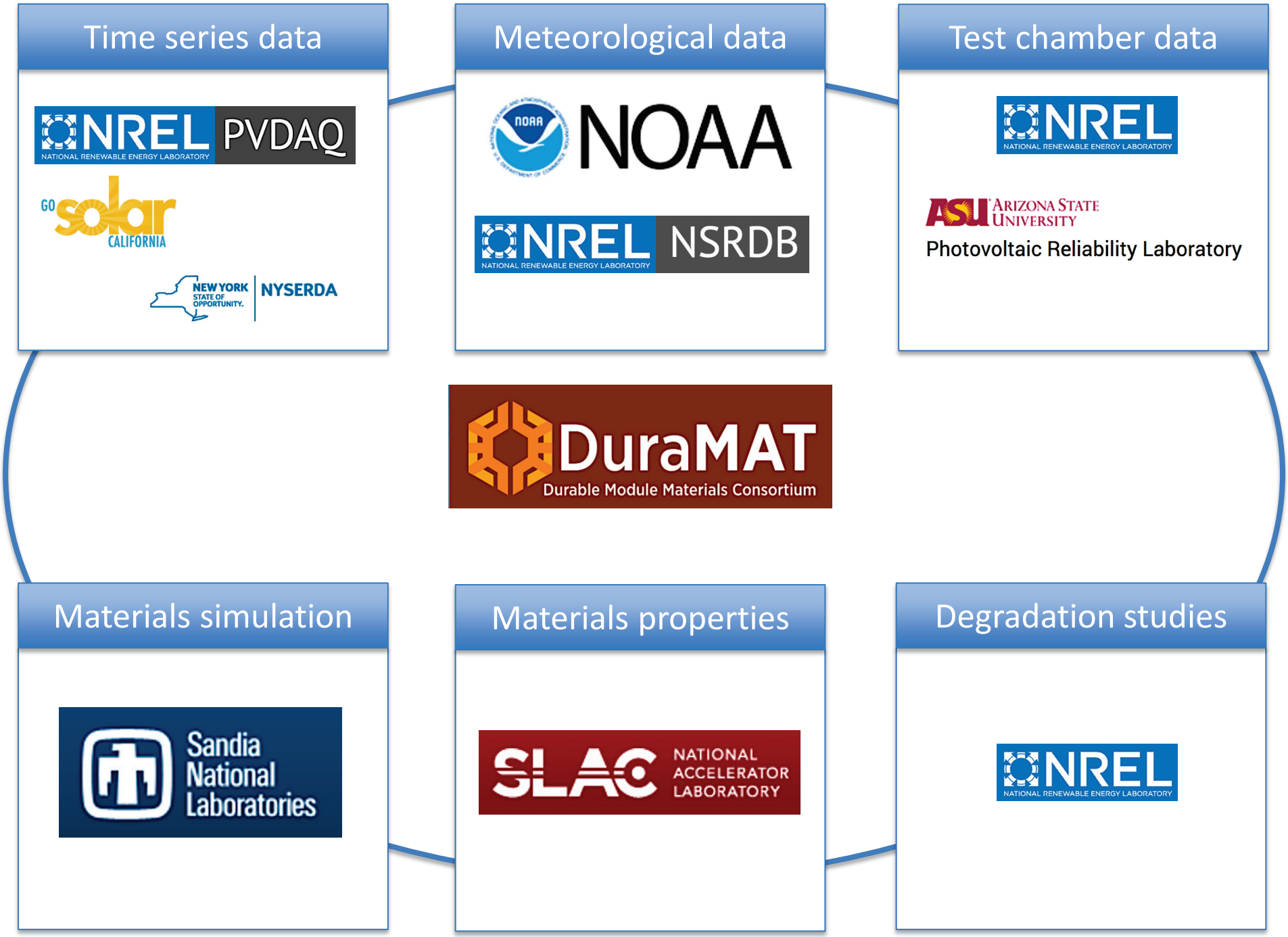
European Union Open Data Portal

PV community and open data

- Researchers have joined the big data movement!
- Projects of different sizes and scopes
- These communities and their data can be disconnected

DuraMAT

- The goal of DuraMAT is to create a centralized data community
- Connect researchers of different backgrounds and specialties together
- Provide data access and data sharing
- Host data that otherwise does not have a platform for open sharing
- Guide data producers and consumers to find new research opportunities


Time series data
NREL PVDAQ
NATIONAL RENEWABLE ENERGY LABORATORY
GO solar CALIFORNIA
NEW YORK STATE OF OPPORTUNITY. NYSERDA
Meteorological data
NOAA
NREL NSRDB
NATIONAL RENEWABLE ENERGY LABORATORY
Test chamber data
NREL
NATIONAL RENEWABLE ENERGY LABORATORY
ASU ARIZONA STATE UNIVERSITY
Photovoltaic Reliability Laboratory
DuraMAT
Durable Module Materials Consortium
Materials simulation
Sandia National Laboratories
Materials properties
SLAC NATIONAL ACCELERATOR LABORATORY
Degradation studies
NREL
NATIONAL RENEWABLE ENERGY LABORATORY

Time series data

NREL PVDAQ
NATIONAL RENEWABLE ENERGY LABORATORY

GO solar CALIFORNIA

NEW YORK STATE OF OPPORTUNITY. NYSERDA

Meteorological data

NOAA

NREL NSRDB
NATIONAL RENEWABLE ENERGY LABORATORY

Test chamber data

NREL
NATIONAL RENEWABLE ENERGY LABORATORY

ASU ARIZONA STATE UNIVERSITY
Photovoltaic Reliability Laboratory

DuraMAT
Durable Module Materials Consortium

We are always looking for more data

Materials simulation

Sandia National Laboratories

Materials properties

SLAC NATIONAL ACCELERATOR LABORATORY

Degradation studies

NREL
NATIONAL RENEWABLE ENERGY LABORATORY

Data diversity

- The data hub will not discriminate based on size or amount of data
  - Data sets can be a single spreadsheet up to an entire database
- Data standards are being developed to maintain some uniformity
  - Establish units of measurement
  - Define which properties and metrics to support
- Metadata is also a focus
  - Metadata will help others understand collection methods and equipment used
  - Help to search, sort data
  - Important for combining data from different sources

Data diversity

- Currently data from 106 sites in US
- Amount of data depends on system site
- Well documented metadata
  - Location (latitude, longitude, elevation, azimuth)
  - Module and inverter makes and models
  - Panel tilt, rated power
- Wide range of performance metrics
  - Energy output, POA irradiance, efficiencies, so on
- Performance data sampled at several frequencies
  - Minutely, hourly, daily, weekly, monthly

Data diversity

- Data collected on 504 installations
- Well documented metadata
  - Location (zip code, county)
  - Mounting method (rooftop, ground), tilt, azimuth
  - Module and inverter make, model, quantity
  - Rated power in AC and DC
- Collection ran from 2010 to 2016 at a frequency of 15 min
- Single performance metric
  - Energy output

Data diversity

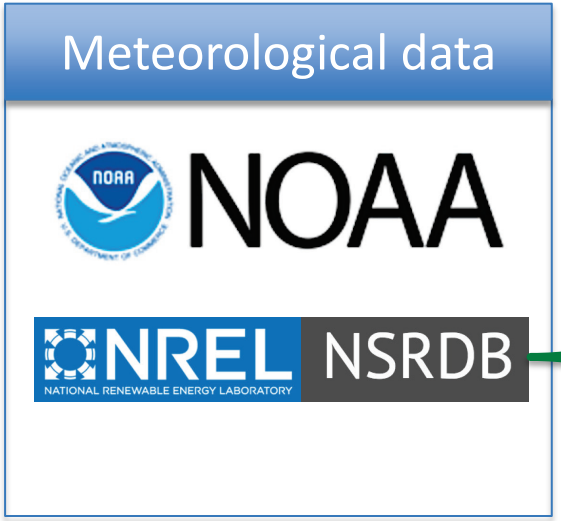
- Data collected on 256 installations
- Some metadata available
  - Location (address)
  - Rated power in DC
- Systems range in age
- Data sampled hourly, daily, and monthly
- Single performance metric
  - Energy output
- Includes temperature

Data diversity

Meteorological data

- Provides a wealth of data products
  - Daily, weekly, monthly normals and summaries
  - Precipitation data
  - Local climatological information
  - Weather radar
- Data sampling varies source to source
  - As fine as 5 minutes
  - As course as month-to-month

Data diversity

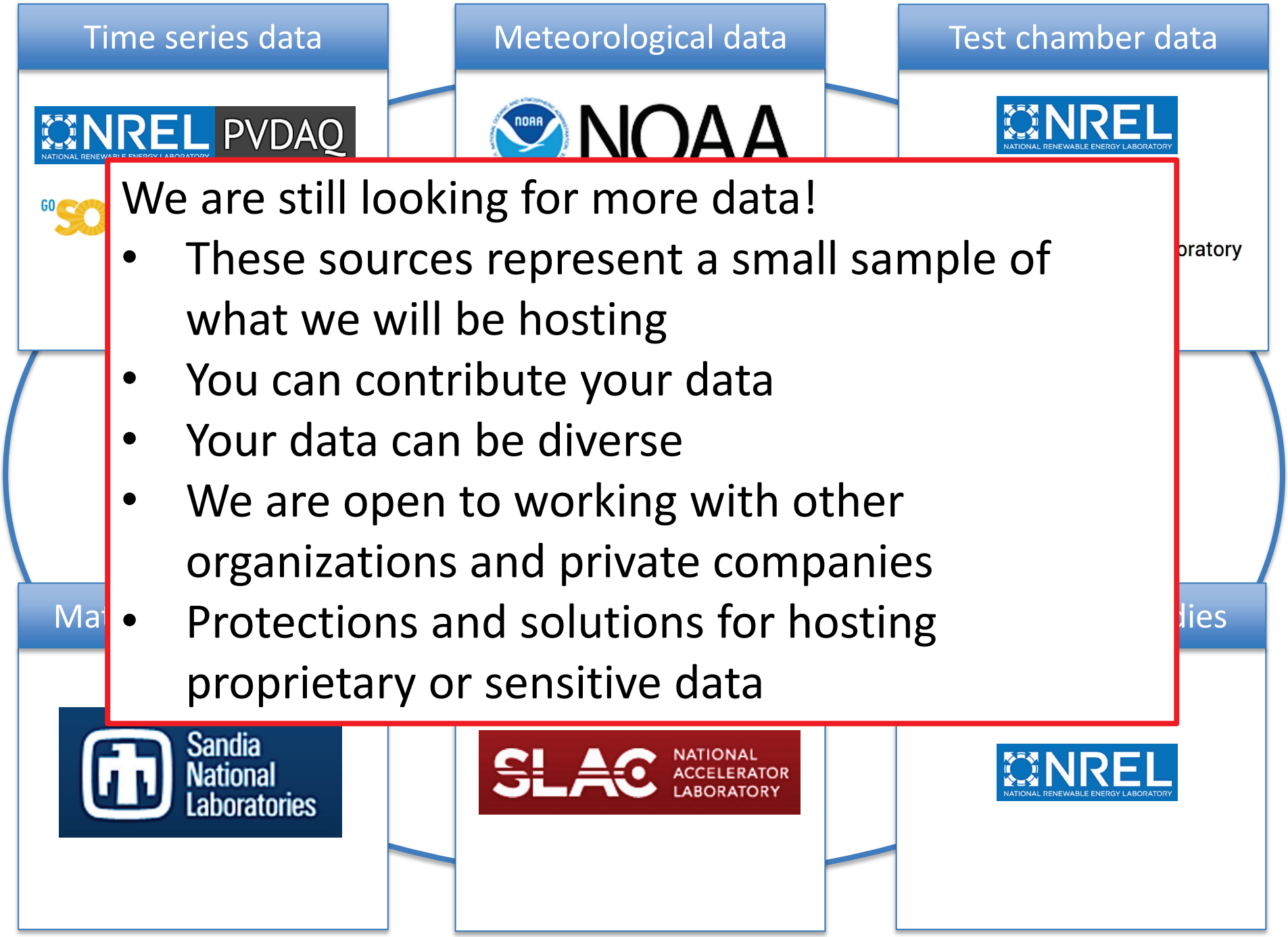


- Time series solar irradiance and meteorological data
- Many data fields reported
  - Temperature, relative humidity, pressure, precipitation, cloud cover and type
  - Irradiance data (global, direct, diffuse)
- Data collected every 30 minute or 60 minutes
  - Spans from 1961 up to 2015
- Spatial resolution of about $4km^2$
  - Covers most of US, parts of Canada, Central and South America, India

Data diversity

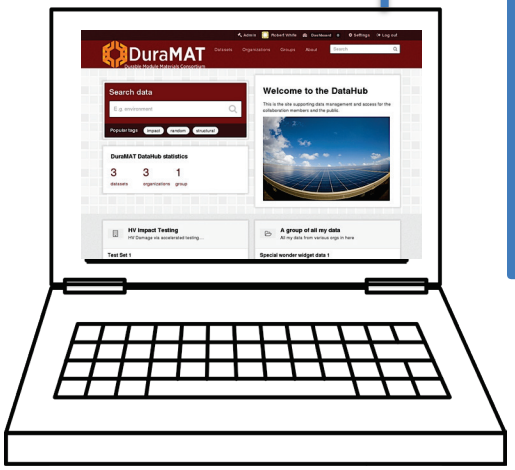
Degradation studies

- Literature survey done by Dirk Jordan at NREL
- Spreadsheet of about 11,000 systems
- System metadata
  - Location, climate classification
  - System size
  - Age of system
  - Technology (mono/multi crystalline Si, CdTe, CIGS)
- Degradation symptoms
  - Delamination, discoloration, corrosion, hot spots


Time series data
Meteorological data
Test chamber data
NREL PVDAQ
NOAA
NREL
NATIONAL RENEWABLE ENERGY LABORATORY
GOSO
oratory
We are still looking for more data!
• These sources represent a small sample of what we will be hosting
• You can contribute your data
• Your data can be diverse
• We are open to working with other organizations and private companies
• Protections and solutions for hosting proprietary or sensitive data
Mat
dies
Sandia National Laboratories
SLAC NATIONAL ACCELERATOR LABORATORY
NREL
NATIONAL RENEWABLE ENERGY LABORATORY

Accessibility

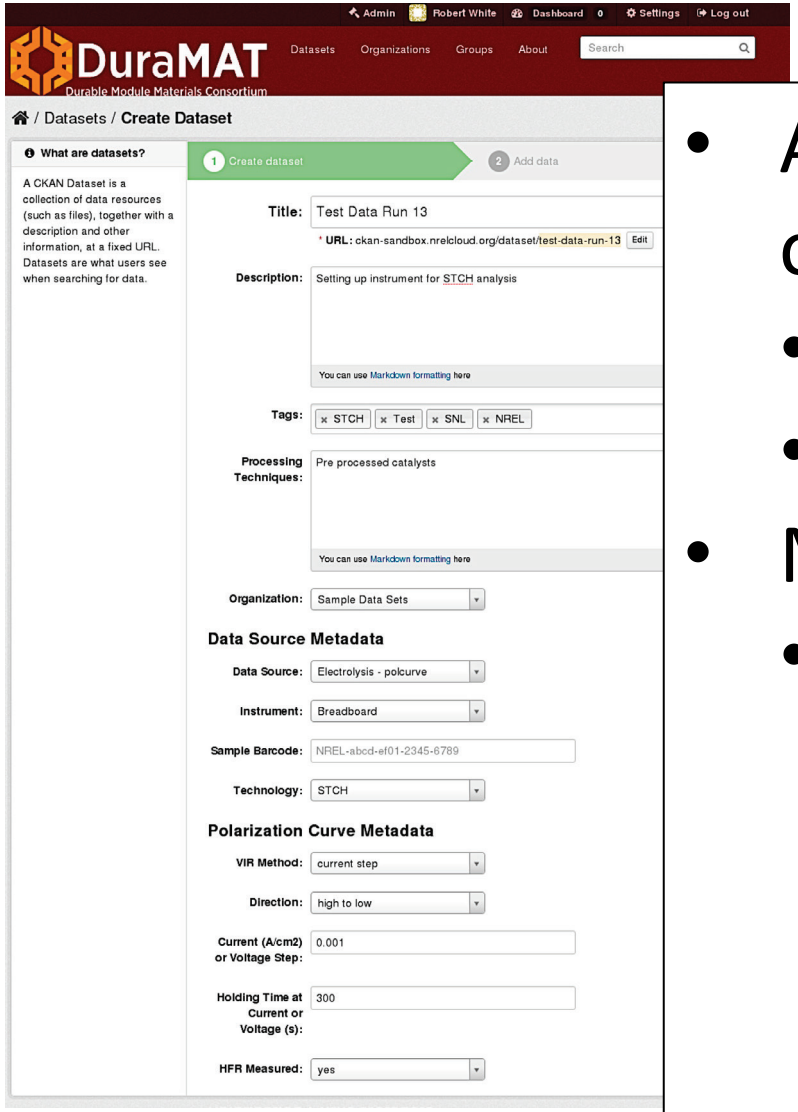
DuraMAT
Durable Module Materials Consortium



Data access via website

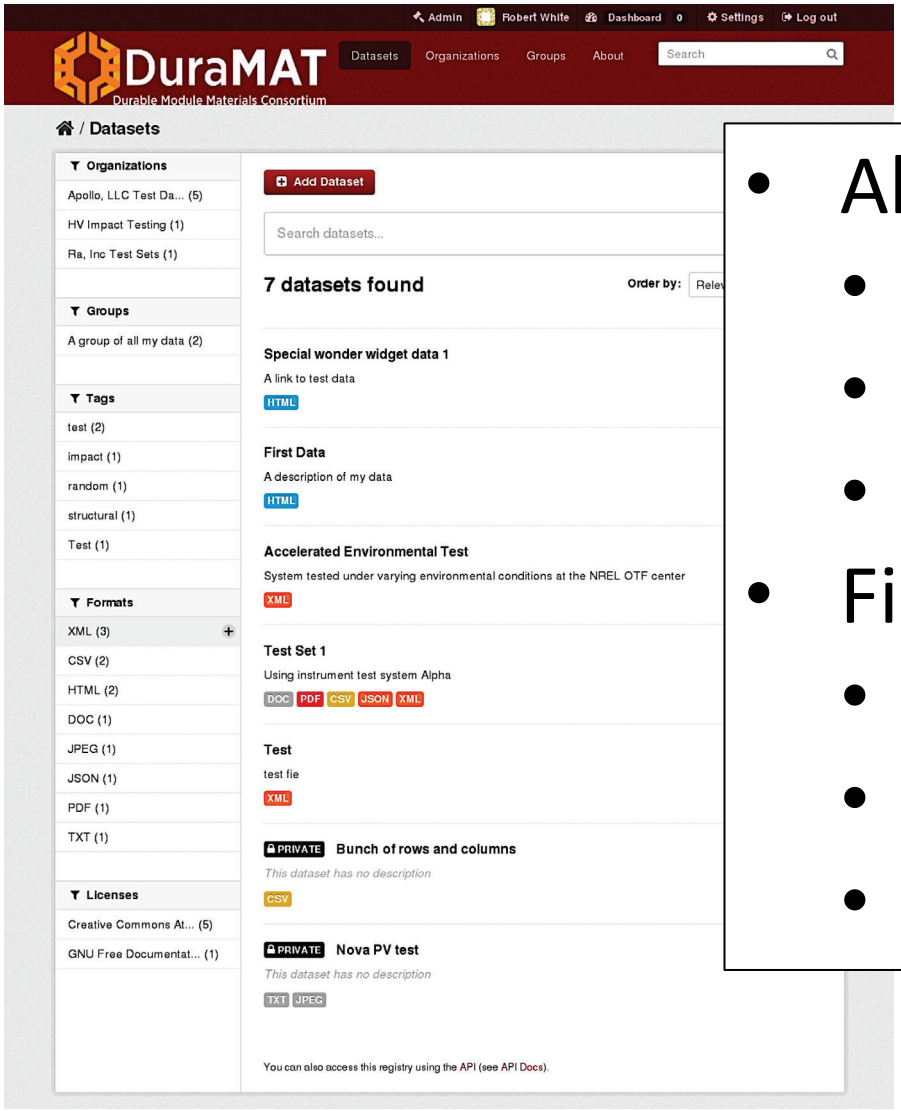
- Upload your data
- Browse available data
- Quick data analytics
- Download data

Uploading data



- Authors control presentation of data
  - Title
  - Description
- Metadata support
  - Options specific to data source
    - Automated collection
    - Test chamber
    - IV curve

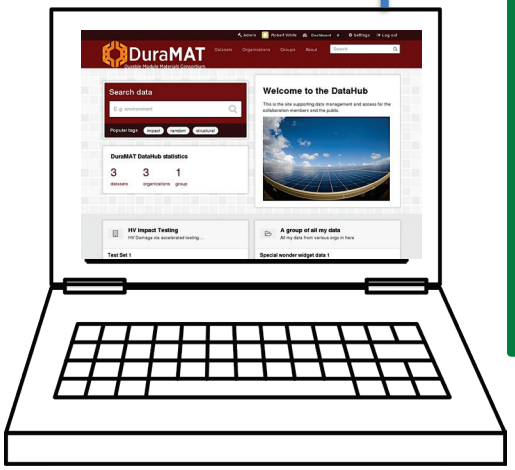
Browsing data sets



- All data sets displayed
  - Title
  - Format
  - Description
- Filter and searching
  - Keyword
  - Organization
  - Tags

Accessibility

DuraMAT
Durable Module Materials Consortium



Access via computer code

- Upload, browse, and download data
- Support larger queries for large-scale analytics projects (API)

Cross-source data aggregation

- Data is heterogeneous
  - Different collection frequencies, performance metrics, units of measurement
- Collecting and aggregating data is an important aspect this project
- Enabled by the centralized data hub design
  - Sources hosted in the data hub can be searched ‘behind the scenes’
  - Common data fields will be considered across multiple sources and aggregated
  - Units will be converted to maintain standardized output
- Need to build these tools and workflows

Cross-source data aggregation

Generate query

Performance data between 2010 and 2012 for CdTe systems

DuraMAT
Durable Module Materials Consortium

Search sources

Filter performance fields, module technology, year

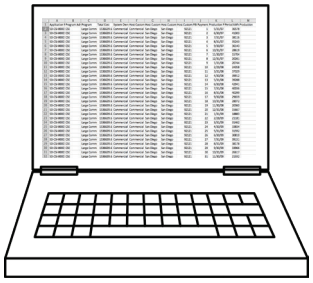
NREL PVDAQ
NATIONAL RENEWABLE ENERGY LABORATORY

NEW YORK STATE OF OPPORTUNITY. NYSERDA

GO solar CALIFORNIA

Process

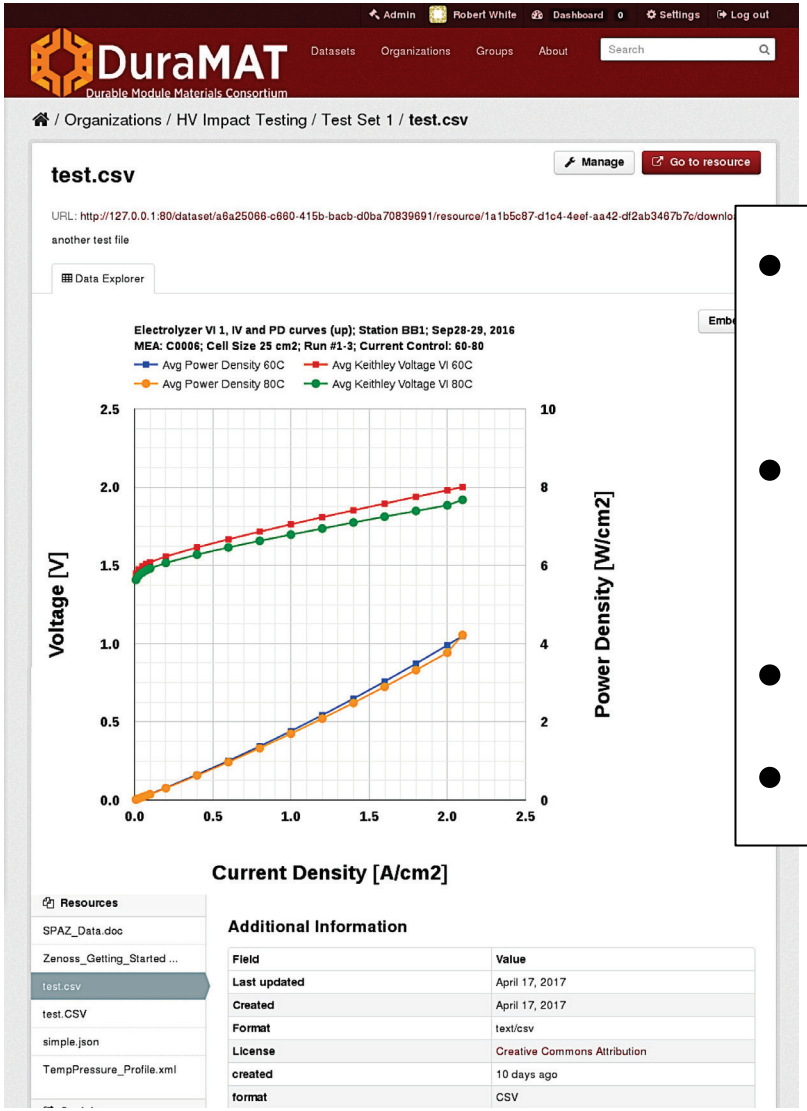
Find shared data fields, resample data frequency, convert units



Data analytics

- DuraMat is also developing software tools for analyzing data hosted on the hub
- Tools on website for analysis
  - Spreadsheet, graphing, mapping
- Open source software tools
  - Research methods and findings will be interpretable and transparent
  - Users (you) will be able to provide input (or directly contribute) to the development
- Analytics projects
  - We are open to collaborating with interested parties and partners
  - Leverage our software development and analytics experience with your domain knowledge

Website analytics



- Built-in graphing and charting for quick analysis
- Automatically works with certain data types
- Spreadsheet capability
- Mapping capability

Analytics projects

- Analysis of degradation modes
  - Data set from Dirk Jordan
- Time series analysis
  - Data from PVDAQ
- RdTools collaboration with NREL, kWh Analytics
  - Development of clear sky detection methods
- Potential induced degradation collaboration with NREL
- Coatings database and mining with SLAC

Degradation modes

- Data from literature survey from Dirk Jordan
- Failure modes and times studied across different mountings, climates, and technologies
- Further analysis being done at LBNL

PROGRESS IN PHOTOVOLTAICS: RESEARCH AND APPLICATIONS
Prog. Photovolt: Res. Appl. 2017; **25**:318–326
Published online 30 January 2017 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/pip.2866

RESEARCH ARTICLE

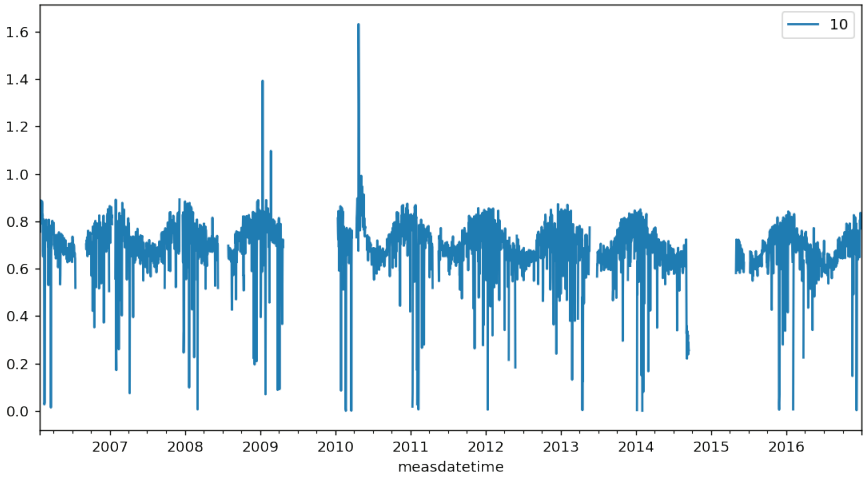
Photovoltaic failure and degradation modes

Dirk C. Jordan[1]*, Timothy J. Silverman[1], John H. Wohlgemuth[1], Sarah R. Kurtz[1] and Kaitlyn T. VanSant[2]

[1] National Renewable Energy Laboratory (NREL), 15013 Denver West Parkway, Golden, Colorado 80401, USA
[2] Colorado School of Mines, 1500 Illinois Street, Golden, Colorado 8040, USA

Time series analysis

- Investigation of degradation rates
- Currently using PVDAQ data
  - Expand study to encompass multiple data sources (NYSERDA, California Solar Initiative, so on)
- Correlate degradation rates
  - Technologies
  - Climates
  - Meteorological effects



RdTools clear sky detection

- Ongoing work by NREL and kWh Analytics
- Open source PV degradation analysis toolkit
- Clear sky detection for increased certainty in degradation rates
  - Working with Mike Deceglie, Dan Ruth, and Ambarish Nag at NREL
  - Current clear sky detection method is from PVLib (software from SNL)

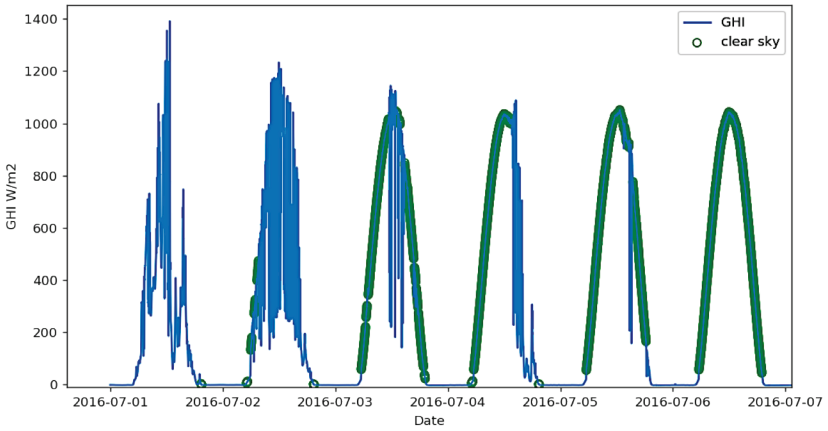
https://pvpmc.sandia.gov/applications/pv_lib-toolbox/

Reno, M.J. and C.W. Hansen, "Identification of periods of clear sky irradiance in time series of GHI measurements" Renewable Energy, v90, p. 520-531, 2016.

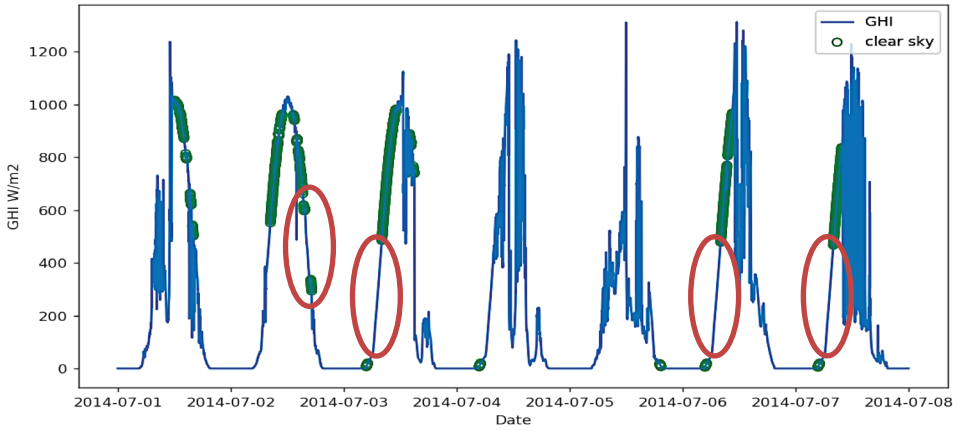
Clear sky detection

- Scalability challenges with current methodology
  - Single input - irradiance
  - Uses five metrics to determine clear and obscured skies for given periods of time

Sandia RTC

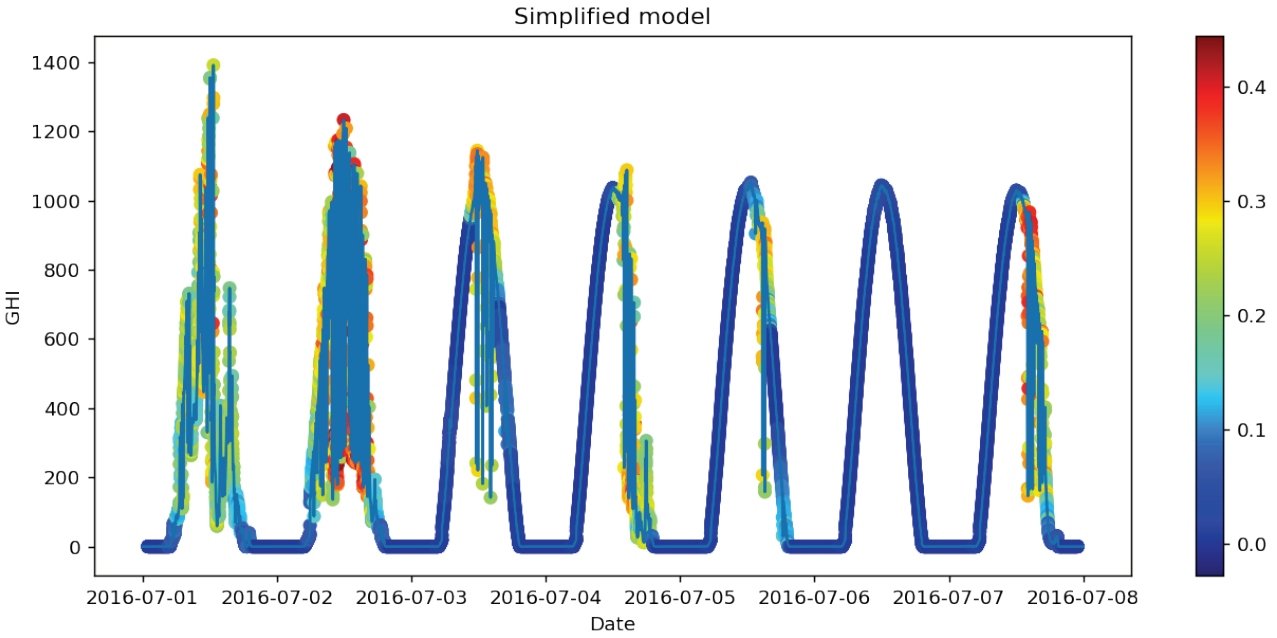


NREL SRRL Baseline Measurement System



Clear sky detection

- Use modern statistical and machine learning methodology
- Develop a simpler method with fewer parameters



Prospective projects

- Potential induced degradation
  - Collaboration with Peter Hacke (NREL)
  - Develop models to identify and quantify PID
- Coatings data
  - Discussions with Laura Schelhas and Mike Toney (SLAC)
  - Develop and mine coatings properties and correlate with PV performance

We are still seeking collaborations

Acknowledgements

- Anubhav Jain
- Michael Kim
- Robert White
- Jonathan Trinastic
- Birk Jones
- Dirk Jordan
- Mike Deceglie
- Dan Ruth
- Ambarish Nag
- Laura Schelhas
- Mike Toney
- Peter Hacke
- Teresa Barnes

Moving forward

Please talk to me if you have…

- Data to share
- Analytics projects open to collaboration

bhellis@lbl.gov

Thank you!