Automatic Detection of Soiling Zones and Rates for Optimized PV Plant Cleaning

Envision Energy USA

Zoe DeFreitas, Sr Performance Engineer

2017 9th PV Performance Modeling and Monitoring Workshop
Solar Research Institute of HIT

Weihai, China December 5-7, 2017

Outline

Company Introduction

Motivation

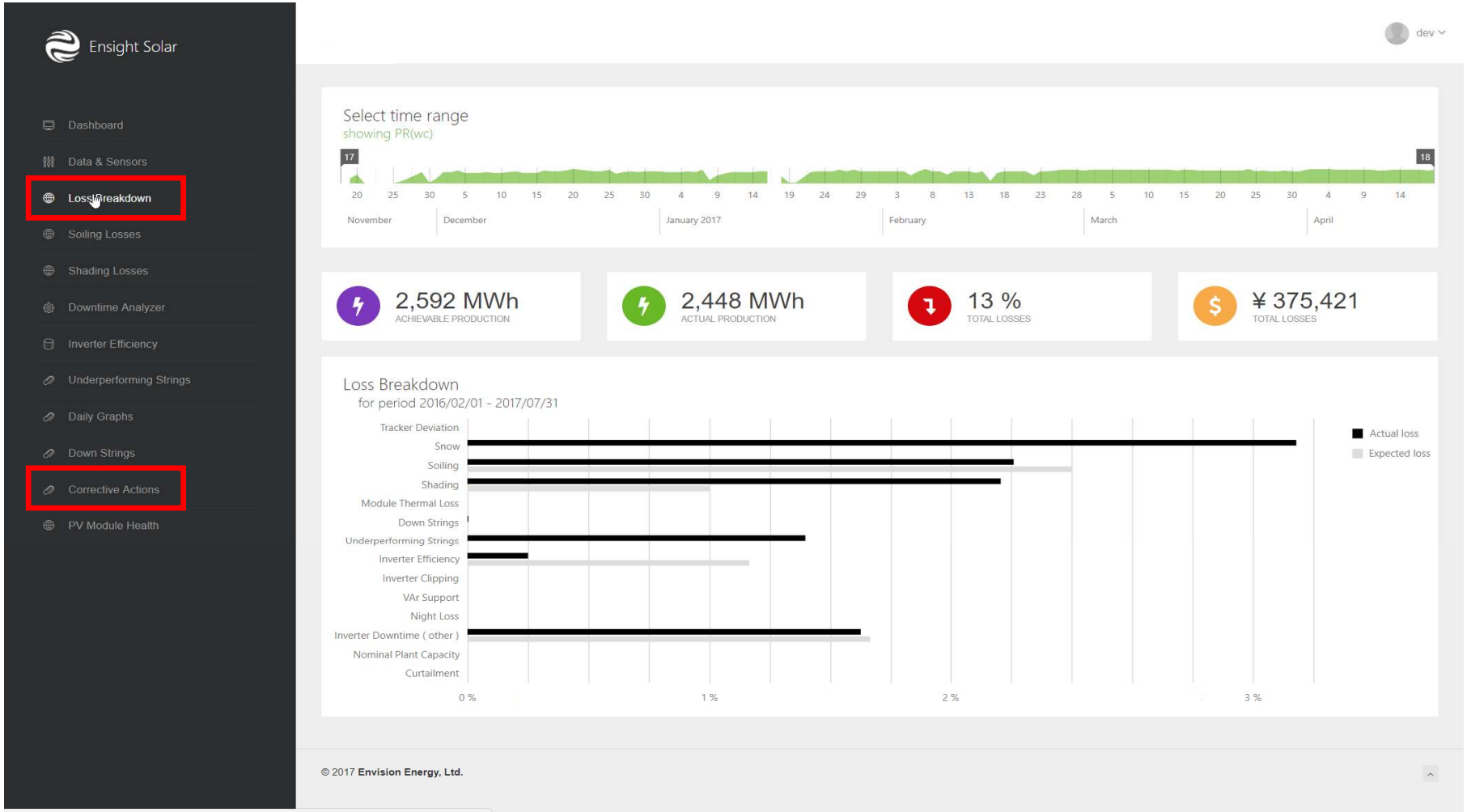
Our Soiling Analytics Process

- Data Filtering
- Soiling Rate Extraction
- Soiling Zone Clustering
- Wash Optimization

Conclusion

Introducing Envision Energy - A Global Leader in Renewable Energy Solutions

Ensight - Analytics for Solar PV Power Plants



What is our Motivation?

Problem Statement

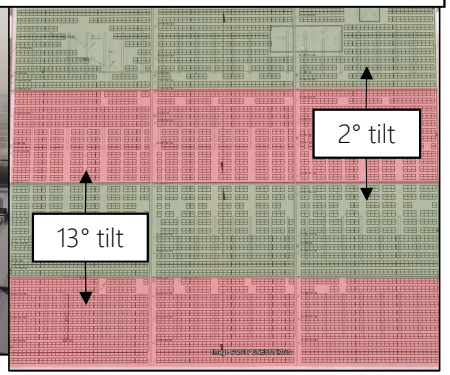
- **Soiling rates can vary** depending on a site's location and also within a plant.
- **Local soiling factors** including pollution sources, roads, agriculture, prevailing winds, and tilt can have a more local impact.
- A method is presented to **automatically detect soiling rates and cluster them by zones within a plant**
- With clustered soiling rates, the **O&M team can determine an optimum wash plan to maximize profits**

A PV plant suffering from extreme soiling due to pollution.

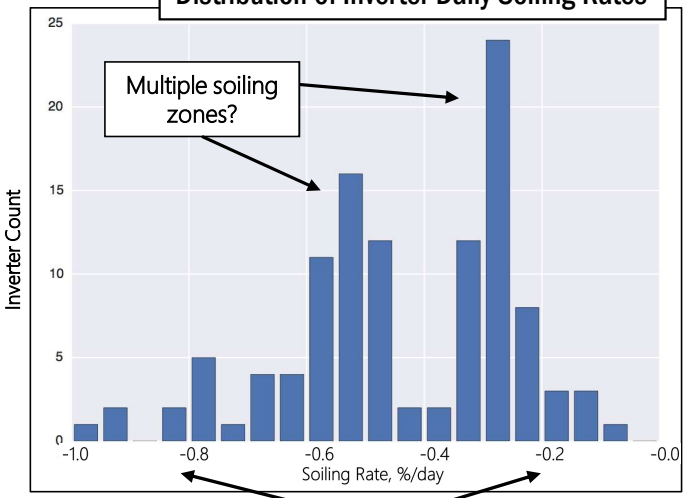
- **The plant:** 6MW rooftop site with 1 active chimney and 2 array tilts
- **Excessive soiling** due to pollution from chimney
- **Large spread in soiling rates** from -0.1 to-1.0% energy loss per day

Rooftop Site with Heavy Pollution From Chimney

Site Drawings with Two Different Tilts

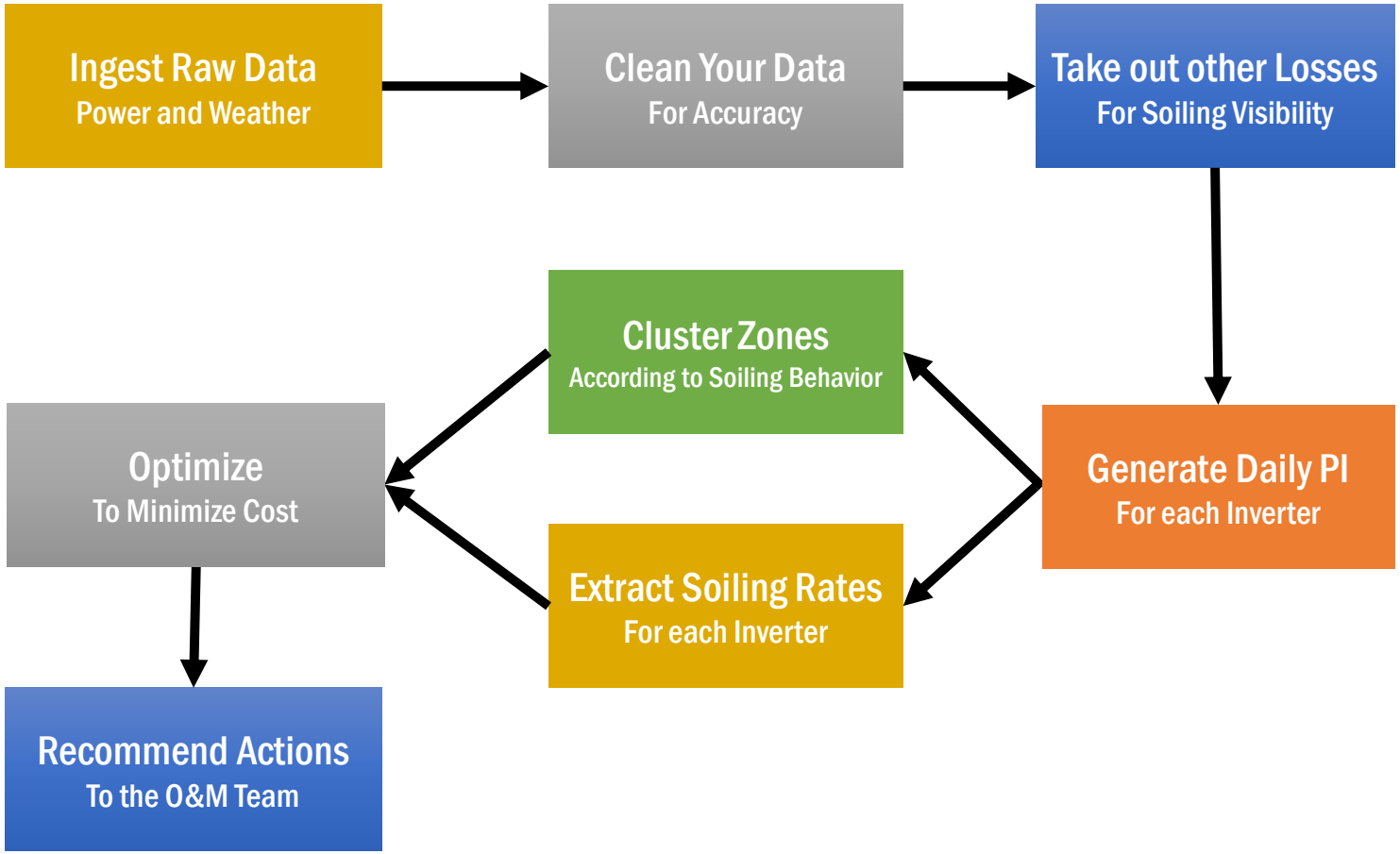


Distribution of Inverter Daily Soiling Rates



Large spread in soiling rates

A Quick Overview of our Soiling Analytics Process



Data Cleaning and Exclusion of Other Losses Before Soiling Analysis

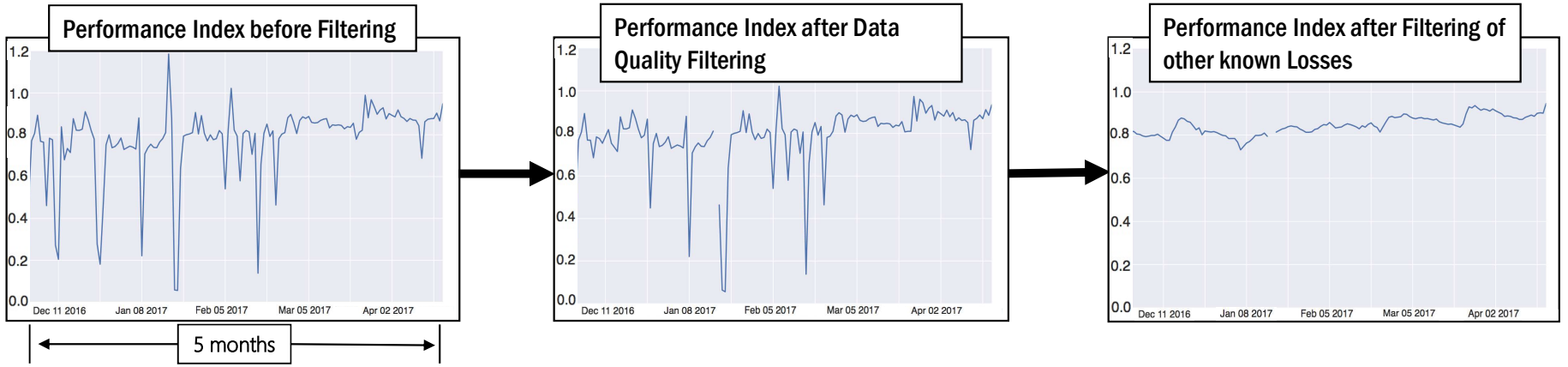
Clean Your Data*
Stuck & out-of-range sensors

→

Take out other Losses
Shade and Snow

→

Generate Daily PI
For each Inverter



- This site experienced outages and stuck sensor values
- There were also other losses:
  - Shading (row-to-row)
  - Shading from chimney,
  - Snow.

- We ended up with a "Soiling Performance Index" which we then used for finding rates and for detecting different soiling zones at the plant.

**K.A. Klise and J.S. Stein (2016), Automated Performance Monitoring for PV Systems using Pecos, 43rd IEEE Photovoltaic Specialists Conference (PVSC), Portland, OR, June 5-10.*

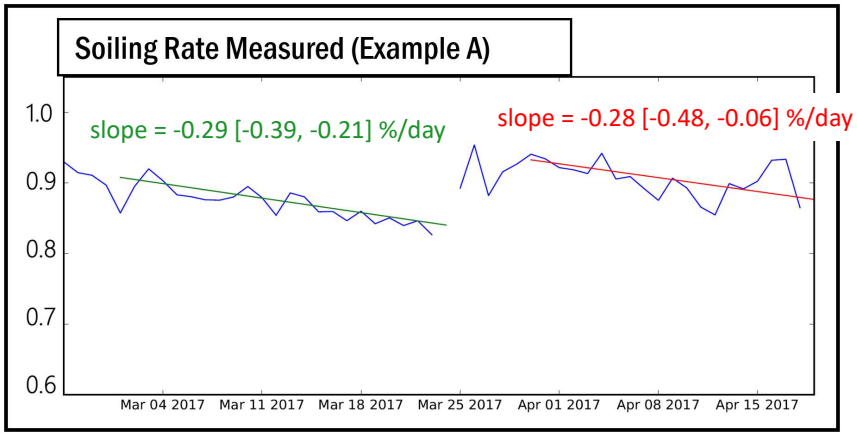
How do we Extract Soiling Rates?

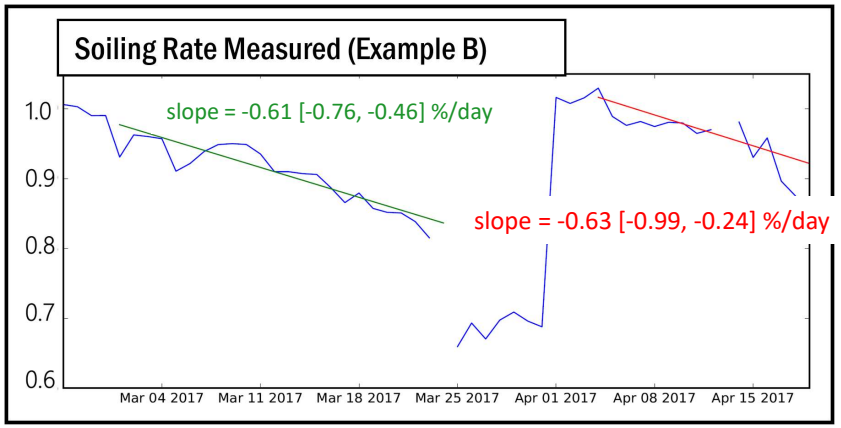
Generate Daily PI
For each Inverter

→

Extract Soiling Rates
For each Inverter

- **Detect soiling periods:** With a combination of rainfall data and a jump detection algorithm
- **Extract rates**: Using a method similar to the Theil-Sen regression technique described in a 2016 NREL paper*
- **Individual rates have high uncertainty.** Assuming groups of similar inverters have rates drawn from the same normal distribution, we can expect ~$1/\sqrt{N}$ lower uncertainty in the average of their soiling rates.
- **Grouping rates is practical** for optimizing O&M wash schedules.





**Michael G. Deceglie, et al., "A Scalable Method for Extracting Soiling Rates from PV Production Data", PVSC 2016*

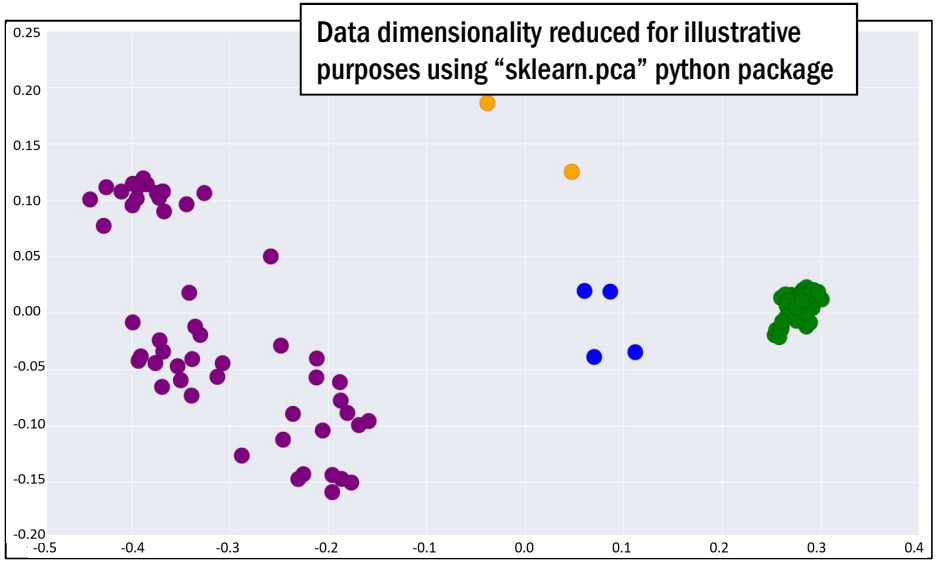
Is there a way to Automatically Group Inverters According to their Soiling Behavior?

Generate Daily PI
For each Inverter

→

Cluster Zones
According to Soiling Behavior

- Clustering is a form of unsupervised Machine Learning: Meaning that there is no training data with correct outputs that are known ahead of time.
- Created a daily "PI Diff" (Performance Index differential) since we are interested in clustering by changes over time
- Calculated the distance between the inverters' daily PI diffs
- Passed these distances into DBSCAN for clustering.
  - DBSCAN is a density based clustering algorithm that is good for handling noise.
  - Number of clusters are not known ahead of time.
  - Tune two other parameters:
    - minimum number of points in a core cluster,
    - minimum distance of points from others in the core cluster



**Using Python packages: numpy, scipy and sklearn*

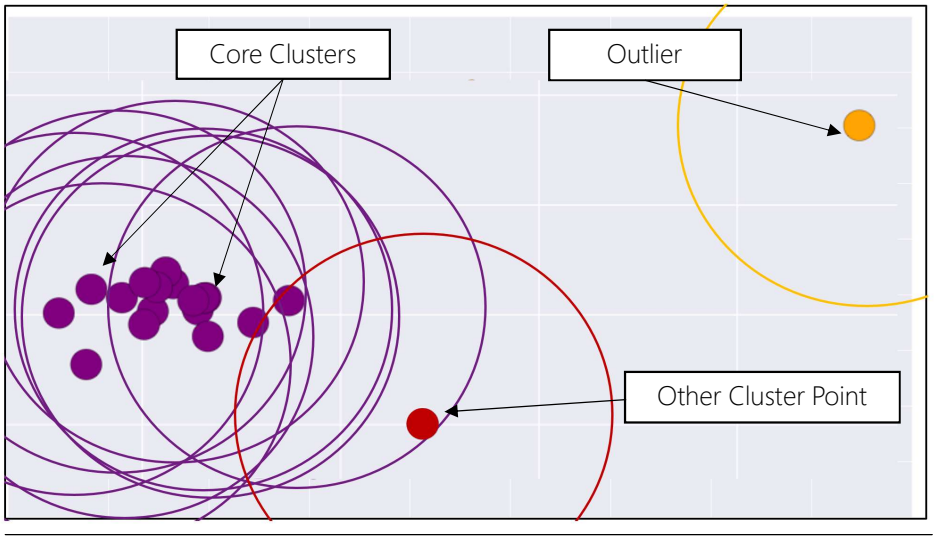
Is there a way to Automatically Group Inverters According to their Soiling Behavior?

Generate Daily PI
For each Inverter

→

Cluster Zones
According to Soiling Behavior

- **Clustering** is a form of unsupervised Machine Learning: Meaning that there is no training data with correct outputs that are known ahead of time.
- **Created a daily "PI Diff"** (Performance Index differential) since we are interested in clustering by changes over time
- **Calculated the distance** between the inverters' daily PI diffs
- Passed these distances into **DBSCAN for clustering**.
  - DBSCAN is a density based clustering algorithm that is good for handling noise.
  - Number of clusters are not known ahead of time.
  - Tune two other parameters:
    - **minimum number of points** in a core cluster,
    - **minimum distance** of points from others in the core cluster



**Using Python packages: numpy, scipy and sklearn*

Zone Rates Provide a Better Estimate

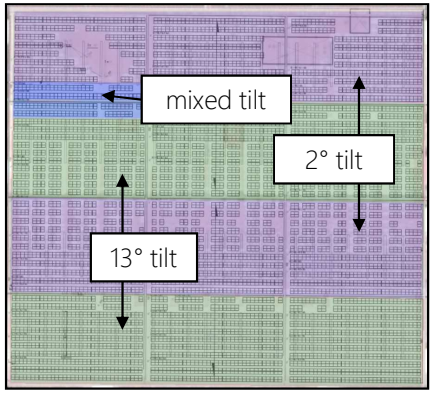
Generate Daily PI
Inverter by Inverter

→

Cluster Zones
According to Soiling Behavior

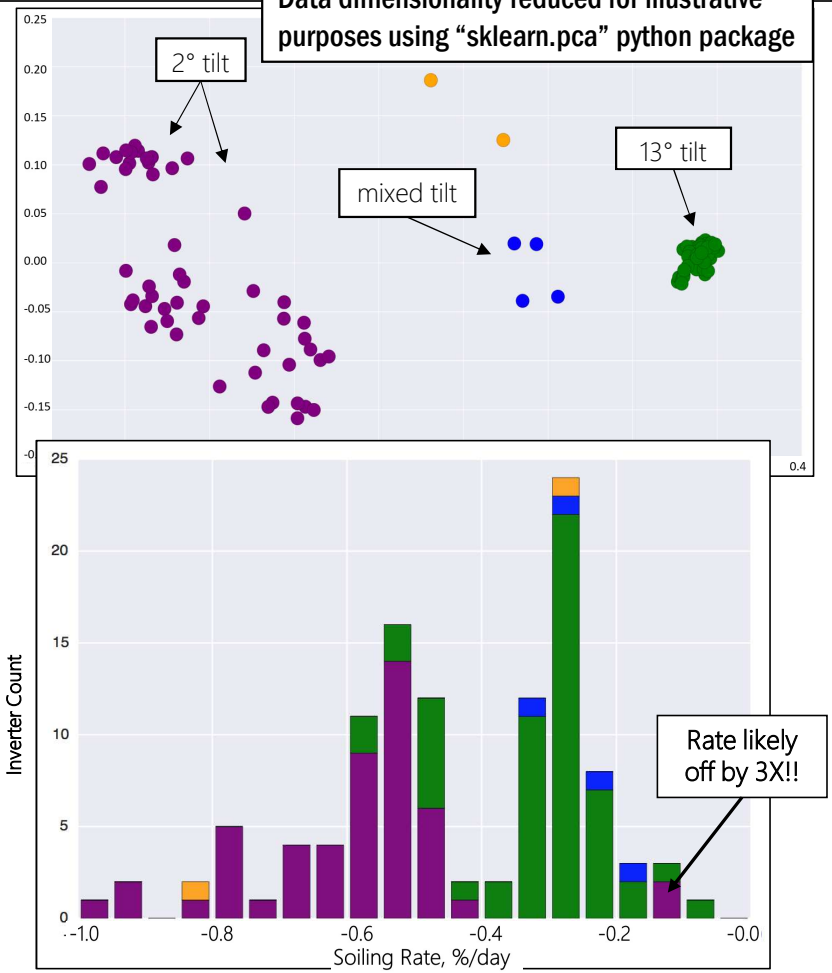
- Clustered into three zones which have a verifiable system design difference

Zone	Soiling rate, %/day
2° tilt	-0.62
13° tilt	-0.28
Mixed tilt	-0.27



- Soiling Rates measured by individual inverters can sometimes be wildly incorrect
- Can apply a rate for inverters where performance index was too noisy to identify a rate at all

Data dimensionality reduced for illustrative purposes using "sklearn.pca" python package



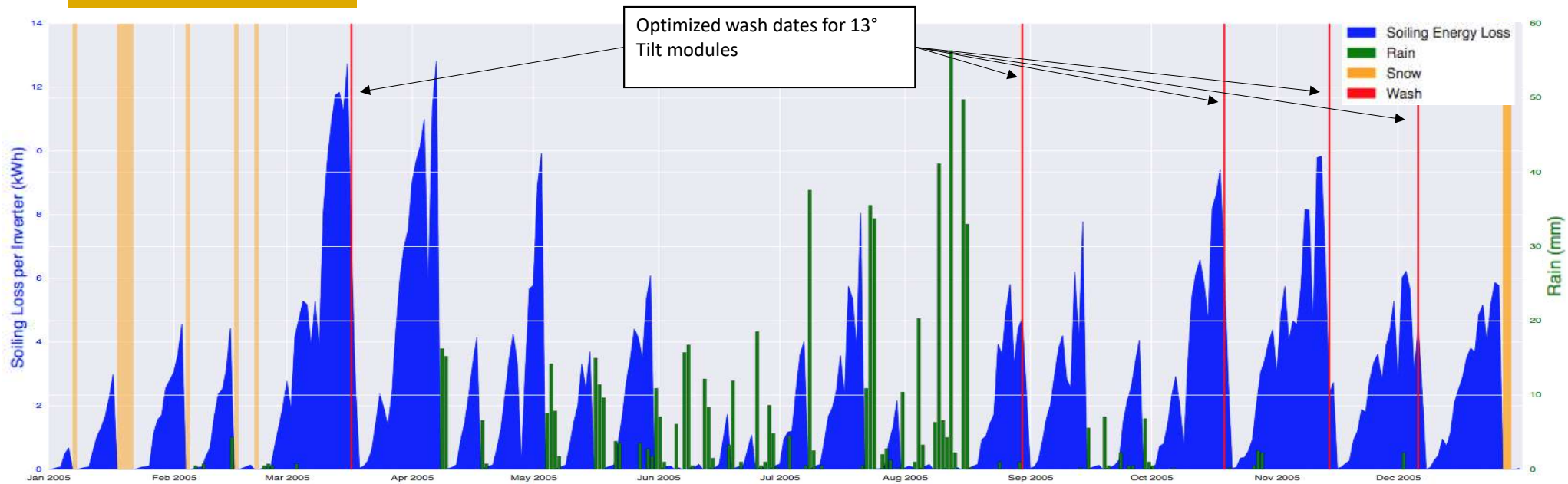
We can recommend an Optimum Wash Schedule for each Zone

Cluster Zones
According to Soiling Behavior

Extract Soiling Rates
For each Inverter

Optimize
To Minimize Cost

- Modeling energy* with a typical weather year
- Assuming that in between rain and snow events there is soiling according to each zone's soiling rate
- Schedule washes to minimize energy loss through the year.



* *https://github.com/pvlib/pvlib-python*

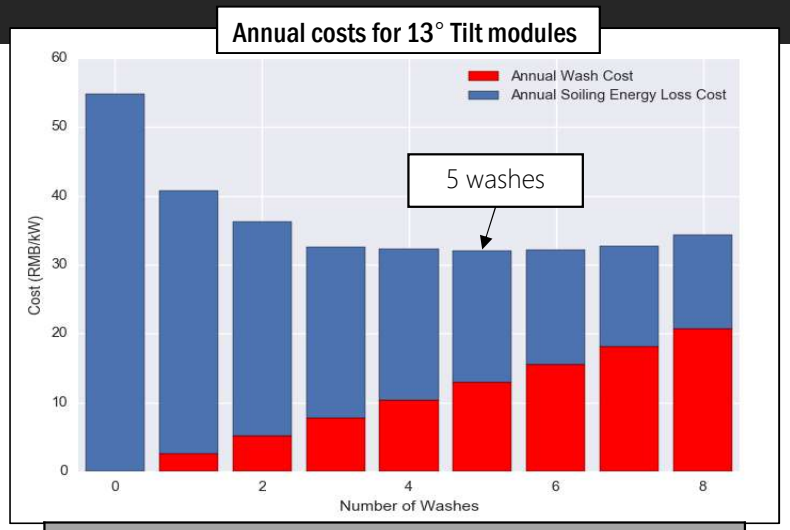
We can recommend Actions

Optimize
To Minimize Cost

→

Recommend Actions
To the O&M Team

- With PPA rates & wash costs we can minimize overall costs due to soiling
  - By recommending an annual wash plan for each zone with typical weather assumptions
  - By alerting in real-time when cost of energy loss due to soiling are approaching wash costs



Zone	Soiling rate, %/day	Washes per year
2° tilt	-0.62	13
13° tilt	-0.28	5
Mixed tilt	-0.27	5

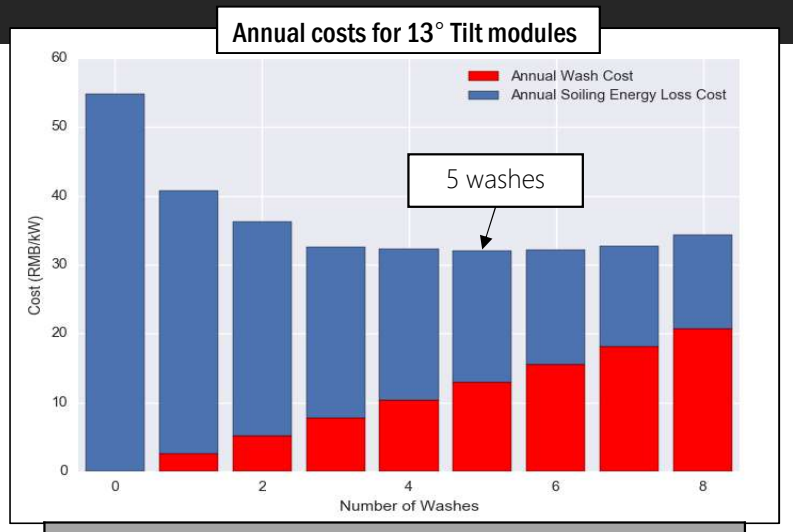
We can recommend Actions

Optimize
To Minimize Cost

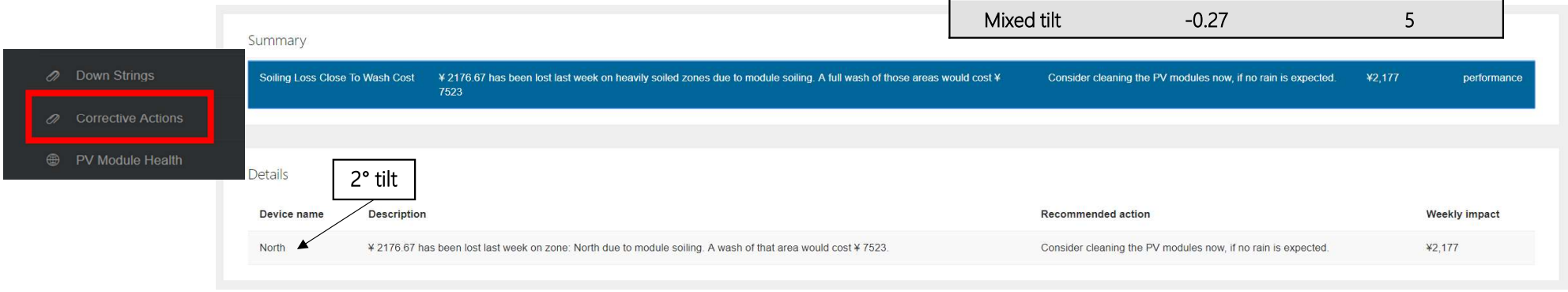
→

Recommend Actions
To the O&M Team

- With PPA rates & wash costs we can minimize overall costs due to soiling
  - By recommending an annual wash plan for each zone with typical weather assumptions
  - By alerting in real-time when cost of energy loss due to soiling are approaching wash costs



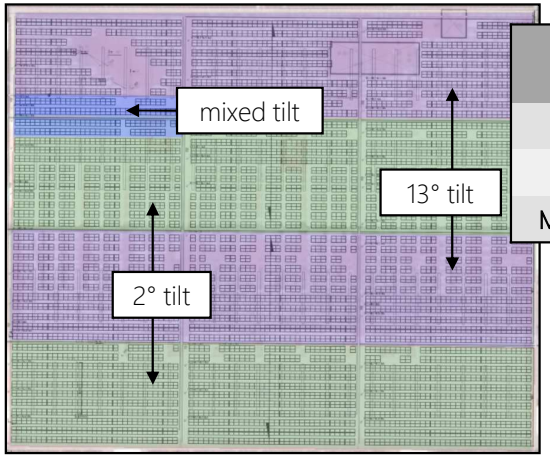
Zone	Soiling rate, %/day	Washes per year
2° tilt	-0.62	13
13° tilt	-0.28	5
Mixed tilt	-0.27	5



Conclusion - Some sites are better described according to multiple soiling zones

Summary:

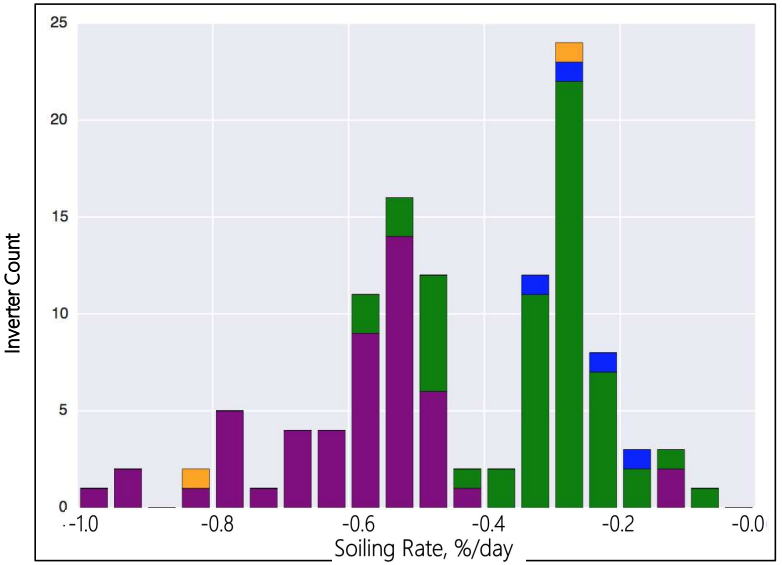
- Using Theil-Sen linear regression we were able to extract soiling rates for each inverter
- Using DBscan clustering algorithm we were able to identify multiple soiling zones at this site.
- Average zone soiling rates are more accurate and practical for making decisions
- Using these rates in an energy model along with PPA rates and cost of washing, we can recommend an optimum wash schedule for each zone & give actions in real time.



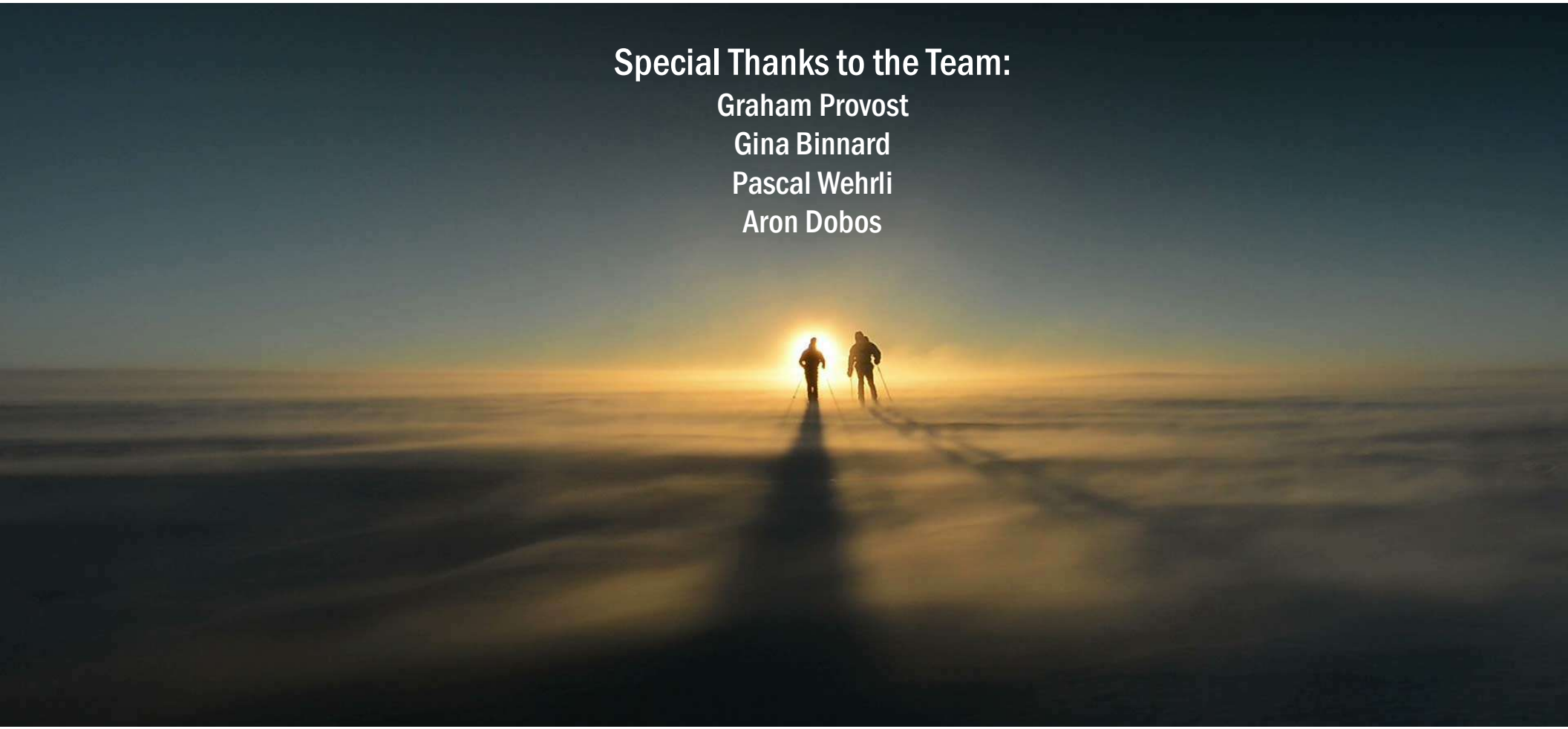
Zone	Soiling rate, %/day	Washes per year
2° tilt	-0.62	13
13° tilt	-0.28	5
Mixed tilt	-0.27	5

Next Steps:

- Test on other heavily soiled sites
- Work with customers to define actions that would be applicable to them in the field
- Automate the selection of the DBSCAN parameters



ENVISION



Special Thanks to the Team:

Graham Provost
Gina Binnard
Pascal Wehrli
Aron Dobos