The importance of shared and synthetic datasets

Angèle Reinders (prof.dr.) University of Twente and Eindhoven University of Technology, the Netherlands, e-mail: a.h.m.e.reinders@utwente.nl

Presentation, Online Workshop COST Action PEARL PV, 2 December 2021

WG1: PV Data Management and the PEARL Data Challenge

Contents

Introducing myself

Main research topics

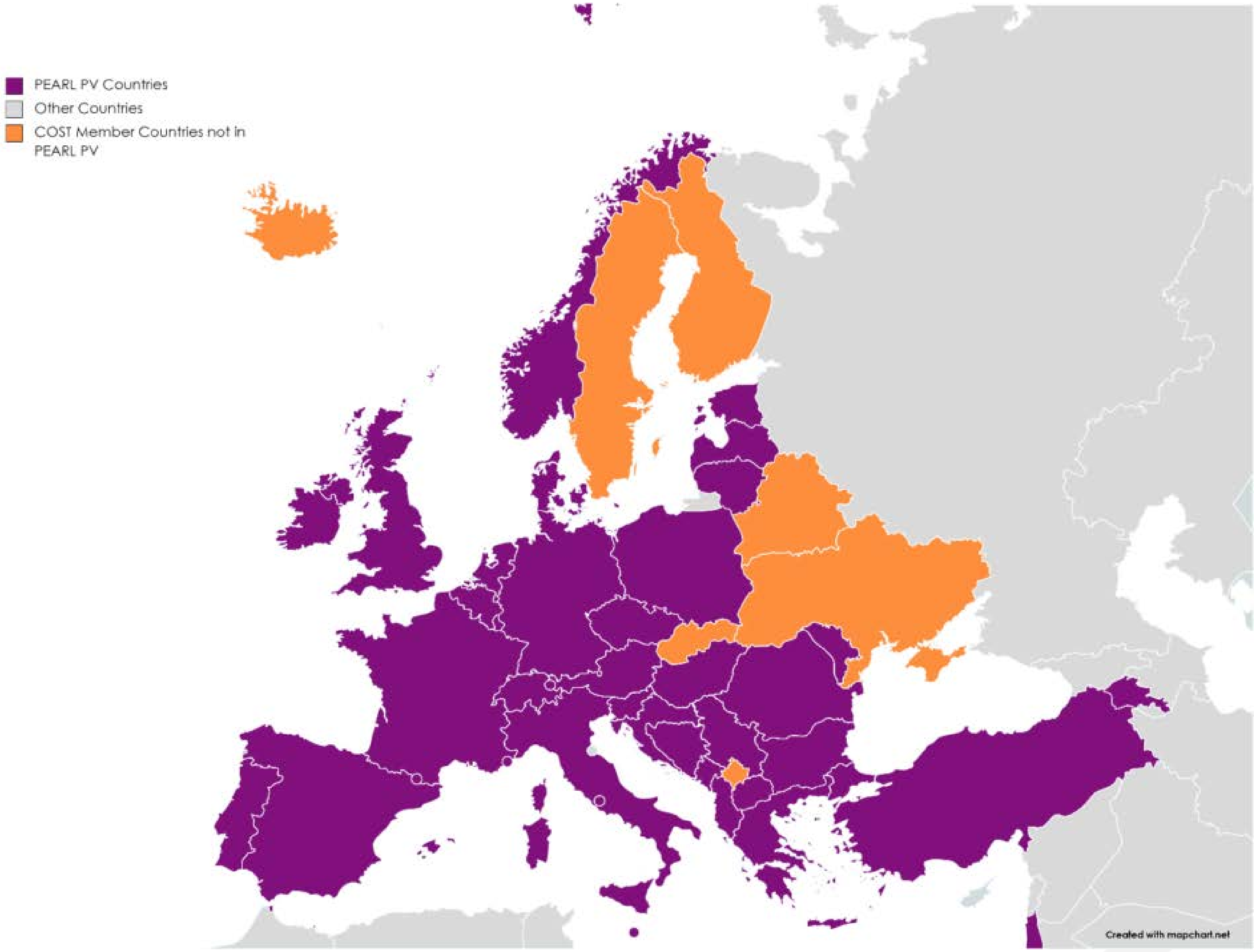
- Performance analysis of PV systems (COST Action PEARL PV)
- Design-driven research on smart grids
- Solar powered mobility with EVs
- Luminescent solar concentrators

Angèle Reinders (prof.dr.)

Short resume

- Professor of Design of Sustainable Energy Systems, Energy Technology Group, Fac. of Mechanical Engineering, Eindhoven University of Technology, the Netherlands, 2018 – present
- Associate Professor with ius promovendi, Dept. of Design Production and Management, Fac. of Engineering Technology, University of Twente, 2002 - present
- Professor of Energy Efficient Design, Fac. of Industrial Design Engineering, Delft University of Technology, Delft, 2010 – 2017
- PhD in 'Performance Analysis of Photovoltaic Solar Energy Systems', Dept. of Science, Technology and Society, Fac. of Chemistry, Utrecht University, 1999

PEARL PV Network



38 European countries take part in COST Action PEARL PV by December 2021, indicated in purple.
Not shown on this map, but participating, are the USA and Australia.

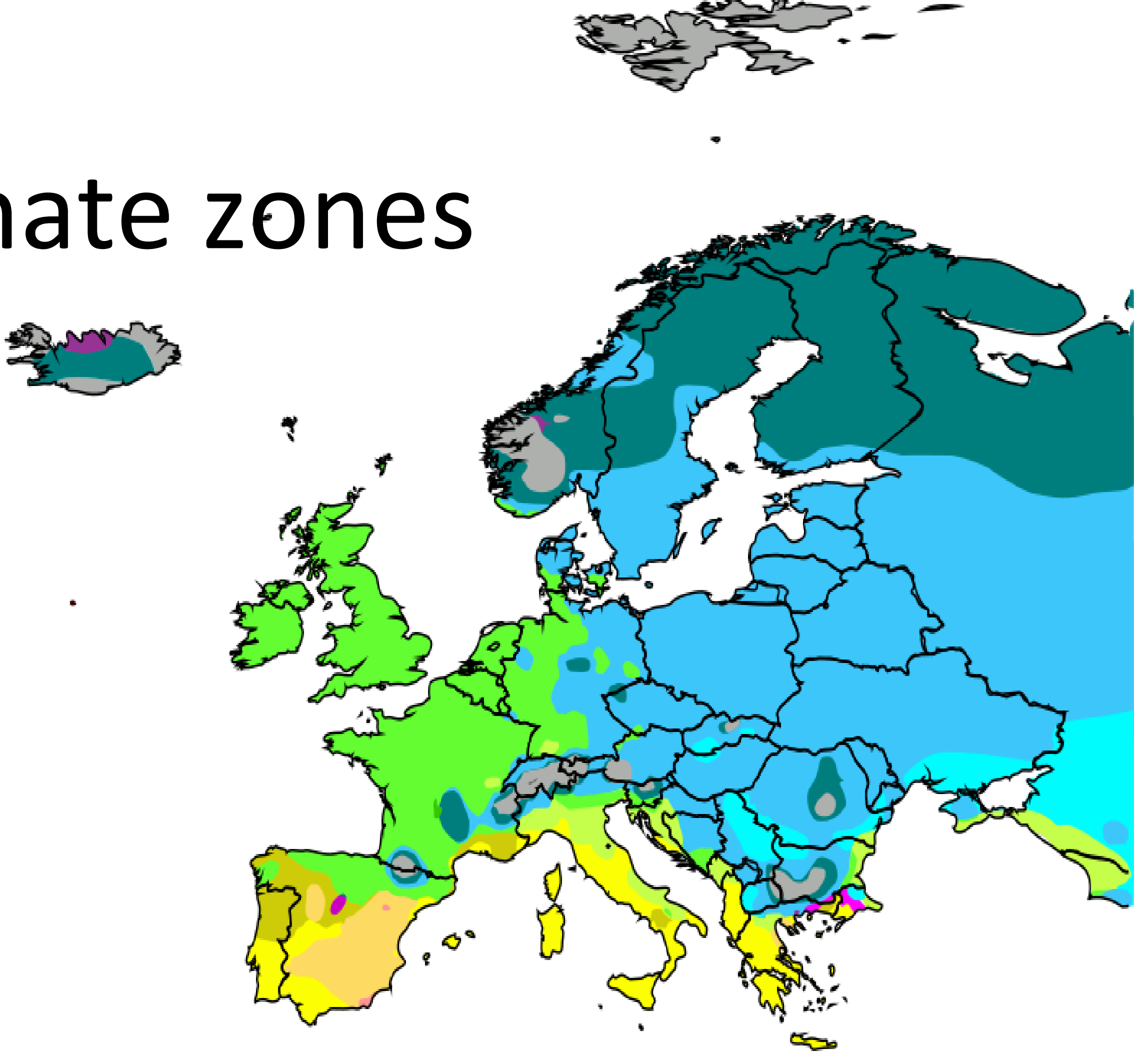
Activities in this network
-Bottom-up Research on PV systems
-Training Schools
-Seminars & Workshops,
-STSM grants & ITC conference grants

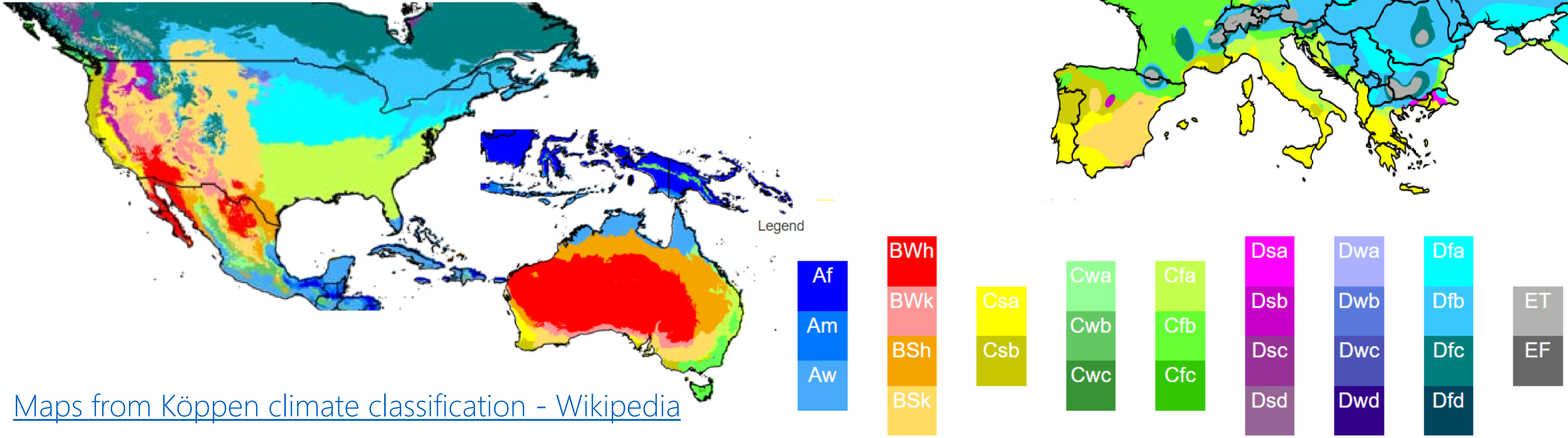
Aim to collaborate by
-Surveys
-Country reports
-Joint papers and other publications
-Joint research proposals
-Shared data bank

PEARL PV network’s climate zones

The geographic coverage of the PEARL PV member countries comprises almost all climate zones according to the Köppen–Geiger climate classification system:
A: Tropical, B: Arid, C:Temperate, D: Continental, E: Polar.

Therefore the network seems very suitable for solar PV research





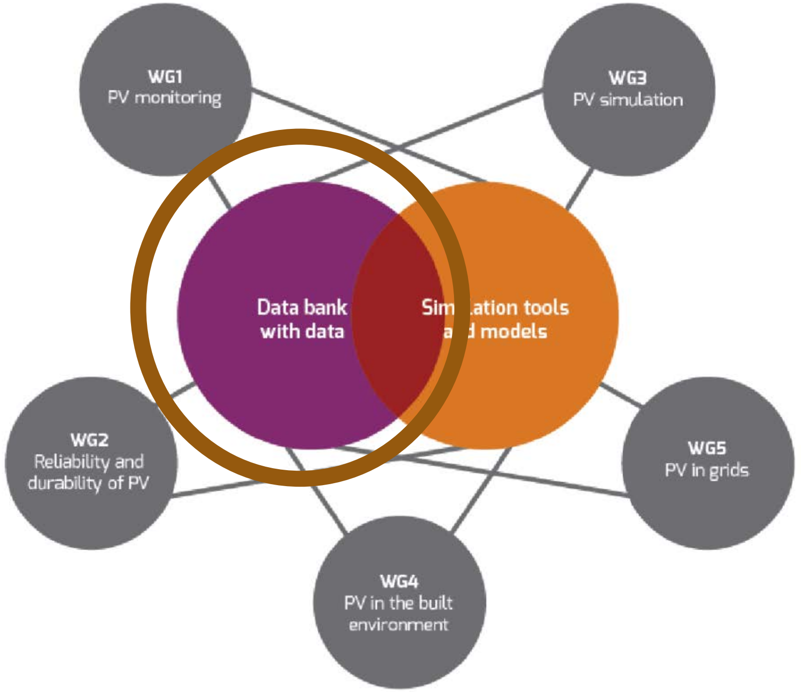
[Maps from Köppen climate classification - Wikipedia]

Objectives of PEARL PV

cost
EUROPEAN COOPERATION
IN SCIENCE & TECHNOLOGY

To **improve** the **energy performance** and **reliability** of PV systems leading to (i) **lower costs** by a higher yield, (ii) a **longer lifetime** and (iii) a **reduction of perceived risk**;

- By analyzing data of the long time monitored long-term performance of PV systems and of their defects and failures;
- To quantitatively determine the absolute influences of (i) components' rated performance, (ii) system design, (iii) installation type, (iv) operation and maintenance practice, (v) interactions with grids, (vi) geographic location and (vii) weather and climate conditions, on performance degradation over time and failure modes;
- To (i) improve the electrical design of PV systems, (ii) achieve optimal sizing via the use of simulation models, (iii) enhance expected system efficiency, (iv) ease maintenance, (v) achieve high reliability and (vi) demonstrate excellent durability.



All information about PEARL PV can be found on the website: Welcome - PEARL PV (pearlpv-cost.eu) including the MoU with research plan.

Working Group's data needs

Each Working Group has specific data needs based on the focus of the WG's research, see [1]

- WG1 PV Monitoring has set requirements for essential data and nice-to-have data that should be entered in the data bank. These data will be analyzed regarding actual monitored long-term performance, defects and failures in PV systems. Close collaboration with IEA-PVPS-Task 13
- WG2 Reliability and Durability of PV objectives: to define reliability and durability metrics for PV modules, components and systems, to identify relevant data to be collected to measure reliability and durability, and to share knowledge on mass PV data analysis methods for the identification of issues causing decrease in power output (e.g. shading, physical degradation…) and correlations of failure modes with climatic conditions.
- WG3 PV simulation considers the use of modelling tools to simulate the performance of photovoltaic devices and systems. WG3 could use the data server as a shared simulation tool facility.
- WG4 PV in the Built Environment will focus on the data coming from various built contexts, i.e. urban and rural environments from which information on urban morphology and discrete urban geometry can also be obtained. Integration in a 2D GIS, 3D GIS and BIM systems will be examined and implemented. The data of particular interest are derived from BIPV (Building Integrated PV) systems, by type of building and position of PV systems (rooftops, facades, shades, window glazing, etc.).
- WG5 PV in Grids: analyzes data and metadata of thousands of PV systems connected to grids in Europe, studies are conducted on the relationship between the PV production and the local consumption, the possible use of batteries, or the economic viability of alternative options. Fault detection toolbox to improve the energy yield of grid-connected PV systems and reduce their power instability.

[1] Reinders, A., Slooten, F. van, Moser, D., Sark, W. van, Oreski, G., Ottersboeck, B., Pearsall, N., Devetaković, M., Leloux, J., Capeska Bogatinoska, D., Braun, C., Gerd Imenes A., Driesse, A., DEVELOPMENT OF A BIG DATA BANK FOR PV MONITORING DATA, ANALYSIS AND SIMULATION IN COST ACTION 'PEARL PV', Conference Proceedings of EU PVSEC, 2019

Why it is relevant to share data?

COST EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

- <u>Four good reasons:</u>

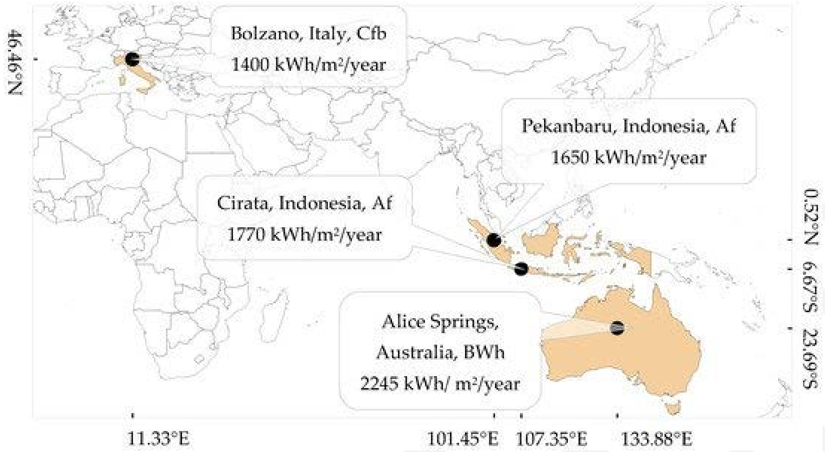
- To be able <u>to compare</u> the performance and reliability of (fleets of) PV systems installed in different climate zones, with different components in various installation categories, different age and different conditions of operation and maintenance (otherwise said, different meta-data)
- <u>To exchange knowledge</u> and tools regarding PV monitoring, methods for data processing, data analytics and machine learning, and different simulation approaches, and, in that way, enhance international collaboration and the quality of PV system research in general
- <u>To develop a professional research infrastructure</u> for PV system research
- <u>To increase the statistical validity</u> of PV system research results

Examples of projects with shared data

cost
EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

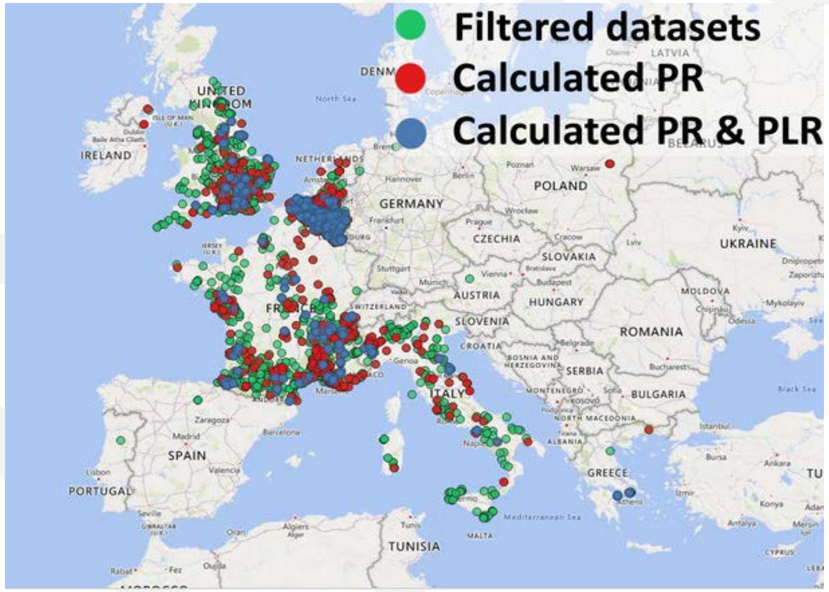
Kunaifi, K., Reinders, A.H.M.E., Lindig, S., Jaeger, M., and Moser, D., Operational performance and degradation of PV systems consisting of six technologies in three climates, Applied Sciences, 10(16), 5412, 2020.

An analysis (and comparison) of monitoring data of PV systems in **Indonesia, Australia and Italy**. Data from **2008 to 2019**, ranging from two to nine years, from **fifteen PV systems with 6 different PV technologies** were analyzed regarding Performance and Ratio and Performance Loss Rate. The study covered CIGS, a-Si, poly-Si, mono-Si, CdTe and HIT PV systems



Lindig, S., Ascensio, J., Leloux, J., Moser, D. and Reinders, A.H.M.E., Performance analysis and degradation of a large fleet of PV systems, IEEE Journal of Photovoltaics, ISSN: 2156-3403, DOI: 10.1109/JPHOTOV.2021.3093049, 2021.

Analysis of monitoring data of **~8400 PV systems with 10-minute recordings for 2010-2016** with mainly crystalline Si PV modules taken from data bank (data from Rbee Solar + ERA5 satellite), with a focus on the determination of Performance and Ratio and Performance Loss Rate

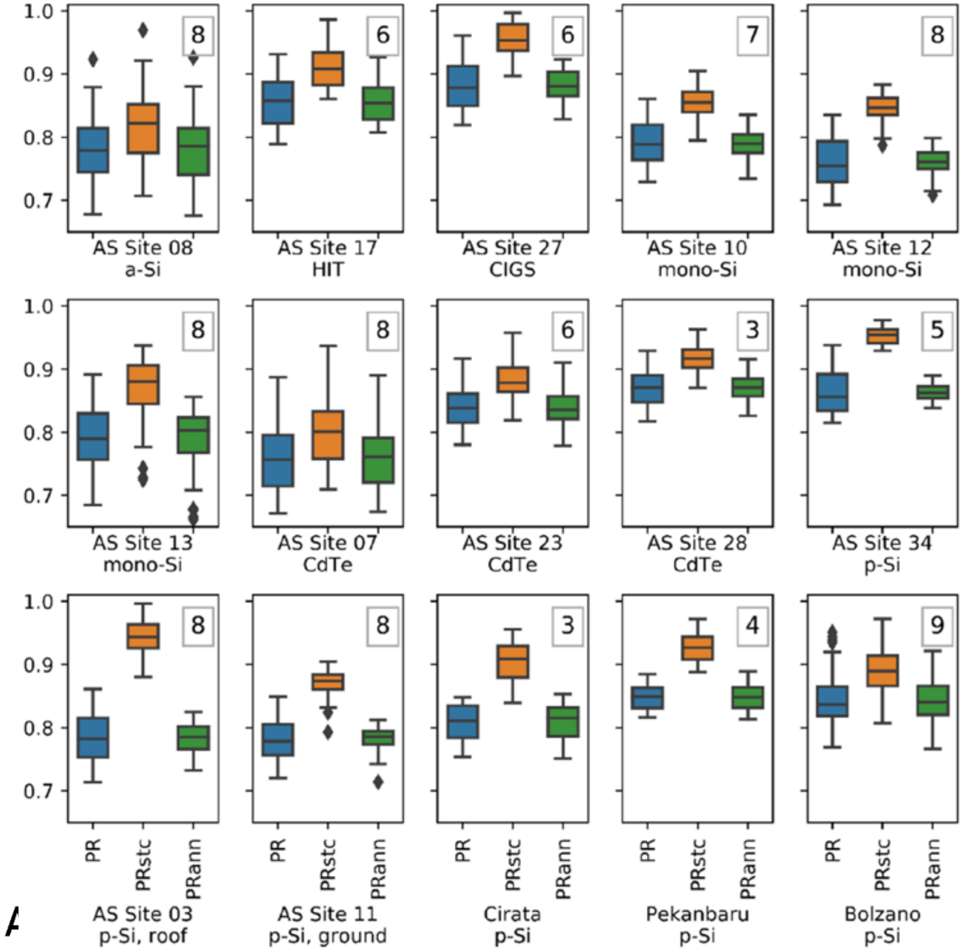


See today's presentation 11:20-11:40 Performance Analysis and Degradation of a Large Fleet of PV Systems Dr. Sascha Lindig (Eurac Research)

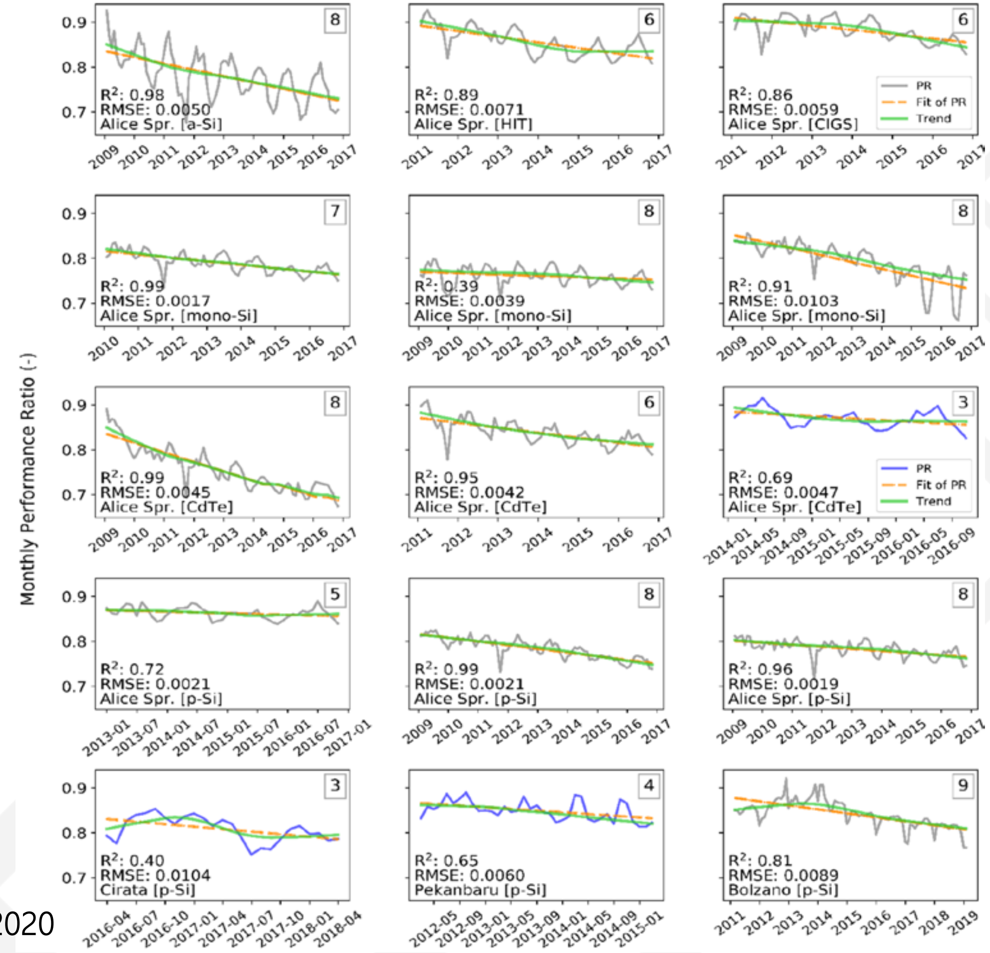
Examples of projects with shared data

EUROPEAN COOPERATION
IN SCIENCE & TECHNOLOGY

Box plots of the performance ratio of each PV system. Left/blue: PR, middle/orange: temperature-corrected PR by STC, right/green: annual-averaged temperature-corrected PR. Numbers at the upper right corners in the boxes indicate the duration of the analyzed data in the number of years.



Annual-averaged monthly temperature-corrected performance ratios, PR_{ann}., of the PV systems (grey or blue lines), their linear fit (dashed orange line), and trend components (green line) over the monitoring period. period. Numbers at the top right of the charts indicate the lengths of the data in years.

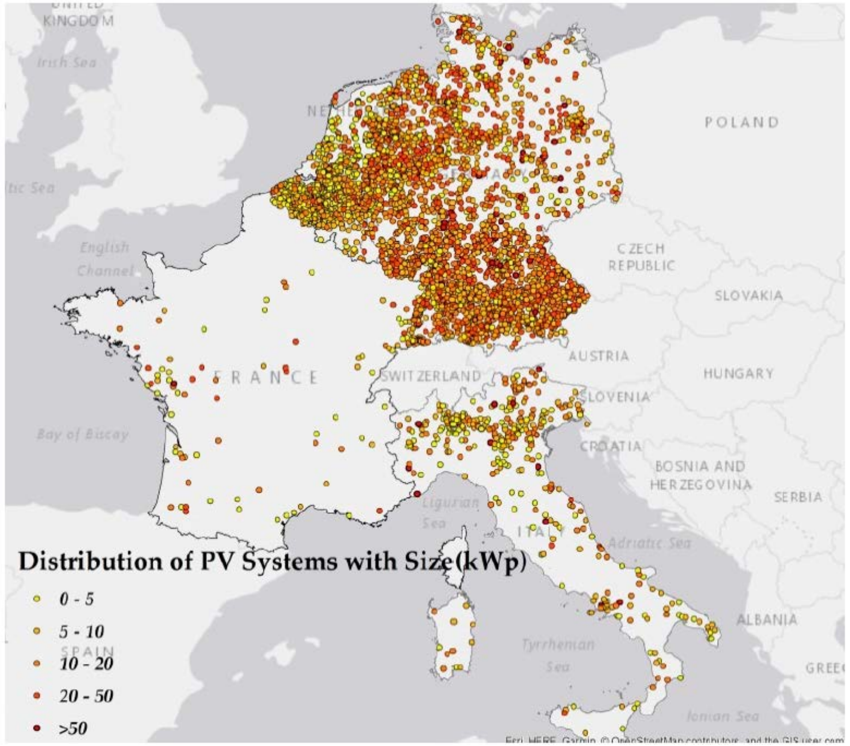


Kunaifi et al, 2020

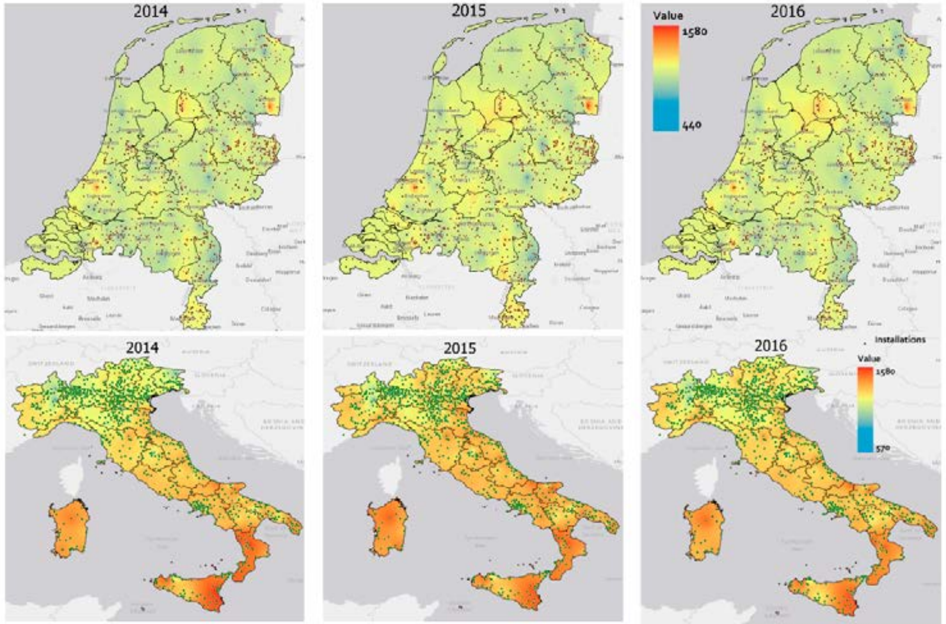
Examples of projects with shared data

Bala Bhavya Kausika, Panagiotis Moraitis and Wilfried G. J. H. M. van Sark, Visualization of Operational Performance of Grid-Connected PV Systems in Selected European Countries, Energies, doi:10.3390/en11061330, 2018

Web scraping techniques were employed to collect detailed yield data at high time resolution (5–15 min) from a large number (31,844) of small-scale systems (each at least 1 year of monitoring data) for the analysis of their operational performance by means of performance ratio.



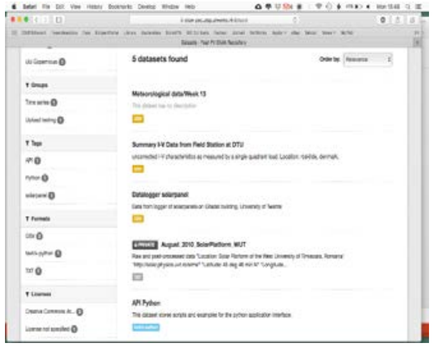
Annual specific yield variation in the Netherlands and Italy using visual interpolation techniques

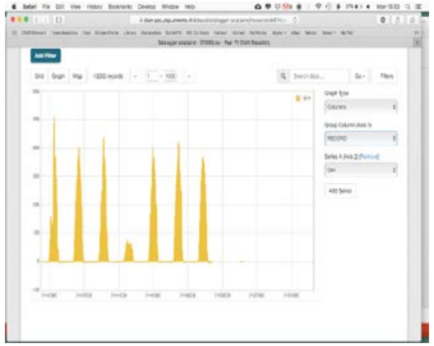


Thoughts about shared data

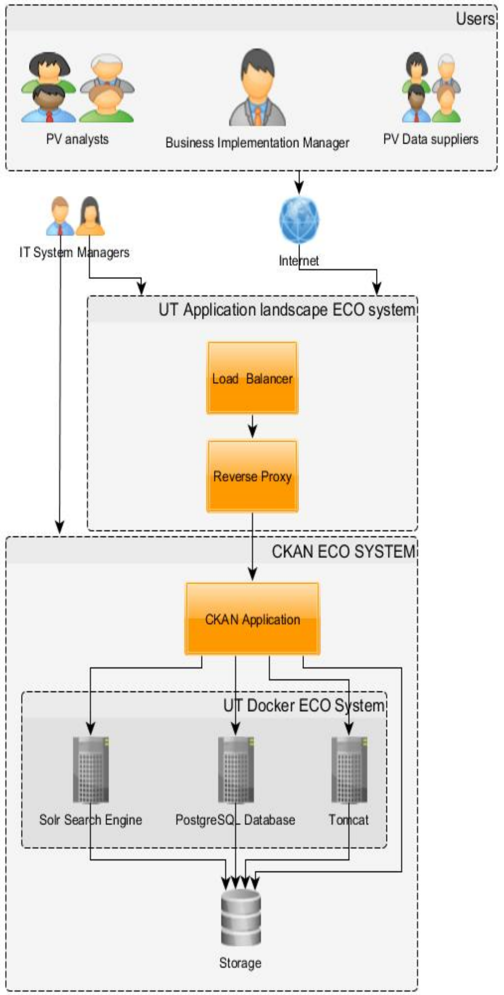
- Different quality of data monitoring systems
- Component faults and operation & maintenance practices are unknown
- This also applies to specific local conditions: shading and animals
- Modern problems such as congestion of grids and the related disconnection of PV systems, are not monitored, neither is curtailment well embedded in monitoring.
- Data streaming of many PV monitoring systems to a central data server would be well possible, but hardly happens.

Introduction to CKAN data server



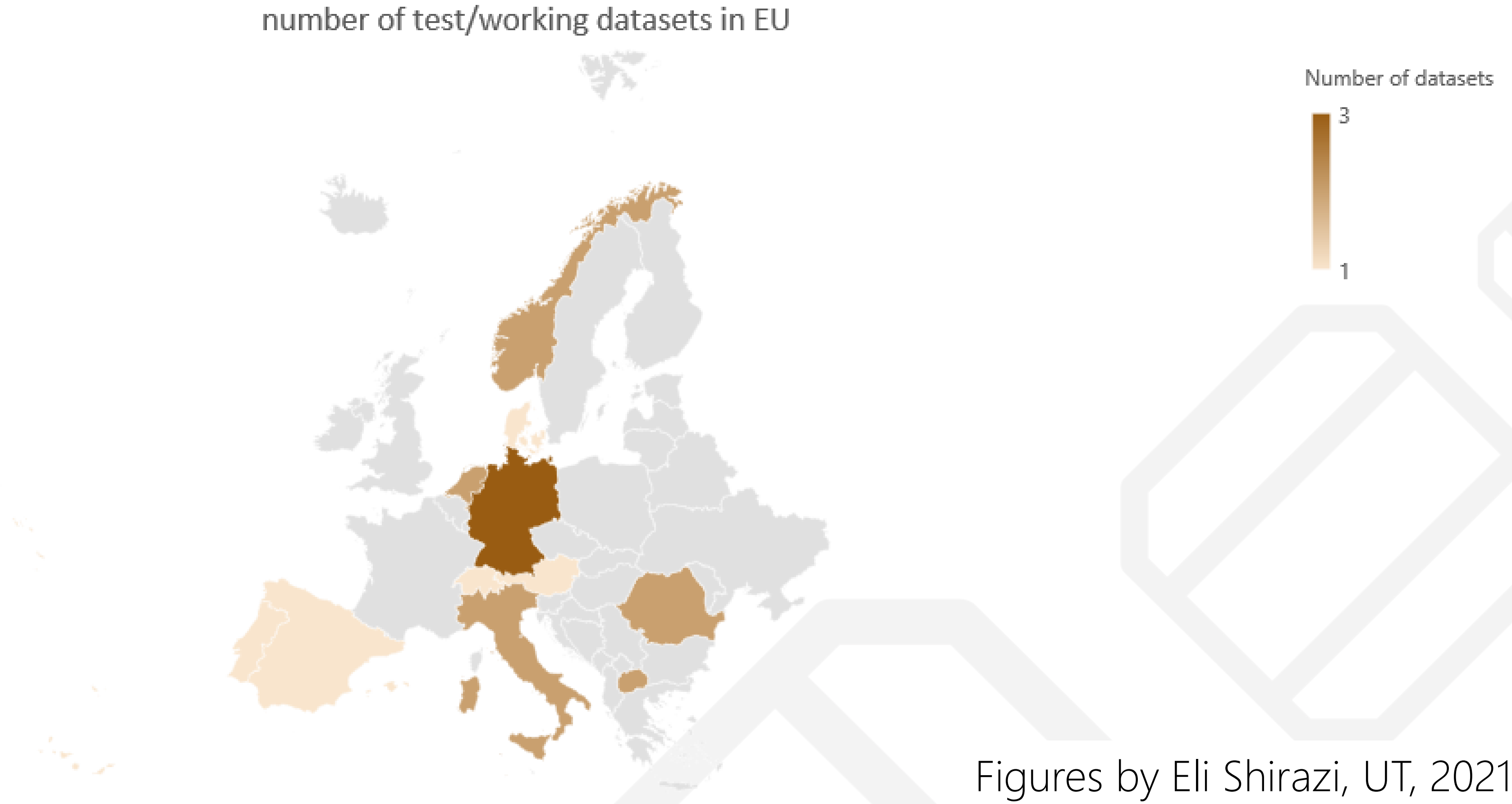


- PEARL PV's CKAN data server was been implemented in 2019 in order to share data in COST Action PEARL PV
- This means in practice: a database server (4 CPUs, 16 GB memory, max 4 TB diskspace) was installed successfully, on which CKAN is running. Upload setting: max. 200 MB/file can be easily increased.
- It can store data and run software in the data server
- In principle it will remain available after this Action has been completed and can be used for the next decades
- Access by https://ckan.pearlpv-cost.eu/ and next **login** (see red circle)
- Suitable for data uploads incl. meta-data
- How does it work? That will be explained today: 11:40-11:50 How to use the Pearl PV CKAN database Dr. Atse Louwen (Eurac Research)
- This is not the only data sharing initiative, others: DuraMAT, IEA PVPS Task13, , BDPV, Sonnenertrag, PVOutput, as well as NREL's PV Fleet Performance Data Initiative



See also: Reinders, A., Slooten, F. van, Moser, D., Sark, W. van, Oreski, G., Ottersboeck, B., Pearsall, N., Devetaković, M., Leloux, J., Capeska Bogatinoska, D., Braun, C., Gerd Imenes A., Driesse, A., DEVELOPMENT OF A BIG DATA BANK FOR PV MONITORING DATA, ANALYSIS AND SIMULATION IN COST ACTION 'PEARL PV', Conference Proceedings of EU PVSEC, 2019

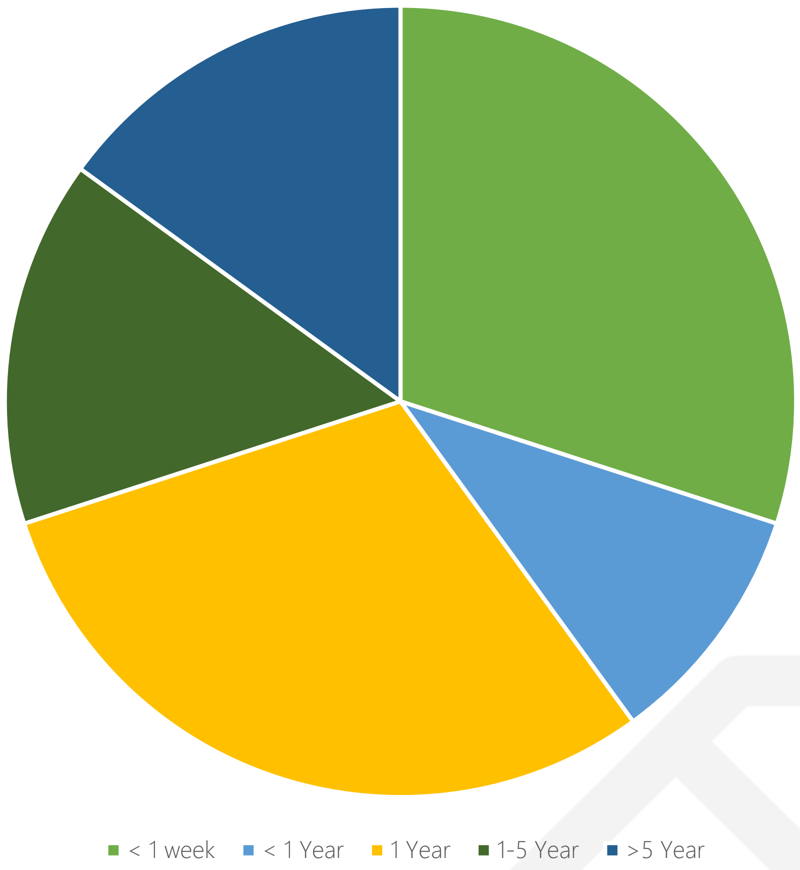
Data in the CKAN data server at present



Figures by Eli Shirazi, UT, 2021

Data in the CKAN data server at present

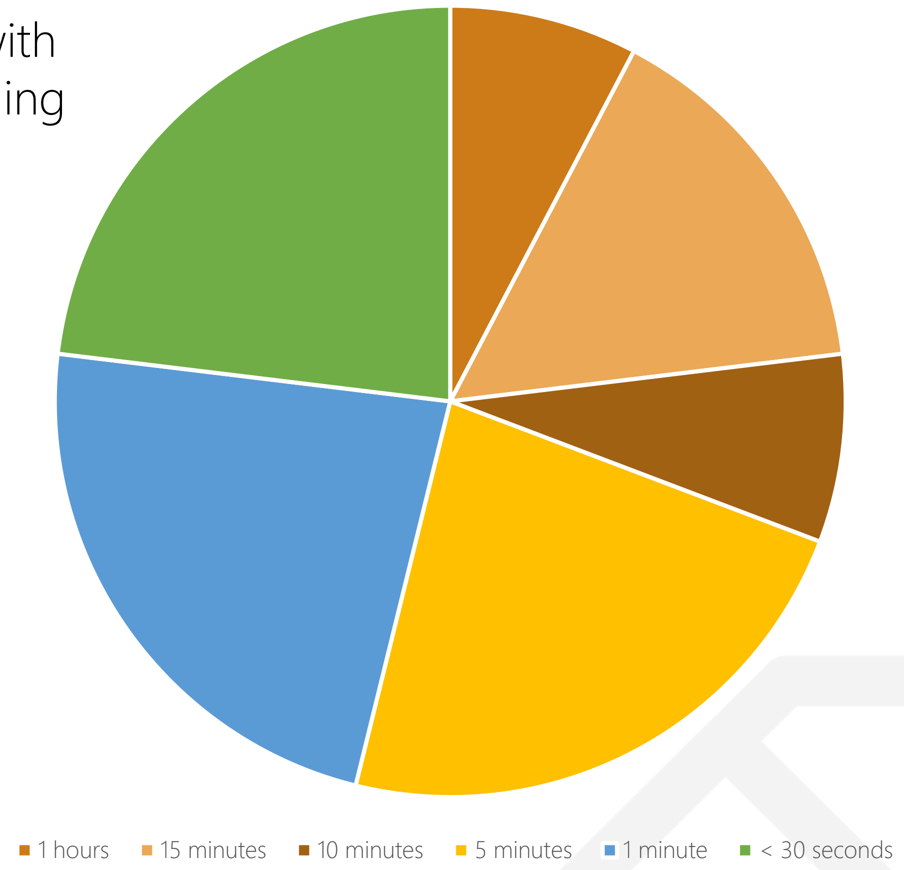
In total ~38 years of data
1. We will need more data!
2. In particular data with a longer monitoring period > 1 year



Figures by Eli Shirazi, UT, 2021

Datasets Resolution

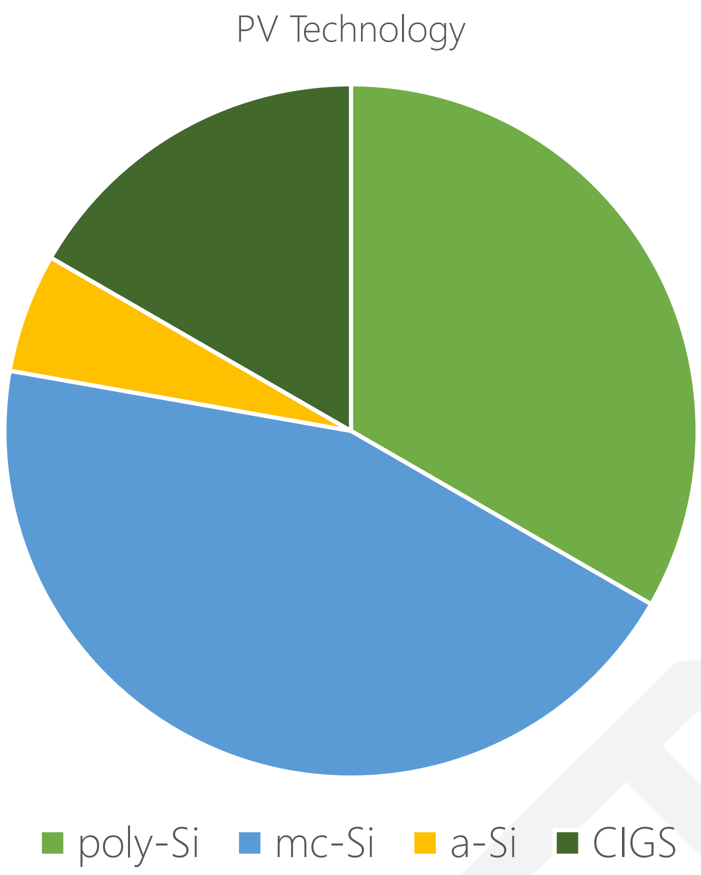
Relative high share of data with very short < 1 minute recording intervals



Figures by Eli Shirazi, UT, 2021

Datasets by PV technologies

The most domination PV technologies in present PV installations are covered.



Figures by Eli Shirazi, UT, 2021

Data of interest to the CKAN data server

- **Dedicated meteo data,** for instance data of spectrally distributed irradiance, see today's presentation at 11:00h on the Use of PEARL PV platform to Share Spectral data by Basant Raj Paudyal (University of Agder)
- **PV monitoring data** with short recording intervals (< 10 minutes) and a long monitoring period (> 1 year).
- PV monitoring data of **fleets with PV systems with different meta-data**
- **Meteo and PV monitoring data from all PEARL PV member countries, incl. USA, Australia and Armenia** Please visit slide 13 to observe whether your country might need to contribute.
- **Artificial or synthetic data sets** for PV system research, to be further explained today by Atse Louwen when introducing the data challenge to you0

Conclusions and outlook

- Please don’t hesitate to use the CKAN data server.
- Your data will be protected by an NDA which complies with the European guidelines for data management.
- Therefore this can be considered to be an opportunity to collaborate and professionalize the PV system research sector.
- --
- Could we please ask you to upload several PV monitoring data sets before the end of 2021?
- Then we will maybe be able to fill up the 4 TB of storage we have ;-)

To end with

- All information about research and network activities can be found in the detailed COST Action PEARL PV Work Plan 2021-2022,

- Atse Louwen, Eli Shirazi, Carolin Ulbrich and Wilfried van Sark are friendly thanked for preparing and implementing this workshop!

Dates of future online Workshops (OW), Training School and Final Conference:
2 December 10-13h: OW, WG1 Data Challenge
15 December 9-12:30h: OW, WG2 on Reliability and Durability of PV
19 January 9-12:30h: OW, WG4 on PV in the Built Environment
2 February 9-12:30h: OW, WG3 on PV Simulation
15 February 9-12:30h: OW, WG5 on PV in Grids
8-11 March: Training School, at University of Twente, Enschede, the Netherlands
14-16 March: Final Conference, at University of Twente, Enschede, the Netherlands

PEARLPV

COST EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

Thanks to all for your interest in collaborating!

Question or comments? Please email me at a.h.m.e.reinders@utwente.nl or feel welcome to respond now.