



# PV performance analysis:

a status update of the Pearl PV data challenge

**Dr. Eli Shirazi**

University of Twente, Netherlands

# Content

- ❖ PEARL PV CKAN Database DEMO
- ❖ PEARL PV Data Challenge
- ❖ Preparing datasets
- ❖ A status update of the data challenge

# ≡ PEARL PV CKAN Database DEMO

- ≡ CKAN = Comprehensive Knowledge Archive Network
- ≡ Open source Open Data storage + distribution
- ≡ PEARL PV CKAN is a customized version
- ≡ Dedicated to PV system + meteo-data
- ≡ Hosted by University of Twente, Enschede, NL
- ≡ 20 years
- ≡ <https://ckan.pearlpv-cost.eu/>

# PEARL PV CKAN

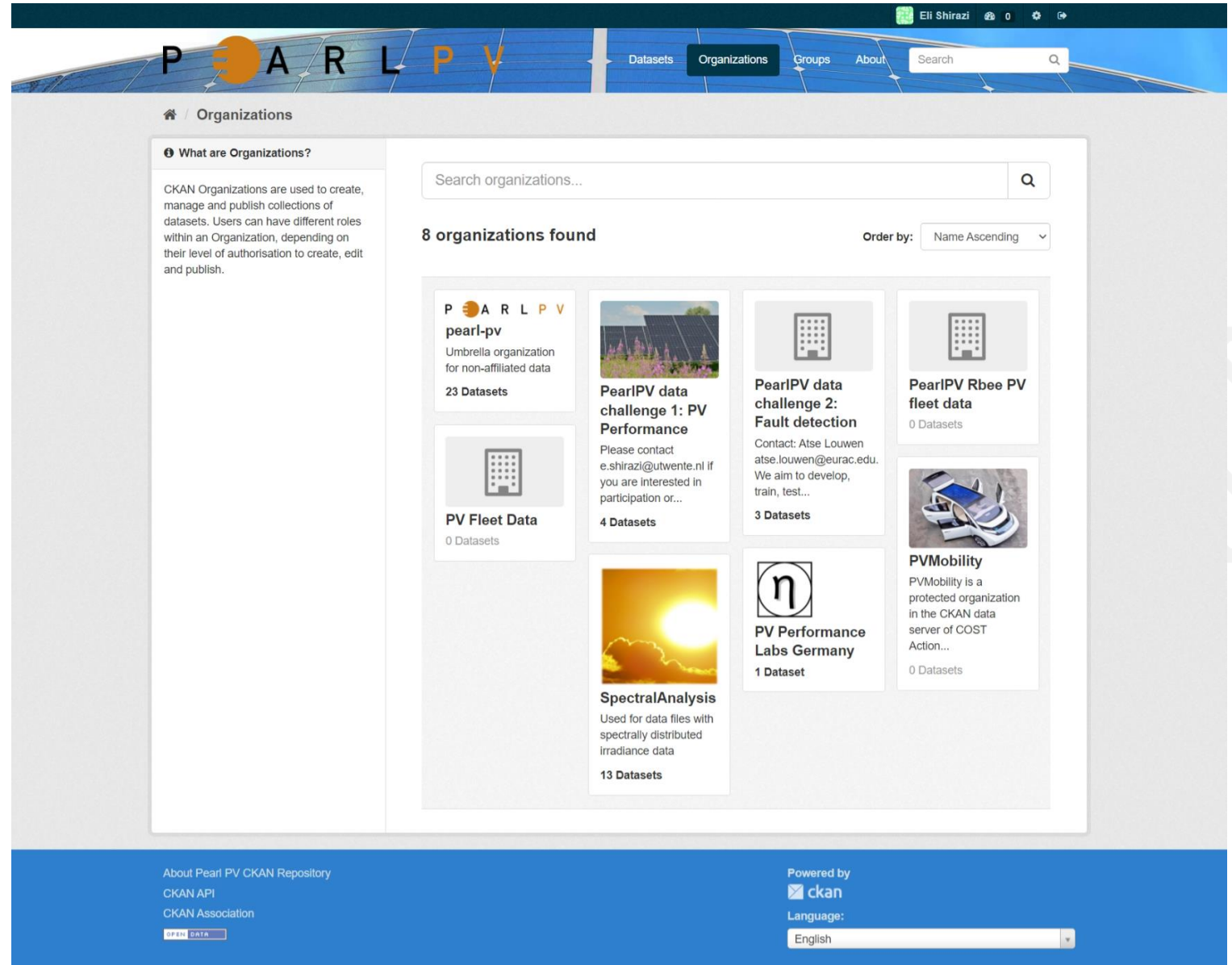
- ≡ Datasets
- ≡ Organizations
- ≡ Groups

4 completely public datasets  
(more inside when you log in!)

The screenshot shows the Pearl PV CKAN Repository website. At the top, there is a navigation bar with links for Datasets, Organizations, Groups, and About, along with a search bar. Below the navigation bar, there is a 'Search data' section with a search input field containing 'E.g. environment' and a search button. Below the search bar, there are 'Popular tags' including APE, Meteorological data, and rooftop. A green arrow points from the text box on the left to the 'Pearl PV CKAN Repository statistics' section, which displays '4 datasets', '5 organizations', and '1 group'. To the right of the search bar, there is a 'Welcome to the Pearl PV CKAN Repository' section with a message about the website's development and a 'Quicklinks to help' section with links to 'Add a dataset' and 'Edit a dataset'. Below the welcome section, there is a 'Spectral irradiance' section with a description of measured spectral irradiance in different locations. At the bottom, there is a 'pearl-pv' section with a description of the umbrella organization for non-affiliated data, and a 'Spectral data 2020.06.15-2021.05.31 - HZB' section with a description of the dataset. The footer contains links to 'About Pearl PV CKAN Repository', 'CKAN API', and 'CKAN Association', along with a 'Powered by ckan' logo and a 'Language' dropdown menu set to 'English'.

# Organizations

- Collection of datasets and users
- Users can have different roles in the organization
  - Member
  - Editor
  - Admin



The screenshot shows the PearlPV CKAN Organizations page. The header features the PearlPV logo and navigation links for Datasets, Organizations, Groups, and About, along with a search bar. The main content area is titled 'Organizations' and includes a sidebar with a 'What are Organizations?' section explaining CKAN Organizations. The main area displays a search bar, a list of 8 organizations found, and a dropdown for 'Order by: Name Ascending'. The organizations listed are:

- pearl-pv**: Umbrella organization for non-affiliated data, 23 Datasets.
- PV Fleet Data**: 0 Datasets.
- PearlPV data challenge 1: PV Performance**: Please contact e.shirazi@utwente.nl if you are interested in participation or..., 4 Datasets.
- PearlPV data challenge 2: Fault detection**: Contact: Atse Louwen atse.louwen@eurac.edu. We aim to develop, train, test..., 3 Datasets.
- PV Performance Labs Germany**: 1 Dataset.
- PVMobility**: PVMobility is a protected organization in the CKAN data server of COST Action..., 0 Datasets.
- SpectralAnalysis**: Used for data files with spectrally distributed irradiance data, 13 Datasets.

The footer contains links for 'About Pearl PV CKAN Repository', 'CKAN API', 'CKAN Association', and 'OPEN DATA'. It also indicates 'Powered by ckan' and a 'Language: English' dropdown menu.

# Registering

- Straightforward
- NDA has been created to ensure users know how to use the database
- E.g.: don't share or publish data without consent of owner

Log in Register

PEARL PV

Datasets Organizations Groups About Search

Registration

**Why Sign Up?**

Create datasets, groups and other exciting things

**Register for an Account**

\* Username:

Full Name:

\* Email:

\* Password:

\* Confirm:

Please read the [NDA](#) and check if agree:

☐ I Accept.

\* Required field

Create Account

About Pearl PV CKAN Repository  
CKAN API  
CKAN Association  
[OPEN DATA](#)

Powered by ckan  
Language:

# ≡ Uploading Data

- ≡ Upload forms customized to allow for detailed description of PV system data.

The screenshot displays the PEARL PV website interface. At the top, the navigation bar includes the PEARL PV logo, links for Datasets, Organizations, Groups, and About, and a search bar. The user profile 'Eli Shirazi' is visible in the top right corner. The main content area shows the 'pearl-pv' organization profile, which includes a description, follower count (11), dataset count (23), and an 'Unfollow' button. A sidebar on the left lists various categories: Organizations (pearl-pv), Groups (Spectral irradiance), and Tags (rooftop, no faults, PV data, Berlin, C-Si). The main content area also features a search bar, an 'Add Dataset' button (circled in green), and a list of datasets. The first dataset listed is 'SUPSI PVLab Solarbrick data', which is private and contains 4 years of 5-minute data. The second dataset is 'Kristiansand PV & meteo 1-min data per month, 2012-2018', also private, with 10-minute resolution data. The third dataset is 'PV Data in Pfaffstaetten/Austria (Karl Berger)', private, with 10-minute resolution data. The 'Add Dataset' button is a blue button with a plus icon and the text 'Add Dataset'.



# ≡ Uploading Data

Many fields that aid in exchange:

- ≡ Tags
- ≡ License
- ≡ Location
- ≡ Author (owner) and maintainer (contact) of data
- ≡ System metadata
- ≡ Description of dataset
- ≡ Dataset visibility!

The screenshot shows the 'Create Dataset' page of the PEARL PV application. The header features the PEARL PV logo and navigation links for Datasets, Organizations, Groups, and About, along with a search bar. The user 'Eli Shirazi' is logged in. The main content area is divided into two columns. The left column contains a 'What are datasets?' section with a brief explanation of CKAN datasets. The right column contains the 'Create Dataset' form, which includes a progress bar with two steps: '1 Create dataset' (active) and '2 Add data'. The form fields are: 'Title' (with a placeholder 'eg. Berlin PV data by months'), 'URL' (with a placeholder 'ckan.pearl-pv-cost.eu/dataset/<dataset>' and an 'Edit' button), 'Description' (with a placeholder 'eg. add any additional information that you consider useful to people who use the data' and a note about Markdown formatting), 'Tags' (with a placeholder 'eg. rooftop, CIGS, Berlin. Hit the Enter button between tags.'), 'License' (a dropdown menu currently showing 'License not specified'), and 'Location name' (partially visible at the bottom). A link to 'http://opendefinition.org/' is provided for license definitions.



# ≡ Uploading Data

- ≡ Many file formats supported
- ≡ Multiple files per dataset possible
- ≡ Additional field to describe the file structure, data formats etc.

The screenshot displays the 'Create Dataset' page of the PEARL PV application. The header features the PEARL PV logo and navigation links for Datasets, Organizations, Groups, and About, along with a search bar. The user 'Eli Shirazi' is logged in. The main content area is divided into two sections. On the left, a sidebar titled 'What are datasets?' explains that a CKAN Dataset is a collection of data resources with a description and a fixed URL. On the right, the 'Create Dataset' form is shown, with a progress bar indicating two steps: '1 Create dataset' (active) and '2 Add data'. The form includes a 'Title' field with the example 'eg. Berlin PV data by months', a 'URL' field with the example 'ckan.pearlpv-cost.eu/dataset/<dataset>' and an 'Edit' button, a 'Description' field with the example 'eg. add any additional information that you consider useful to people who use the data' and a note about Markdown formatting, and a 'Tags' field with the example 'eg. rooftop, CIGS, Berlin. Hit the Enter button between tags.'

# Viewing and Downloading Data

- As easy as finding and selecting a dataset

The screenshot displays the PEARL PV Datasets website interface. The top navigation bar includes the PEARL PV logo, a search bar, and links for Datasets, Organizations, Groups, and About. The main content area is titled 'Datasets' and features a sidebar with filters for Organizations, Groups, and Tags. The main panel shows a search bar with the text 'Search datasets...', a '44 datasets found' result, and a list of datasets. The first dataset listed is 'PV fault dataset - Brazil', marked as 'PRIVATE', with a description: 'The following dataset was used in the paper submitted to Sensors MDPI: Monitoring System for Online Fault Detection and Classification in Photovoltaic Plants by André E....'. The second dataset is 'Bern BIPV string data', also marked as 'PRIVATE', with a description: 'Data from one string of a BIPV system in the Bern area (Switzerland). The string size is 5.23 kWp. Low amounts of near shading from nearby tree and chimney on rooftop. Include...'. A 'CSV' download button is visible next to this dataset. The third dataset is 'Eurac PV power and solar irradiance data', marked as 'PRIVATE', with a description: 'Two years of 1-hour data from Eurac's test facility in Bolzano, for a polycrystalline silicon based systems, including GPOA and GHI'.

# ❖ PEARL PV Data Challenge

- ❖ PEARL PV Data Challenge 1: Comparison of PV performance across Europe
- ❖ PEARL PV Data Challenge 2: Artificially Faulty Solar PV Datasets for Development and Benchmarking of Fault Detection Algorithms

# ❖ Data Challenge 1: comparison of PV performance across Europe

## ❖ Aim:

- ❖ to have a harmonized, high-coverage, open dataset of solar photovoltaic installations across EU
- ❖ investigate performance loss based on location, climate and PV technology
- ❖ By contributing datasets, participants gain access to the datasets of other participants, and can join in a Europe wide intercomparison of performance and degradation of PV systems.
- ❖ The results will be published in a joint paper
- ❖ [PearlPV data challenge 1: PV Performance - Organizations - Pearl PV CKAN Repository \(pearlpv-cost.eu\)](https://pearl-pv.eu/data-challenge-1-pv-performance-organizations)
- ❖ **Please contact [e.shirazi@utwente.nl](mailto:e.shirazi@utwente.nl) if you are interested!**

# ❖ Data Challenge 1: Required Data

PV system metadata
GPS Coordinates
Fixed installation, tilt and azimuth
Tracking Mode
PV module technology
Total number of PV modules in the system
Number of modules connected in one string
Number of strings connected to each inverter
Total number of PV strings in the system
Soiling
Shading

Variables
Resolution: $\leq 1$ hour
Horizon: $> 1$ year
plane-of-array (POA) irradiance
DC output power
module temperature
global horizontal irradiance (GHI)
AC output power

## ❖ Data Challenge 2: Fault Detection

- ❖ Aim: development and comparison of fault detection algorithms
- ❖ By sharing datasets where participants introduce and label faults procedurally/manually
  - ❖ Room for discussion on how exactly
  - ❖ Multiple methods are fine (e.g. also datasets with actual labelled faults, or faults physically introduced by disconnecting)
- ❖ Ambition is to publish the results in a joint paper

<https://ckan.pearlpv-cost.eu/organization/about/pearlpv-data-challenge-2-fault-detection>

# ❖ Data Challenge 2: Fault Detection

- ❖ Progress so far:
  - ❖ 7 participants
  - ❖ 2 datasets already uploaded
  - ❖ At least 3-4 more pledged
- ❖ Due to busy period with many proposals original timeline was too ambitious
- ❖ For the journal paper we would like to have a larger group
- ❖ **Please contact [atse.louwen@eurac.edu](mailto:atse.louwen@eurac.edu) if you are interested!**

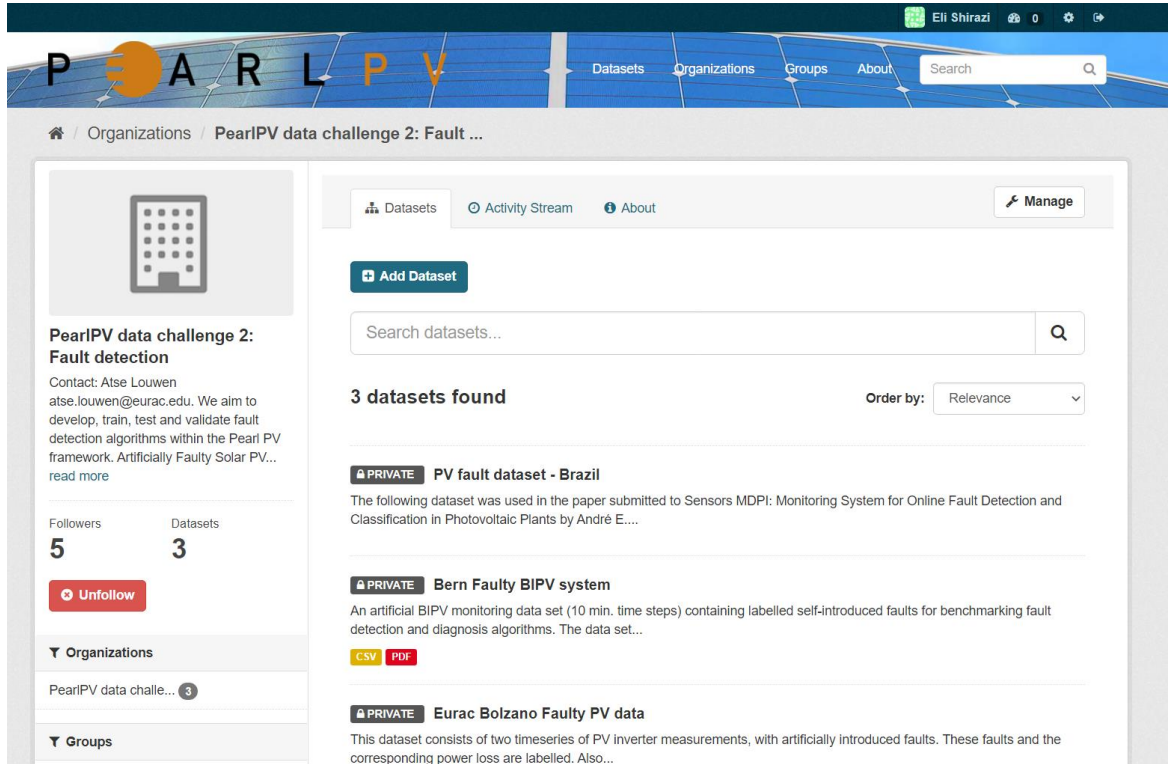


# CKAN Data server

The screenshot displays the CKAN Data server interface for the 'Organizations' section. The header features the 'PEARLPV' logo and navigation links for 'Datasets', 'Organizations', 'Groups', and 'About'. A search bar is located in the top right corner. The main content area shows a list of 8 organizations found, ordered by name ascending. The organizations are:

- pearl-pv**: Umbrella organization for non-affiliated data. 23 Datasets.
- PV Fleet Data**: 0 Datasets.
- PearlPV data challenge 1: PV Performance**: Please contact e.shirazi@utwente.nl if you are interested in participation or... 4 Datasets. (Highlighted with a red box)
- PearlPV data challenge 2: Fault detection**: Contact: Atse Louwen atse.louwen@eurac.edu. We aim to develop, train, test... 3 Datasets.
- PearlPV Rbee PV fleet data**: 0 Datasets.
- PVMobility**: PVMobility is a protected organization in the CKAN data server of COST Action... 0 Datasets.
- PV Performance Labs Germany**: 1 Dataset.
- SpectralAnalysis**: Used for data files with spectrally distributed irradiance data. 13 Datasets.

# Data Challenges



The screenshot shows the PearlPV data challenge 2: Fault detection page. The header includes the PearlPV logo and navigation links for Datasets, Organizations, Groups, and About. A search bar is present. The main content area has tabs for Datasets, Activity Stream, and About. A 'Manage' button is in the top right. Below the tabs is an 'Add Dataset' button and a search bar. The search results show 3 datasets found, ordered by Relevance. The datasets listed are: 'PV fault dataset - Brazil' (PRIVATE), 'Bern Faulty BIPV system' (PRIVATE), and 'Eurac Bolzano Faulty PV data' (PRIVATE). Each dataset entry includes a brief description and file format icons (CSV, PDF).

PearlPV data challenge 2: Fault detection

Contact: Atse Louwen  
atse.louwen@eurac.edu. We aim to develop, train, test and validate fault detection algorithms within the Pearl PV framework. Artificially Faulty Solar PV...  
[read more](#)

Followers: 5 Datasets: 3

[Unfollow](#)

Organizations

PearlPV data challe... 3

Groups

**3 datasets found** Order by: Relevance

**PRIVATE** PV fault dataset - Brazil

The following dataset was used in the paper submitted to Sensors MDPI: Monitoring System for Online Fault Detection and Classification in Photovoltaic Plants by André E....

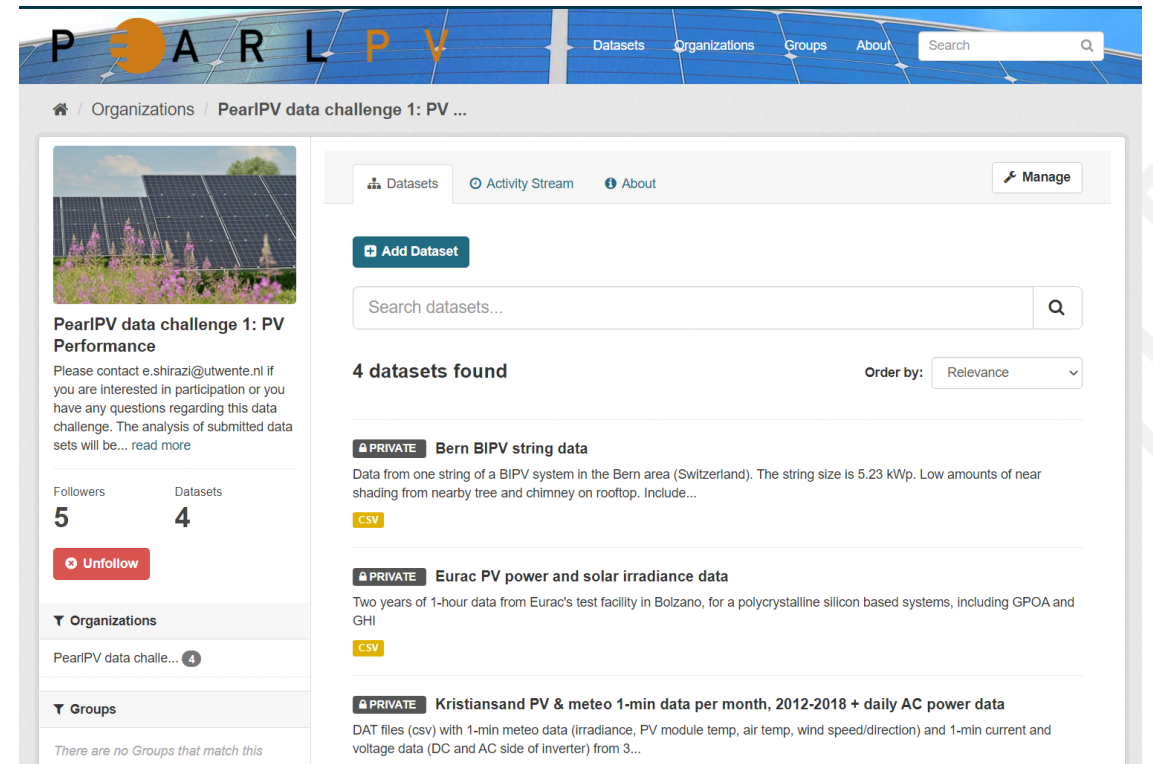
**PRIVATE** Bern Faulty BIPV system

An artificial BIPV monitoring data set (10 min. time steps) containing labelled self-introduced faults for benchmarking fault detection and diagnosis algorithms. The data set...

CSV PDF

**PRIVATE** Eurac Bolzano Faulty PV data

This dataset consists of two timeseries of PV inverter measurements, with artificially introduced faults. These faults and the corresponding power loss are labelled. Also...



The screenshot shows the PearlPV data challenge 1: PV Performance page. The header includes the PearlPV logo and navigation links for Datasets, Organizations, Groups, and About. A search bar is present. The main content area has tabs for Datasets, Activity Stream, and About. A 'Manage' button is in the top right. Below the tabs is an 'Add Dataset' button and a search bar. The search results show 4 datasets found, ordered by Relevance. The datasets listed are: 'Bern BIPV string data' (PRIVATE), 'Eurac PV power and solar irradiance data' (PRIVATE), and 'Kristiansand PV & meteo 1-min data per month, 2012-2018 + daily AC power data' (PRIVATE). Each dataset entry includes a brief description and file format icons (CSV).

PearlPV data challenge 1: PV Performance

Please contact e.shirazi@utwente.nl if you are interested in participation or you have any questions regarding this data challenge. The analysis of submitted data sets will be... [read more](#)

Followers: 5 Datasets: 4

[Unfollow](#)

Organizations

PearlPV data challe... 4

Groups

**4 datasets found** Order by: Relevance

**PRIVATE** Bern BIPV string data

Data from one string of a BIPV system in the Bern area (Switzerland). The string size is 5.23 kWp. Low amounts of near shading from nearby tree and chimney on rooftop. Include...

CSV

**PRIVATE** Eurac PV power and solar irradiance data

Two years of 1-hour data from Eurac's test facility in Bolzano, for a polycrystalline silicon based systems, including GPOA and GHI

CSV

**PRIVATE** Kristiansand PV & meteo 1-min data per month, 2012-2018 + daily AC power data

DAT files (csv) with 1-min meteo data (irradiance, PV module temp, air temp, wind speed/direction) and 1-min current and voltage data (DC and AC side of inverter) from 3...

[PearlPV data challenge 1: PV Performance - Organizations - Pearl PV CKAN Repository \(pearl-pv-cost.eu\)](https://pearl-pv-cost.eu/organizations/pearl-pv-data-challenge-1-pv-performance/)

[PearlPV data challenge 2: Fault detection - Organizations - Pearl PV CKAN Repository \(pearl-pv-cost.eu\)](https://pearl-pv-cost.eu/organizations/pearl-pv-data-challenge-2-fault-detection/)

# Preparing Datasets

## Data Cleaning

- Outliers, missing datapoints

## Data Filtering

Filtering serves to identify and remove data within the time series that are influenced by factors that cannot be modeled

- Irradiance threshold filtering
- Power threshold filtering
- Inverter saturation and curtailment
- Clear-sky filters
- Shading, soiling, and snow corrections

## Cleaning Datasets

Data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, or incomplete data within a dataset.

- ≡ Handle missing values
- ≡ Filter unwanted outliers
- ≡ Data normalization



Image courtesy: [www.finereport.com](http://www.finereport.com)

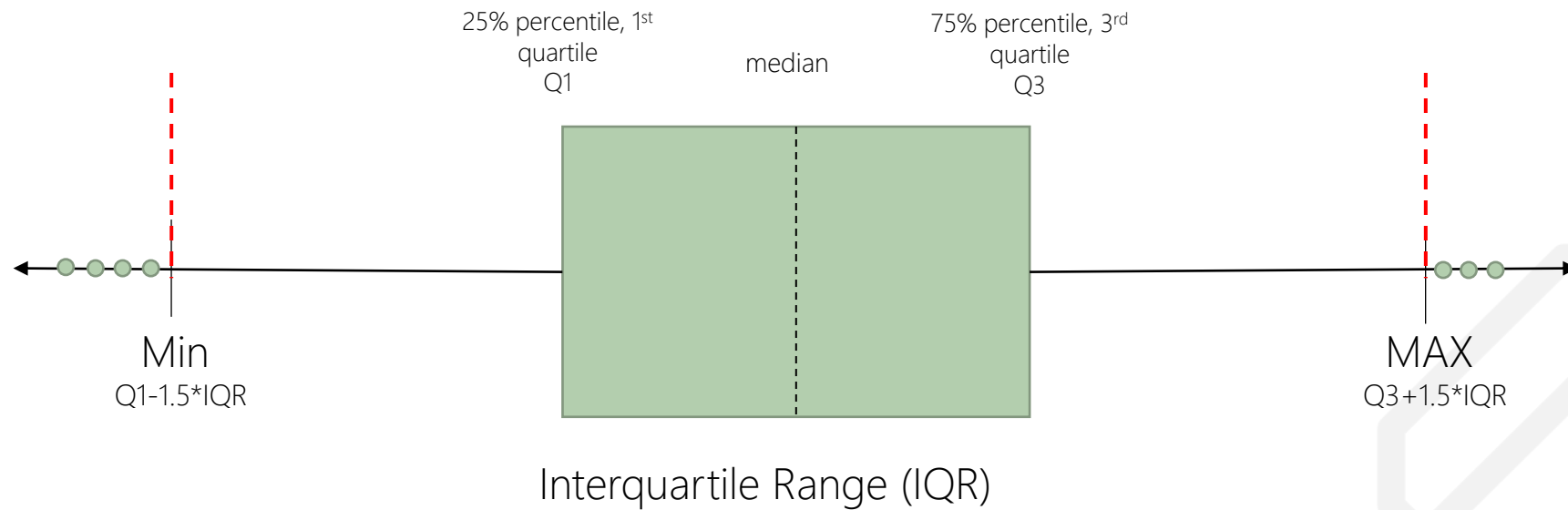


## ❖ Cleaning Datasets – Handling missing values

---

- ❖ Two main strategies to handle missing values in a dataset:
  - ❖ Drop observations that have missing values
  - ❖ Input missing values based on other observations

# Cleaning Datasets – Filter Unwanted Outliers







## ❖ Cleaning Datasets – Data Normalization

---

Common normalization techniques:

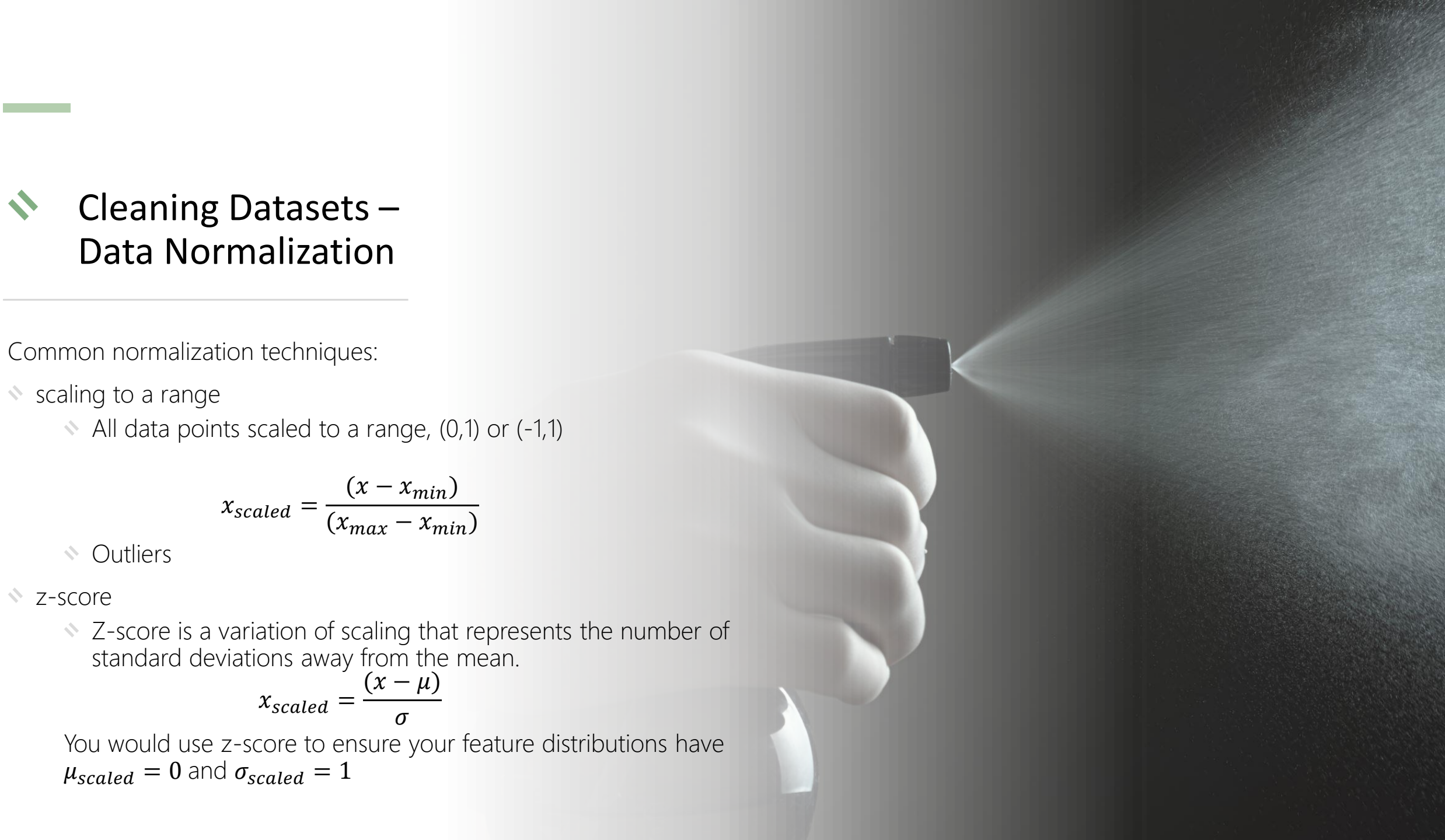
- ❖ scaling to a range
  - ❖ All data points scaled to a range, (0,1) or (-1,1)

$$x_{scaled} = \frac{(x - x_{min})}{(x_{max} - x_{min})}$$

- ❖ Outliers
- ❖ z-score
  - ❖ Z-score is a variation of scaling that represents the number of standard deviations away from the mean.

$$x_{scaled} = \frac{(x - \mu)}{\sigma}$$

You would use z-score to ensure your feature distributions have  $\mu_{scaled} = 0$  and  $\sigma_{scaled} = 1$





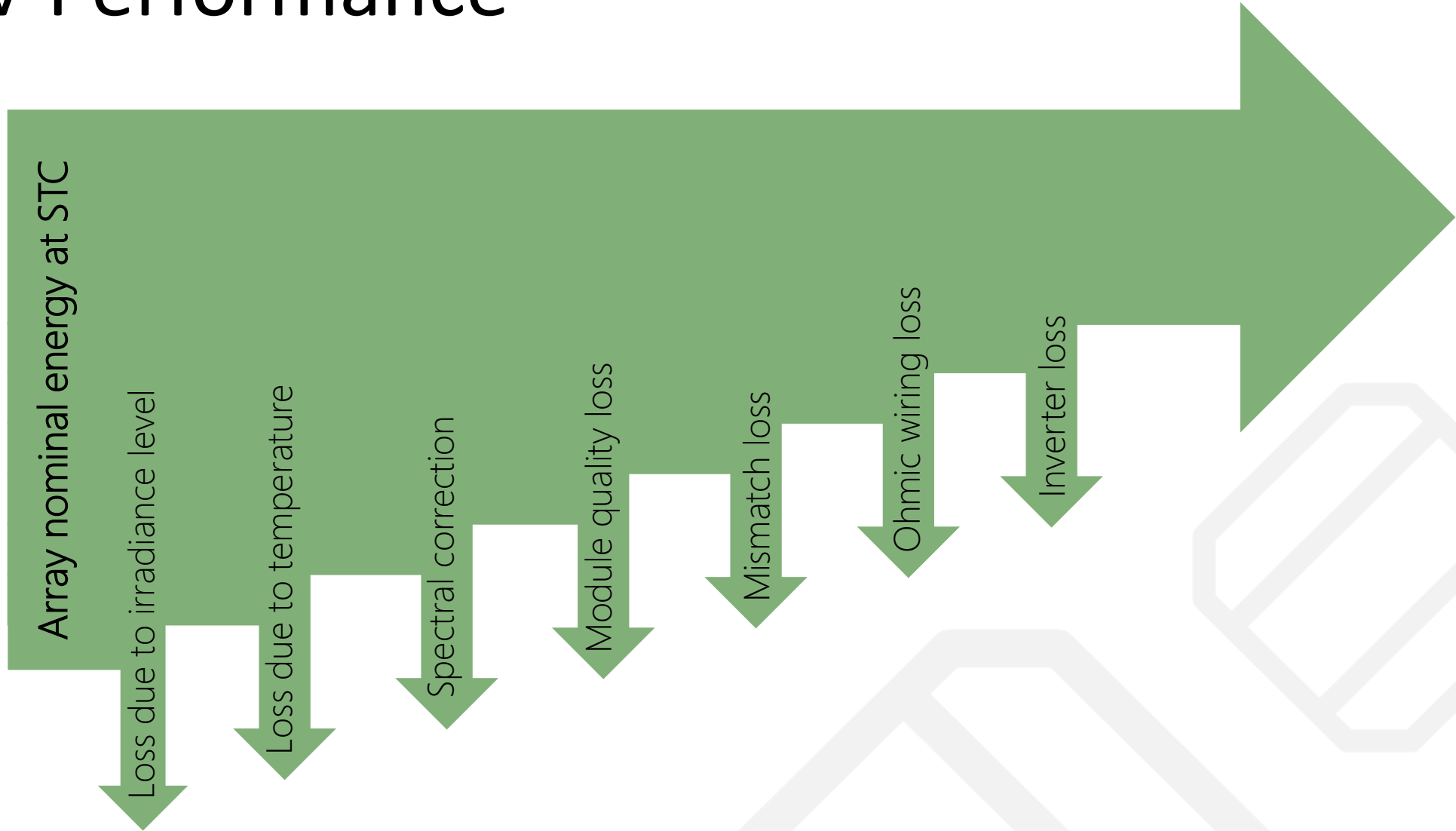


# Why PV system Performance?

---

- ❖ photovoltaic (PV) system's performance is essential to:
- ❖ evaluate if a system is operating within the boundaries of long-term yield assessments and warranties
- ❖ provide more accurate values to be used in yield assessments not only in terms of absolute value but also in terms of uncertainty.

# ≡ PV Performance



# KPI

Following key performance indicators (KPI's) are used:

- ❖ The **final yield**  $Y_f$ , is defined as the energy output ( $E_{AC}$ ,  $E_{DC}$  the total AC or DC power produced by the solar PV system over a defined period) per the system's rated capacity ( $P_{rated}$ ).

$$Y_f = \frac{E_{AC}}{P_{rated}}, Y_f = \frac{E_{DC}}{P_{rated}}$$

- ❖ The **reference yield**, is defined as the solar radiation at the plane of array ( $H_{POA}$ ) divided by the irradiance at the standard test conditions ( $G_{STC} = 1 \text{ kW/m}^2$ ).

$$Y_{ref} = \frac{H_{POA}}{G_{STC}}$$

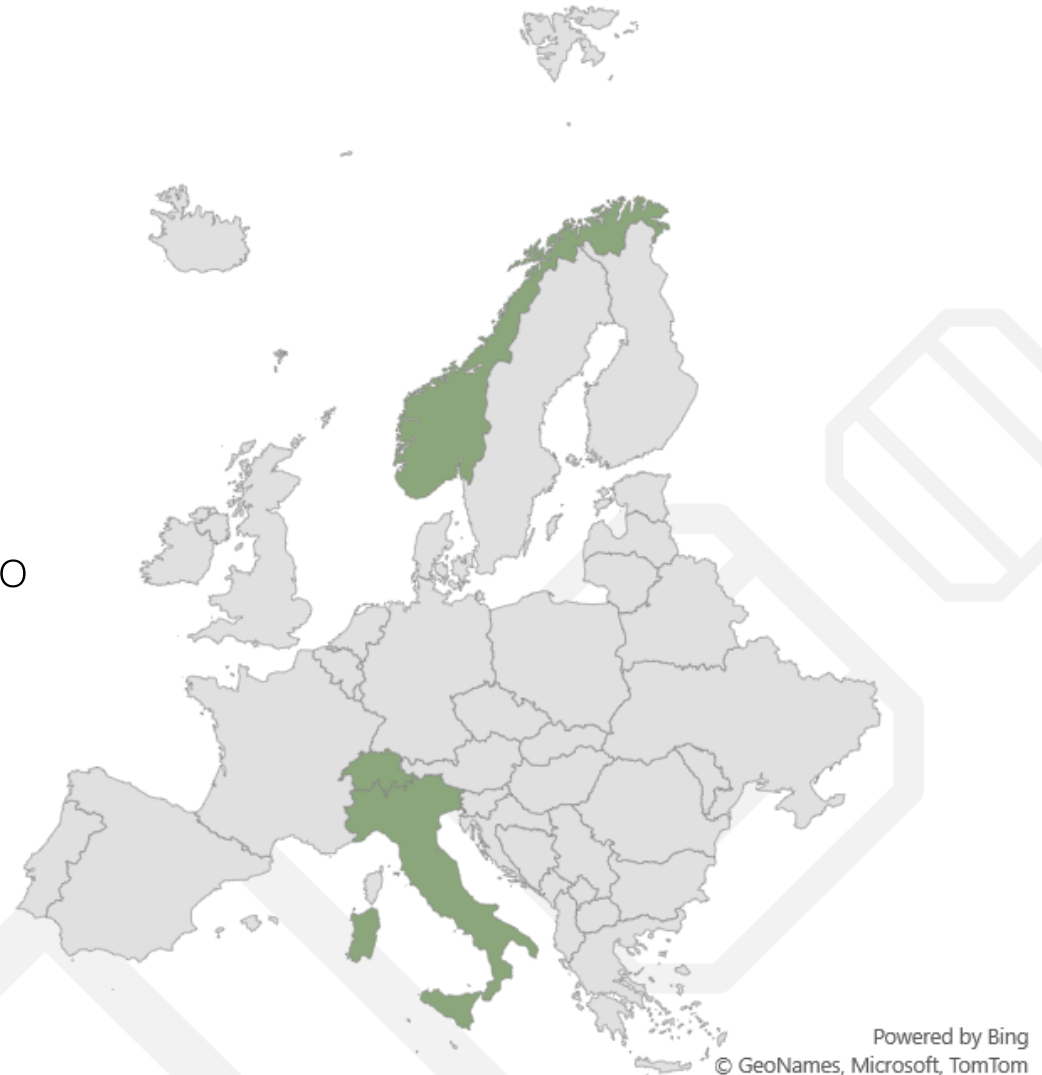
- ❖ The **performance ratio** (PR):

$$PR = \frac{Y_f}{Y_{ref}} = \frac{E_{AC}/P_{rated}}{H_{POA}/G_{STC}}$$

# Contributing Countries

## Datasets:

- Dataset 1: SUPSI PVLab Solarbrick data (Switzerland)
- Dataset 2: University of Agder, Kristiansand PV & meteo data, (Norway)
- Dataset 3: Eurac PV power and solar irradiance data (Italy)
- Dataset 4: Bern BIPV string data (Switzerland)



# Data Sets

Dataset	Parameters	Resolution	Duration	Module Level	System Level	Test system	
SUPSI	GHI, GPOA and module temperature current and voltage	5 minutes	+4 years	✓	✗	crystalline-Si	
						Glass/EVA/BS Glass/PVB/Glass	Open-rack Insulated (no air gap)
Kristiansand, University of Agder	irradiance, PV module temp, air temp, wind speed/direction current and voltage	1 minute	+6 years	✓	✓	BAPV flat roof	
						Multi crystalline-Si, Micromorph-Si	
EURAC	GPOA and GHI Module temperature DC and AC Power	1 hour	2 years	✓	✗	polycrystalline silicon	
						free-standing installation	
Bern	GHI/DHI/DNI/ POA irradiance, Air temperature and wind speed DC power/ voltage/ current	1 hour	+6 years	✓	✗	one string of a BIPV system	



# ❖ Kristiansand Dataset

## ❖ Inverters:

- ❖ Suntech = multi-Si, 24 modules of 225 Wp Suntech STP225-20/Wd (eff STC=13.6 %) from Elkem Solar Silicon multi-Si material (5.4 kWp).
- ❖ REC = multi-Si. (166 modules of 225 Wp REC PE-series (eff STC=13.6 %) made from standard multi-Si material (37.4 kWp).)
- ❖ Sharp = thin film micromorphous-Si, 18 modules of 135 Wp Sharp NA-series (eff STC=9.5 %) made from a-Si/micro-Si silicon material (2.4 kWp).







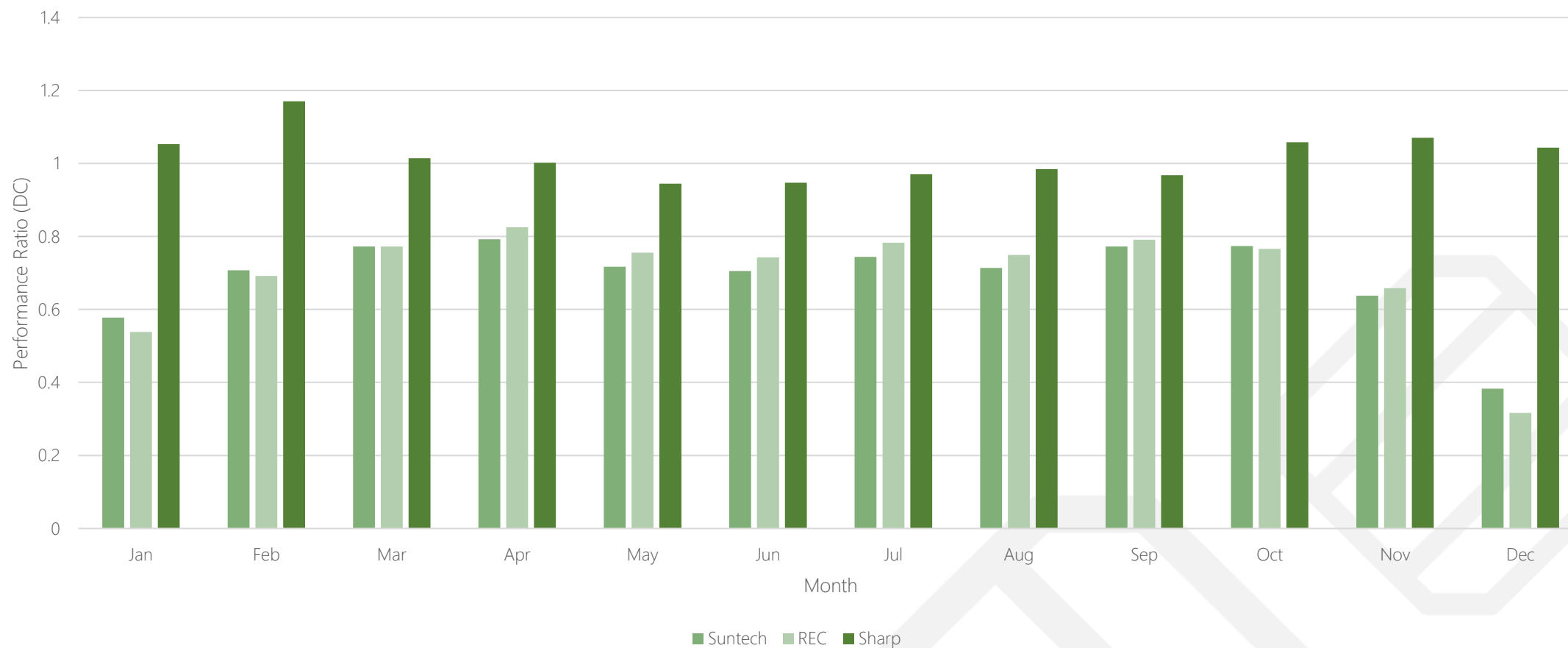
PV site

- ❖ Shadow
- ❖ Snow





# Performance Ratio



# ❖ SUPSI dataset

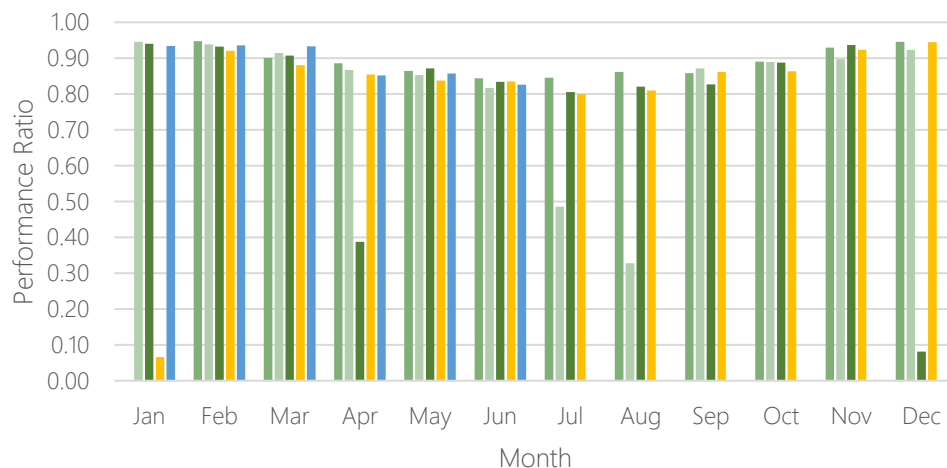
- ❖ C-Si modules, Al BSF:
  - ❖ 1) Glass/EVA/Backsheet
  - ❖ 2) Glass/PVB/Glass
- ❖ Installation configurations:
  - ❖ 1) open-rack configuration
  - ❖ 2) BIPV-insulated configuration (no air gap behind).



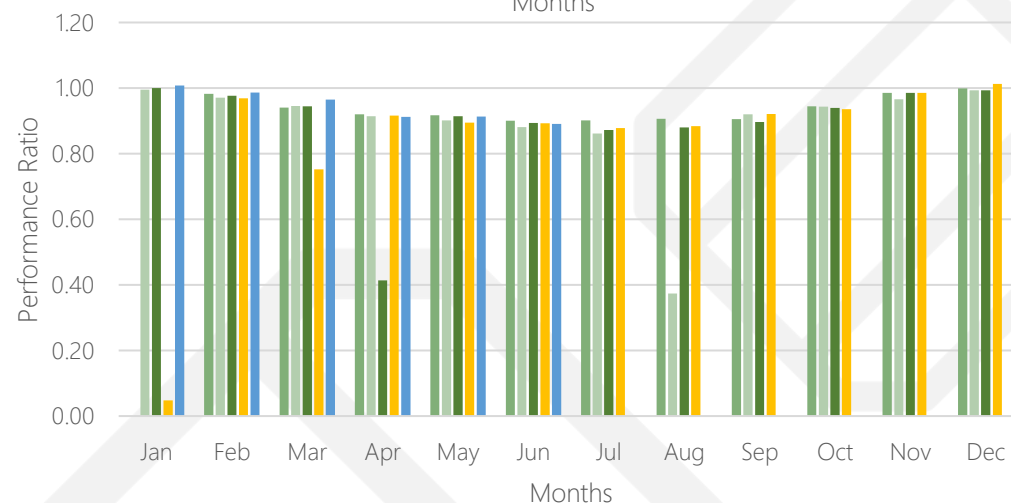
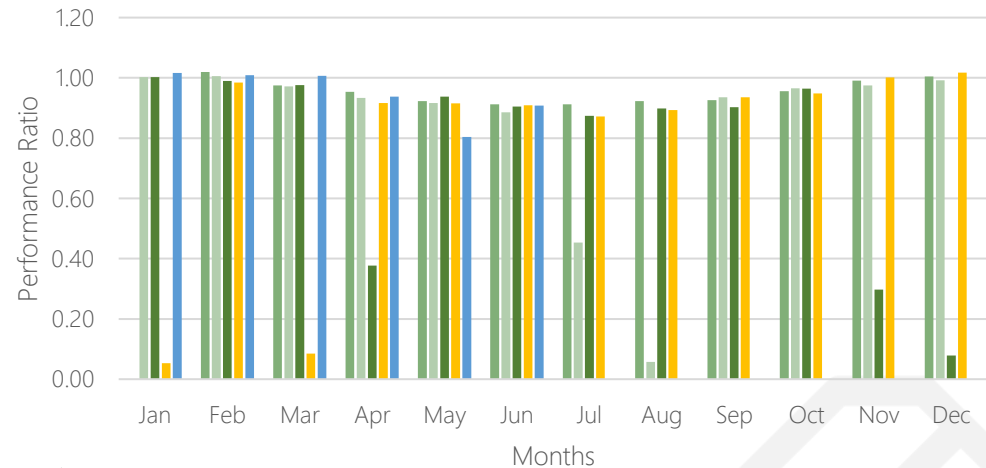
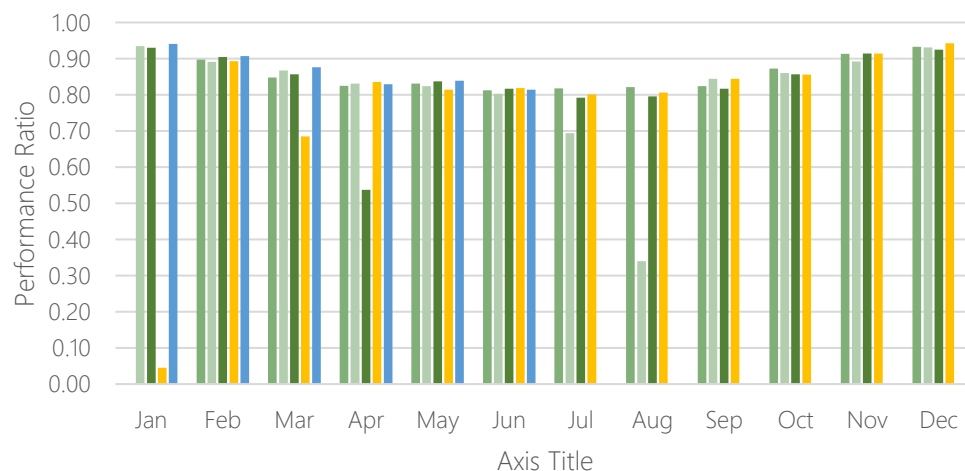
First date of measurements	2014-01-29
Last date of measurements	2018-07-10

# Performance Ratio

G/EVA/BS



G/PVB/G



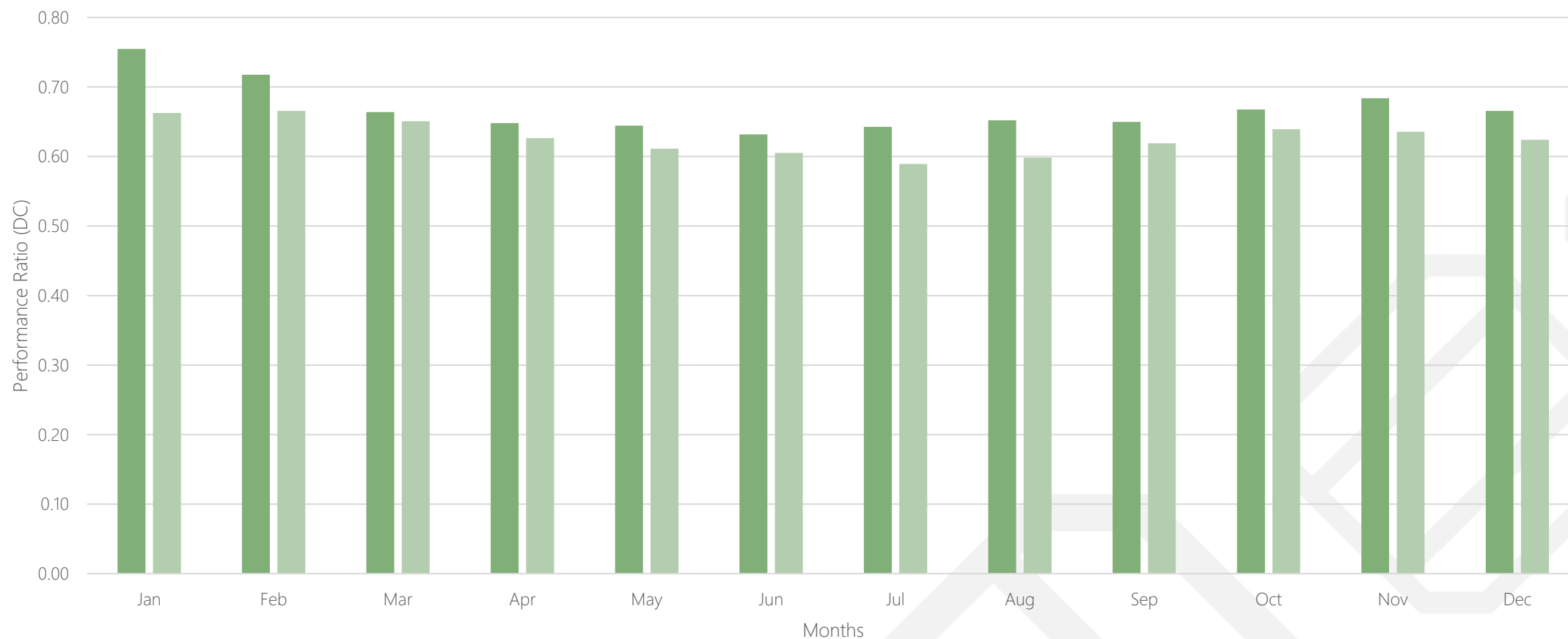
# EURAC dataset

- ✧ Eurac's test facility in Bolzano, Italy
- ✧ Two years
- ✧ 1 hour resolution
- ✧ polycrystalline silicon systems
- ✧ Data:
  - ✧ GPOA
  - ✧ GHI
  - ✧ T\_Module
  - ✧ P\_AC
  - ✧ P\_DC



First date of measurements	2014-01-01
Last date of measurements	2015-12-31

# Performance ratio



# ≡ Bern BIPV

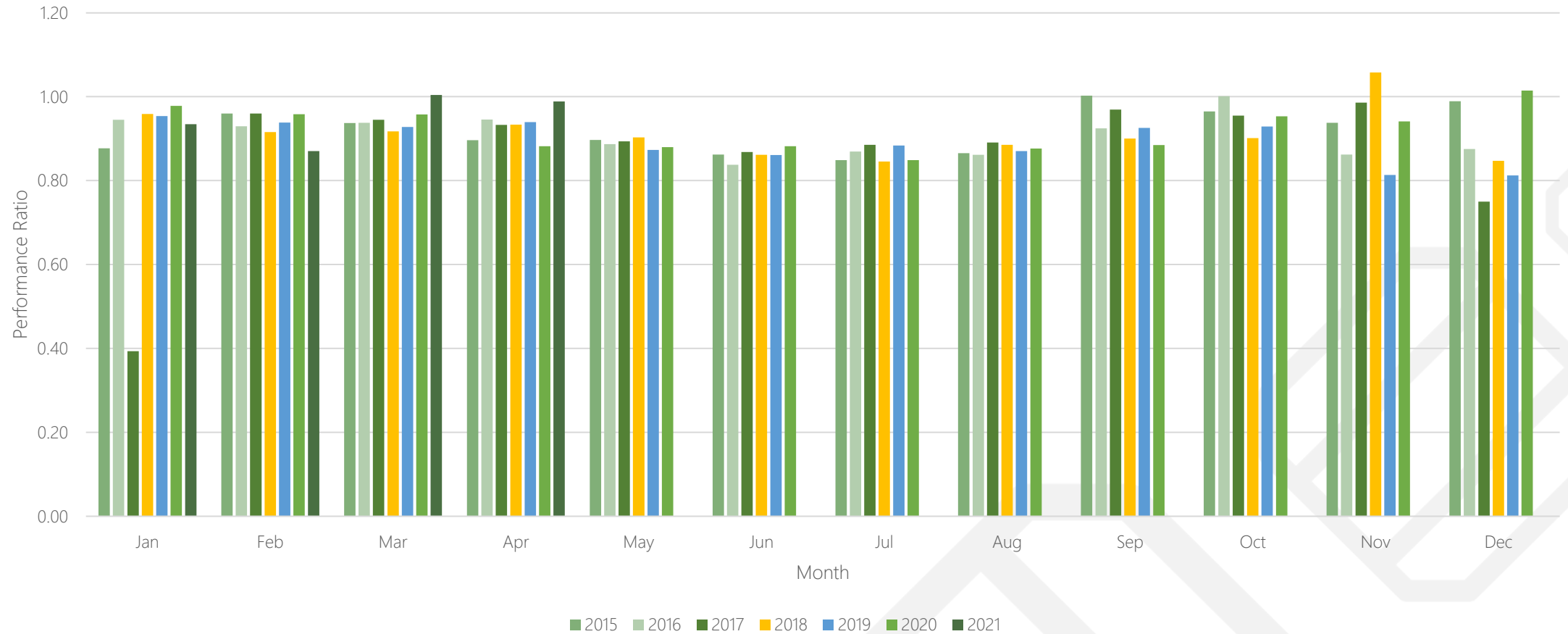
## ≡ Data:

- ≡ one string of a BIPV system in the Bern area (Switzerland).
- ≡ The string size is 5.23 kWp.
- ≡ Low amounts of near shading from nearby tree and chimney on rooftop.

First date of measurements	2015-01-01
Last date of measurements	2021-04-21



# Performance Ratio





# ❖ Contributing to PEARL PV data challenge

- ❖ The results will be published in a joint paper
- ❖ [PearlPV data challenge 1: PV Performance - Organizations - Pearl PV CKAN Repository \(pearlpv-cost.eu\)](#)
- ❖ Please contact [e.shirazi@utwente.nl](mailto:e.shirazi@utwente.nl) if you are interested!
- ❖ [PearlPV data challenge 2: Fault detection - Organizations - Pearl PV CKAN Repository \(pearlpv-cost.eu\)](#)
- ❖ Please contact [atse.louwen@eurac.edu](mailto:atse.louwen@eurac.edu) if you are interested!

# Acknowledgement

- ◊ Norwegian Research Council
- ◊ University of Agder
- ◊ EPFL PV-Lab
- ◊ Baur AG
- ◊ SUPSI PV Lab
- ◊ Eurac Research

# Special Thanks to

- Atse Louwen, Eurac Research, Italy
- Anne Gerd Imenes, University of Agder, Norway
- Ebrar Oezkalay, SUPSI, Switzerland
- Hugo Quest, 3S, Switzerland

P  A R L P V

❖ Thank you for your attention

Question or comments?

Feel free to contact me at [e.shirazi@utwente.nl](mailto:e.shirazi@utwente.nl)