Exceptional service in the national interest

Sandia National Laboratories

Pecos, Open Source Software for PV Performance Monitoring

8th PV Performance Modeling and Monitoring Workshop

Albuquerque, NM, May 9-10, 2017

Katherine A. Klise

Sandia National Laboratories, Albuquerque, NM

U.S. DEPARTMENT OF ENERGY

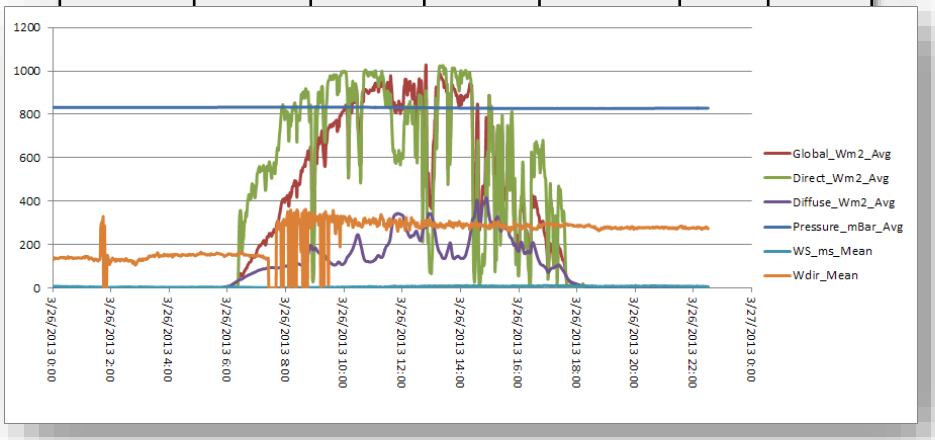
National Nuclear Security Administration

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. SAND2017-4987 C

Overview

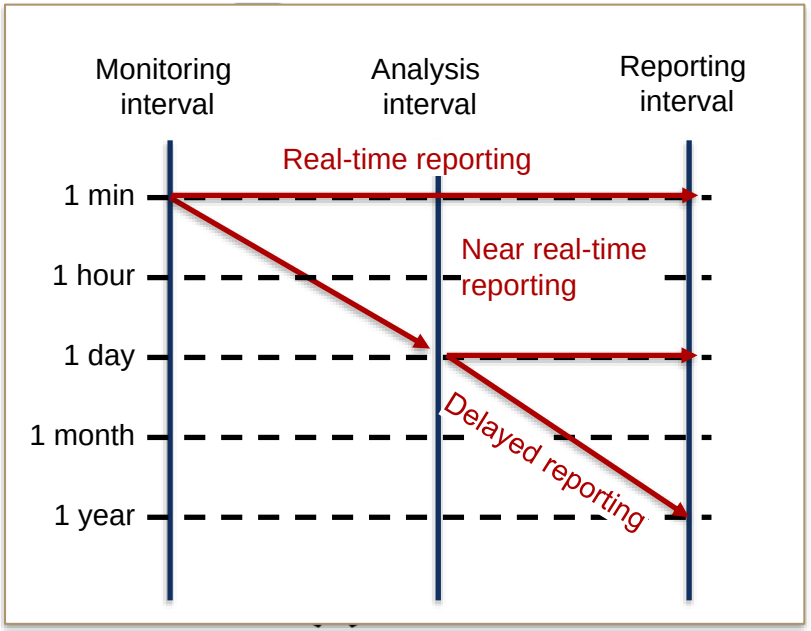
- What is Pecos? (PAY-cose)
  - Software for automated quality control and performance monitoring of time series data
- Why use Pecos?
  - Collect large amounts of data on multiple systems and locations
  - Run automated quality control tests on that data
  - Alert system operators when the system has changed
  - Generate reports
  - Collect performance statistics to track long term system health
  - Compare system performance across sites

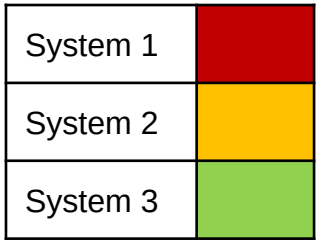
TOA5	CR1000	46385	CR1000.Std.24	CPU:ABQ_RTC_MET_2013_03_21.CR1	58869	DataOut
TIMESTAMP	Global_Wm2_Avg	Direct_Wm2_Avg	Diffuse_Wm2_Avg	Pressure_mBar_Avg	WS_ms_Mean	Wdir_Mean
TS						Deg
	Avg	Avg	Avg	Avg	WVc	WVc
3/26/2013 0:00	-1.16195	-0.45458	0	832.121	6.338	135.7
3/26/2013 0:01	-1.14918	-0.5455	0	832.123	5.8	136.4
3/26/2013 0:02	-1.14918	-0.52277	0	832.106	5.988	131.2
3/26/2013 0:03	-1.14918	-0.45458	0	832.0875	6.838	139.6
3/26/2013 0:04	-1.14918	-0.45458	0	832.0799	6.825	136.8
3/26/2013 0:05	-1.14918	-0.45458	0	832.0693	6.775	137
3/26/2013 0:06	-1.14919	-0.40155	0	832.0547	6.825	135.2
3/26/2013 0:07	-1.14919	-0.31063	0	832.0114	6.85	137.4
3/26/2013 0:08	-1.14921	-0.46217	0	832.0062	7.013	136.3
3/26/2013 0:09	-1.14922	-0.45459	0	832.0159	7	135.1
3/26/2013 0:10	-1.14922	-0.45459	0	832.0093	6.063	136.4
3/26/2013 0:11	-1.14921	-0.45459	0	832.0027	6.825	134.6
3/26/2013 0:12	-1.14921	-0.45459	0	831.9932	6.813	135.8
3/26/2013 0:13	-1.14921	-0.36367	0	831.9811	6.65	137.2
3/26/2013 0:14	-1.14921	-0.28791	0	832.0098	7	137.1
3/26/2013 0:15	-1.14921	-0.45459	0	832.0153	6.738	138.6
3/26/2013 0:16	-1.1492	-0.45459	0	831.9963	6.613	141.1
3/26/2013 0:17	-1.1492	-0.60612	0	832.0099	6.125	139.8
3/26/2013 0:18	-1.1492	-0.84099	0	832.0046	6.113	139.9

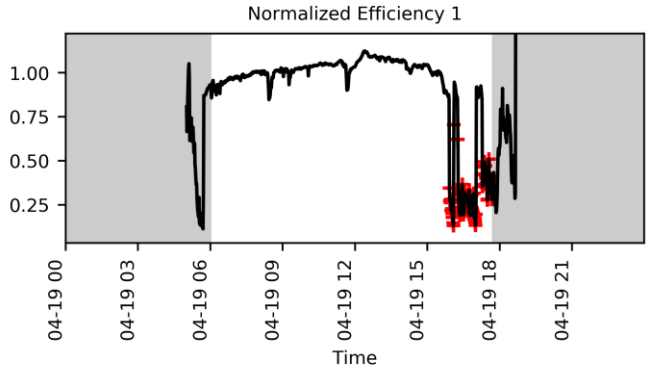


Getting started

- Retrieve data
  - From sensor, database, files, or from the web
  - 3Vs (volume, velocity, and variety)
  - Single or repeat (automated)
- Define analysis
  - Analysis/reporting time interval
  - Filters
  - Integrate models
  - Quality control tests
  - Metrics
- Final product
  - Simple to complex
  - Red/yellow/green approach
  - Time series or interactive graphics
  - Performance history
  - Dashboards hosted on the web
  - Email alerts







Time series data

- Time series data loaded into Pecos as a Pandas DataFrame
  - Powerful time series analysis options
  - Datetime and timezone recognition
  - Merge multiple DataFrames in a single analysis (i.e. electrical and weather)
  - Data can be easily loaded from database, file, or web
- *New* Data acquisition methods recently added to Pecos
  - Transfer data from sensors to an SQL database
- User defines the analysis timeframe (minute, hour, day, month, …)
- Data can be grouped and renamed according to type
- Repeat analysis automated using OS task scheduler (cron, tasks)

From database

```
sql_con= MySQLdb.connect(host=ip_address, port=…)
sql_query = "SELECT * FROM table…"
df = pandas.read_sql(sql_query, con=sql_con)
```

From file

```
df = pandas.read_csv(filename)
```

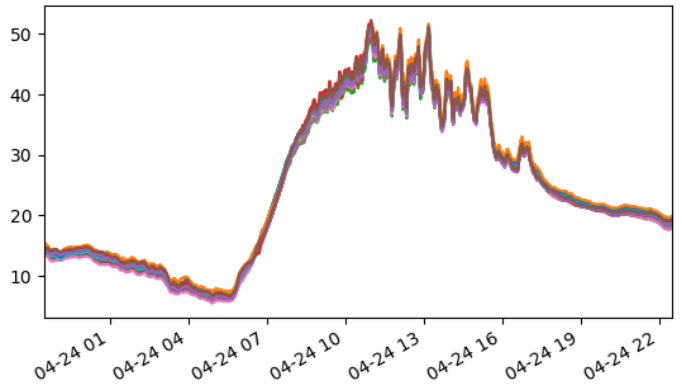
From the web

```
response = requests.get(url=http://developer.nrel.gov/pvdaq/api/…)
data = json.loads(response.text)
df = pandas.DataFrame(data=data['outputs']['data'])
```

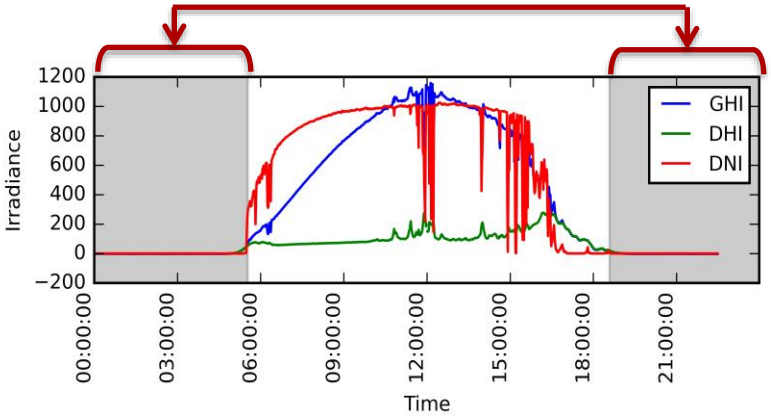
Pre-processing filters

- Filter data
  - Smoothing
  - Upscale/downscale
- Fill missing data
  - Interpolation (linear, polynomial, etc.)
  - Duplicate sensors
  - Historic/regional data
  - Data generated from models
- Time filter
  - Conditional statement that exclude specific timestamps from quality control tests
  - Time filter can be based on:
    - Time of day (i.e. before 8 am and after 5 pm)
    - Sun position (i.e. sun elevation < 10 degrees)
    - Data properties (i.e. irradiance < 200 W/m^2)

Module temperatures from 16 sensors

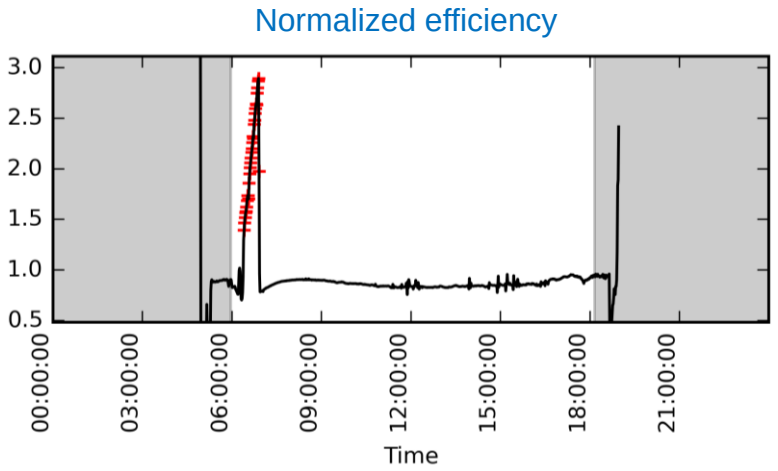


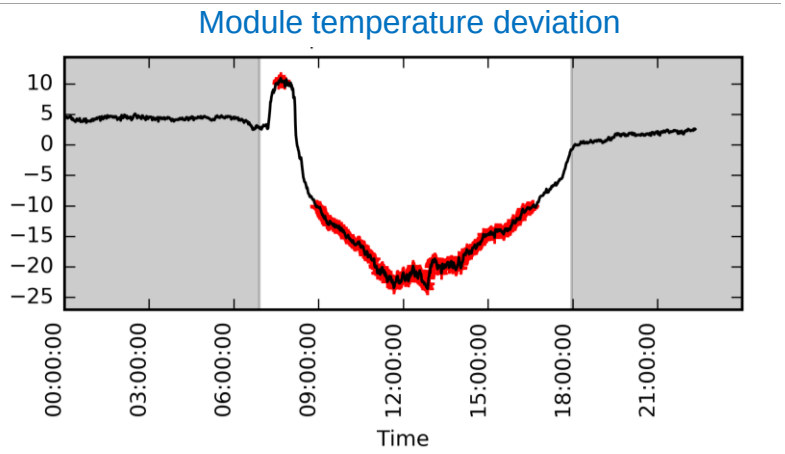
Time filter based on sun position



Composite signals

- Composite signals are used to create new data from existing data or from a model
  - Compute relationships between data columns
  - Compare measured data to a model
    - PVLIB performance model
    - Machine learning
- Examples
  - DC Power from current and voltage
  - Inverter efficiency from DC and AC power
  - Normalized efficiency from power and irradiance
  - Module temperature deviation
  - Relative error between model and data
- Composite signals can be used in the quality control tests

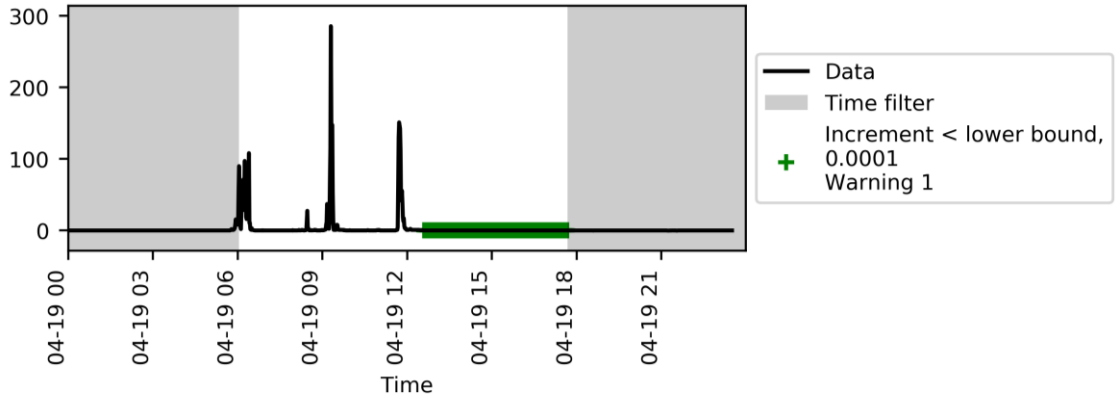




Quality Control tests

- Quality controls tests fall into five categories
  - Timestamp test
  - Missing data test
  - Corrupt data test
  - Range test
  - Dead sensor/abrupt change tests
- When a test fails, information is stored in a summary table which can be included in automated reports and saved to file/database. Graphics can be produced that pin point the data points that caused the test failure.

System Name	Variable Name	Start Date	End Date	Timesteps	Error Flag
PV System 1	Direct_Wm2	2017-04-19 12:36:00	2017-04-19 17:40:00	305	Increment < lower bound, 0.0001



Quality Control tests

- Timestamp test identifies duplicate, non-monotonic, and missing timestamps. *New* Irregular timestamps can be preserved.
- Missing data test identifies column-time pairs that are missing.
- Corrupt data test screens for datalogger values that indicate corrupt data.

Original data

TIMESTAMP	Column A	Column B
1/1/2017 0:00	0	1
1/1/2017 1:00		2
1/1/2017 2:00	2	3
1/1/2017 3:00	3	4
1/1/2017 5:00	5	1
1/1/2017 6:00	6	2
1/1/2017 8:00	8	-999
1/1/2017 7:00	7	-999
1/1/2017 9:00	9	1
1/1/2017 9:00	9.5	2

Missing data

Missing timestamp

Non-monotonic timestamp

Duplicate timestamp

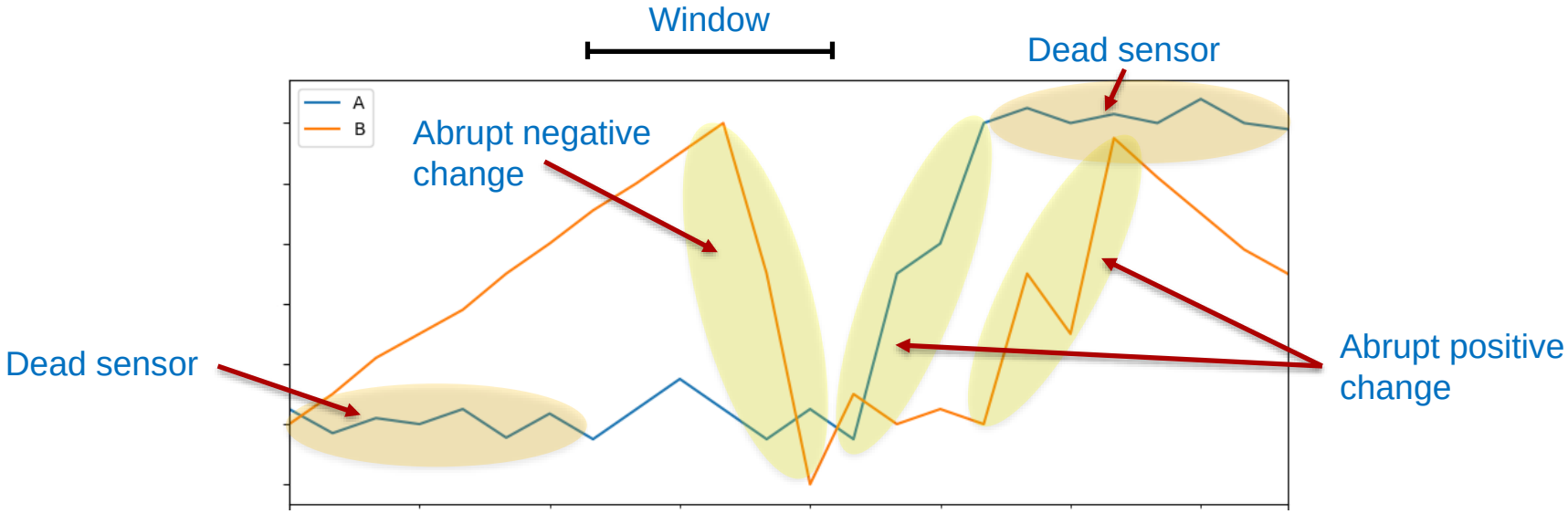
Corrupt data

Corrected data

TIMESTAMP	Column A	Column B
1/1/2017 0:00	0	1
1/1/2017 1:00	NaN	2
1/1/2017 2:00	2	3
1/1/2017 3:00	3	4
1/1/2017 4:00	NaN	NaN
1/1/2017 5:00	5	1
1/1/2017 6:00	6	2
1/1/2017 7:00	7	NaN
1/1/2017 8:00	8	NaN
1/1/2017 9:00	9	1

Quality Control tests

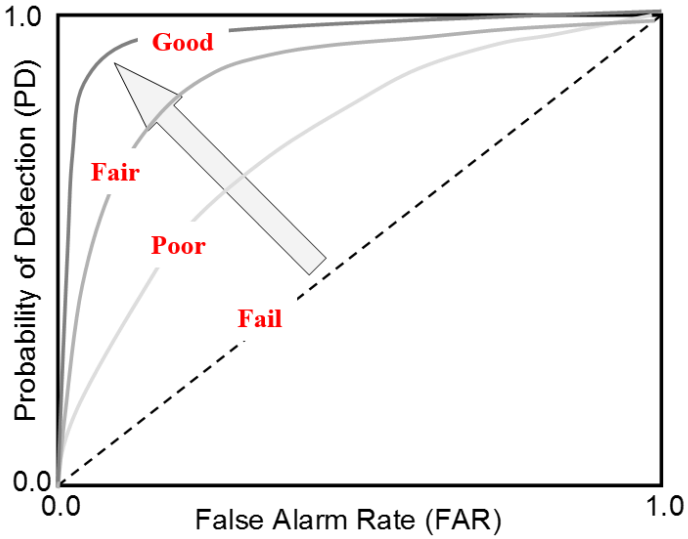
- Range tests checks if data is within expected bounds
  - Ambient temperature should be between -30 and 50 degrees C
  - Normalized efficiency (composite signal) should be between 0.5 and 1
- *New* Dead sensor/abrupt change test checks if the difference between min and max is within expected bounds over a given time span
  - Voltage should not change by more than 80% rating within 15 minutes
  - The rain gauge should not increase by more than 2 inches in an hour
  - If the irradiance sensor changes by less than 0.0001 in 5 hours, it's probably dead



Evaluating quality control tests

- *New* Evaluate how well a quality control test (or set of quality control tests) distinguishes normal from anomalous conditions.
  - Probability of detection
  - False alarm rate
- Strategies to reduce false positives and false negatives
  - Adjust thresholds
  - Specify the minimum number of consecutive failures needed to signal a warning
  - Smooth data before running quality control tests

	Actual normal condition	**Actual anomalous condition**
Estimated normal condition	True negative (TN)	False negative (FN)
Estimated anomalous condition	False positive (FP)	True positive (TP)
	FAR = TN / (TN+FP) FAR = 1-Specificity	PD = TP / (TP+FN) PD = Sensitivity



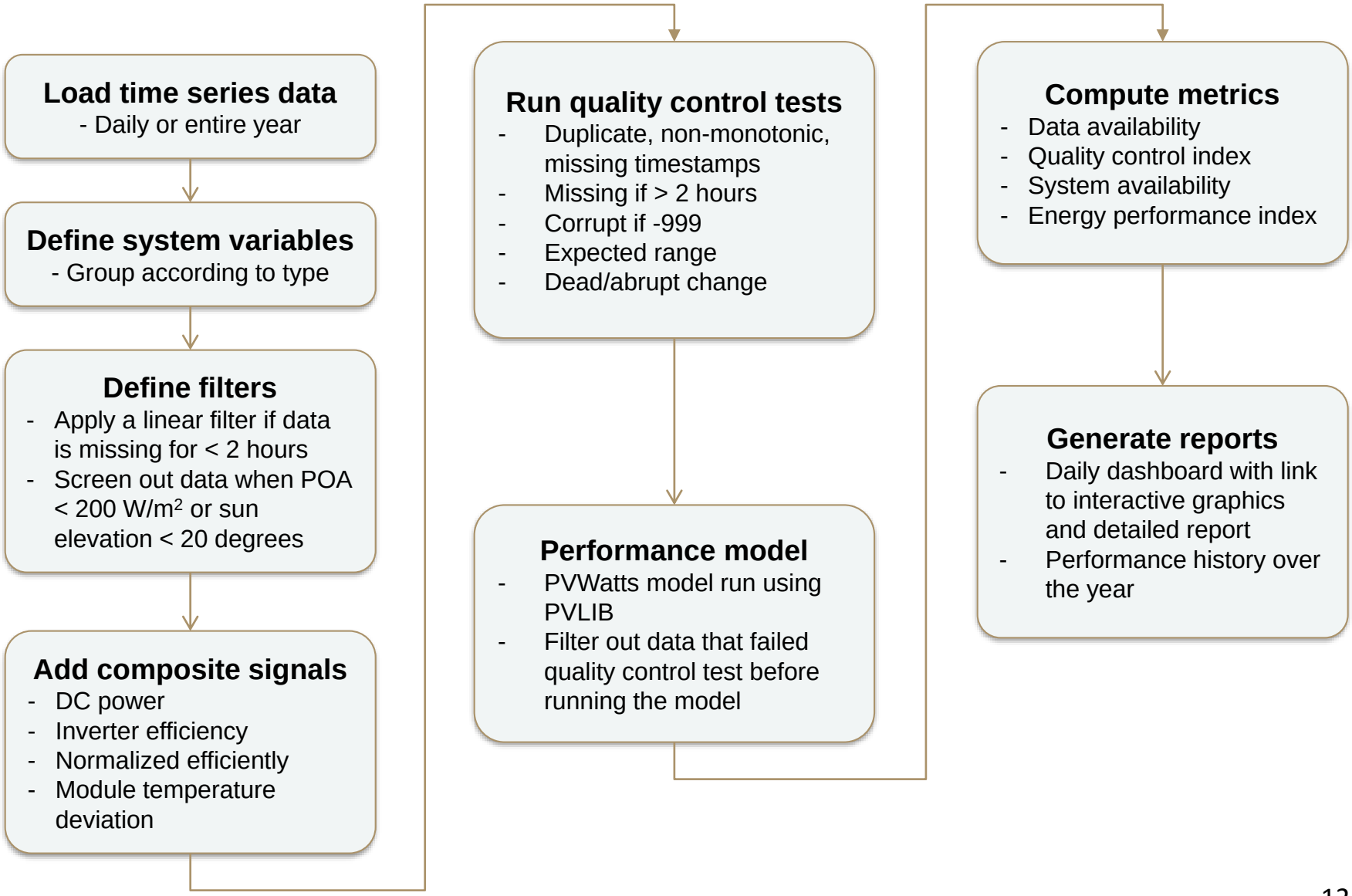
RTC quality control analysis

- Regional Test Center Baseline and Weather systems
  - New Mexico, Florida, Vermont, Nevada
  - 2 strings of 12 Suniva Optimus 270 Black modules
- Quality control tests and performance metrics based on IEC 61724
  - Check for data outside expected range, dead sensors, and abrupt changes
  - Compute in-service and all-in energy performance index
- Analysis run daily (near real-time), results emailed to stakeholders.
- End of year report

Module specs: Pmax = 270 W, Vmp = 31.2 V, Voc = 38.5 V, Imp = 8.68 A, Isc = 9.15 A

Weather data	Baseline PV data
GHI, DNI, DHI, air pressure, wind speed, wind direction, relative humidity	For each string: DC voltage, DC current, AC voltage, AC current, AC power, power factor, frequency, reference cell irradiance, and reference cell temperature

RTC quality control analysis



RTC quality control analysis

- Expected range, dead sensor, abrupt change thresholds

Variable	Expected range	Dead sensor threshold	Abrupt change threshold
DC current and AC current (A)	> 0 and < Imp*1.5	< 0.0001 in 5 hours	
DC voltage and AC voltage (V)	> 0 and < Vmp*12*1.5	< 0.0001 in 5 hours	
DC power* and AC power (W)	> 0 and < Pmp*12*1.5	< 0.0001 in 5 hours	> Pmp*12*0.8 in 15 min
Power factor	> -1 and < 1	< 0.0001 in 5 hours	
Frequency(Hz)	> 57 and < 63	< 0.0001 in 5 hours	
POA, DNI, GHI, and ref cell irradiance (W/m^2)	> 0 and < 1500	< 0.0001 in 5 hours	
DHI (W/m^2)	> 0 and < 500	< 0.0001 in 5 hours	
Air pressure (mbar)	> 800 and < 1020	< 0.0001 in 5 hours	> 100 in 15 minutes
Wind speed (m/s)	> 0 and < 32	< 0.0001 in 5 hours	
Wind direction	> 0 and < 360	< 0.0001 in 5 hours	
Relative humidity	> 0 and < 100	< 0.0001 in 5 hours	> 50 in 15 minutes
Ambient temperature (^oC)	> -30 and < 50	< 0.0001 in 5 hours	> 20 in 15 minutes
Module and ref cell temp (^oC)	> -30 and < 90	< 0.0001 in 5 hours	> 20 in 15 minutes
Module temp deviation (^oC)*	> -10 and < 10		
Inverter efficiency*	> 0.5 and < 1		> 0.25 in 15 minutes
Normalized efficiency*	> 0.5 and < 1		> 0.25 in 15 minutes

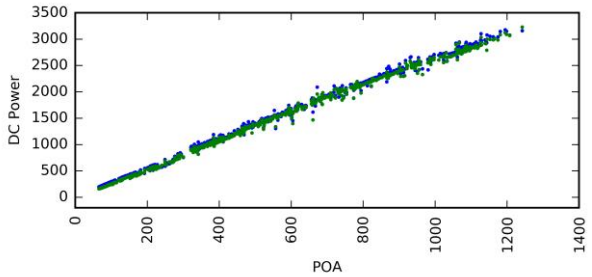
* Composite signal

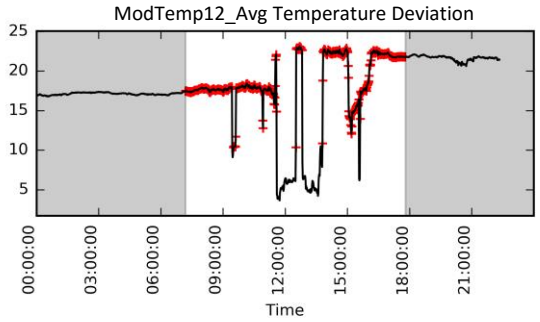
RTC quality control analysis

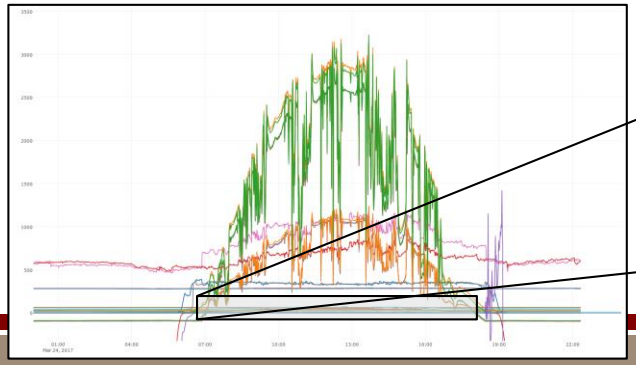
Daily report, red/yellow/green dashboard with links to details and interactive graphics

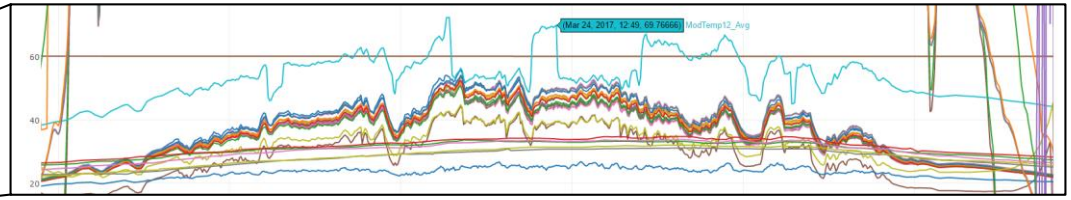
RTC Dashboard for 2017-03-24

	New Mexico	Florida	Vermont	Nevada
Weather	Irradiance 1.00 Wind 1.00 Air Pressure 1.00 Humidity 1.00 Rainfall 1.00 Datalogger 1.00 Detailed Report Interactive Plot	Irradiance 1.00 Wind 1.00 Air Pressure 1.00 Humidity 1.00 Rainfall 1.00 Datalogger 1.00 Detailed Report Interactive Plot	Irradiance 0.57 Wind 1.00 Air Pressure 1.00 Humidity 1.00 Rainfall 1.00 Datalogger 1.00 Detailed Report Interactive Plot	Irradiance 1.00 Wind 1.00 Air Pressure 1.00 Humidity 1.00 Rainfall 1.00 Datalogger 1.00 Detailed Report Interactive Plot
Baseline	Irradiance 1.00 Temperature 1.00 Current 1.00 Voltage 1.00 Power 0.64 Detailed Report Interactive Plot	Irradiance 1.00 Temperature 0.19 Current 1.00 Voltage 1.00 Power 0.93 Detailed Report Interactive Plot	Irradiance 1.00 Temperature 1.00 Current 0.27 Voltage 1.00 Power 0.24 Detailed Report Interactive Plot	Irradiance 1.00 Temperature 1.00 Current 1.00 Voltage 1.00 Power 1.00 Detailed Report Interactive Plot





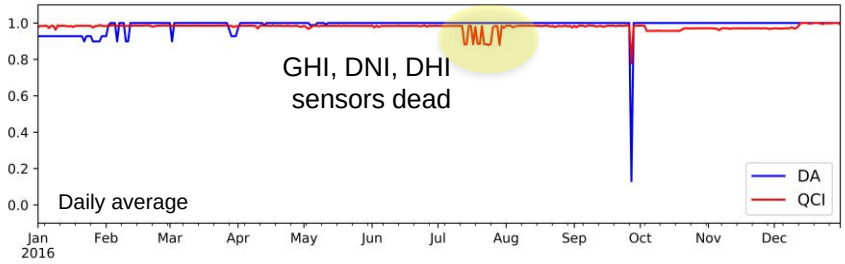


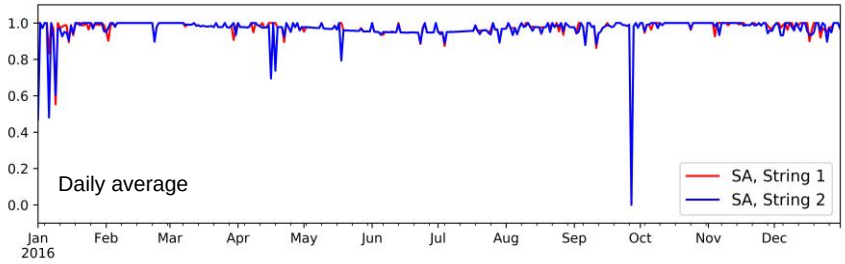


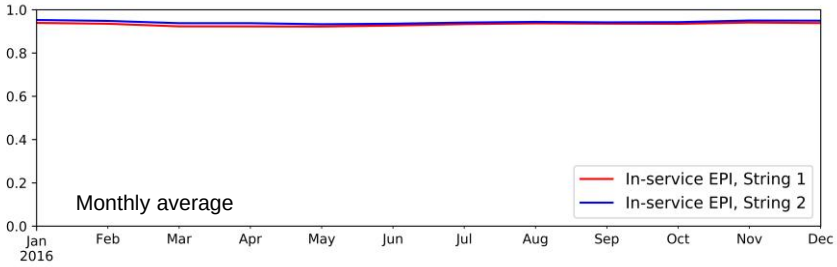
RTC quality control analysis

Yearly report, daily and monthly metrics

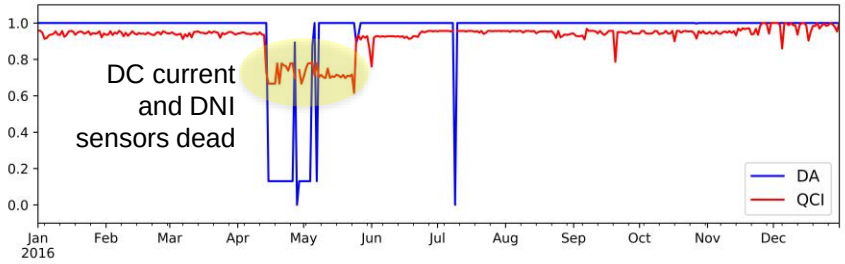
New Mexico



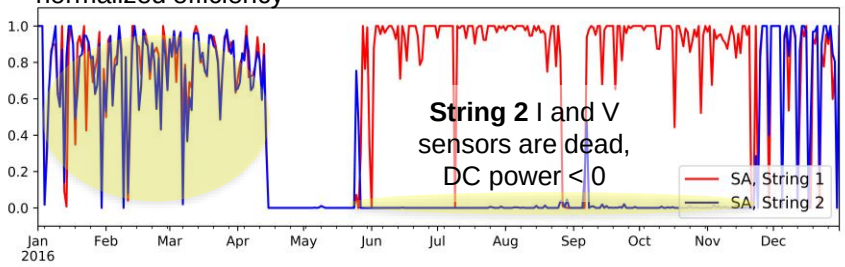


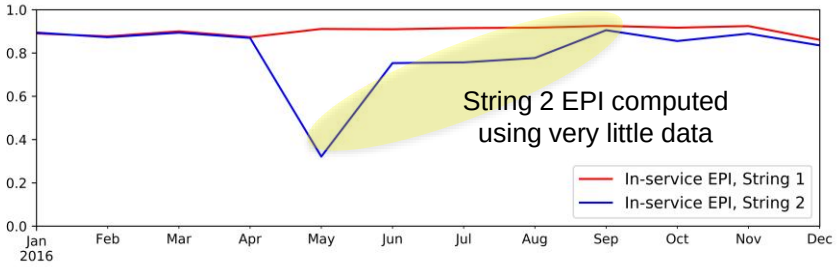


Vermont



Periodic low
normalized efficiency





Pecos

- Open-source python package
  - Python 2.7, 3.4, or 3.5
  - Revised BSD License
- Software repository
  - https://github.com/sandialabs/pecos
- Documentation
  - http://pecos.readthedocs.io
- Software testing
  - https://travis-ci.org/sandialabs/pecos
- 'Getting started' examples included with the software
- Version 0.1.5 (master branch)
  - New features include data acquisition, more flexible dashboards, PD and FAR metrics, compatibility with irregular timestamps, improved efficiency