Anomaly Detection and Classification for Photovoltaic Systems

— **Master's Thesis** —

Author: **Maxime Charrière**
Advisors: Prof. Dr. **Christof Bucher**, Prof. Dr. **Horst Heck**
Place: Bern - August 30, 2024

KEY TAKEAWAYS

Main Principle:
- Leverages data from neighboring PV systems to estimate daily production.

Universal Anomaly Detection Method:
- Operates without the need for meteorological data or additional sensors.

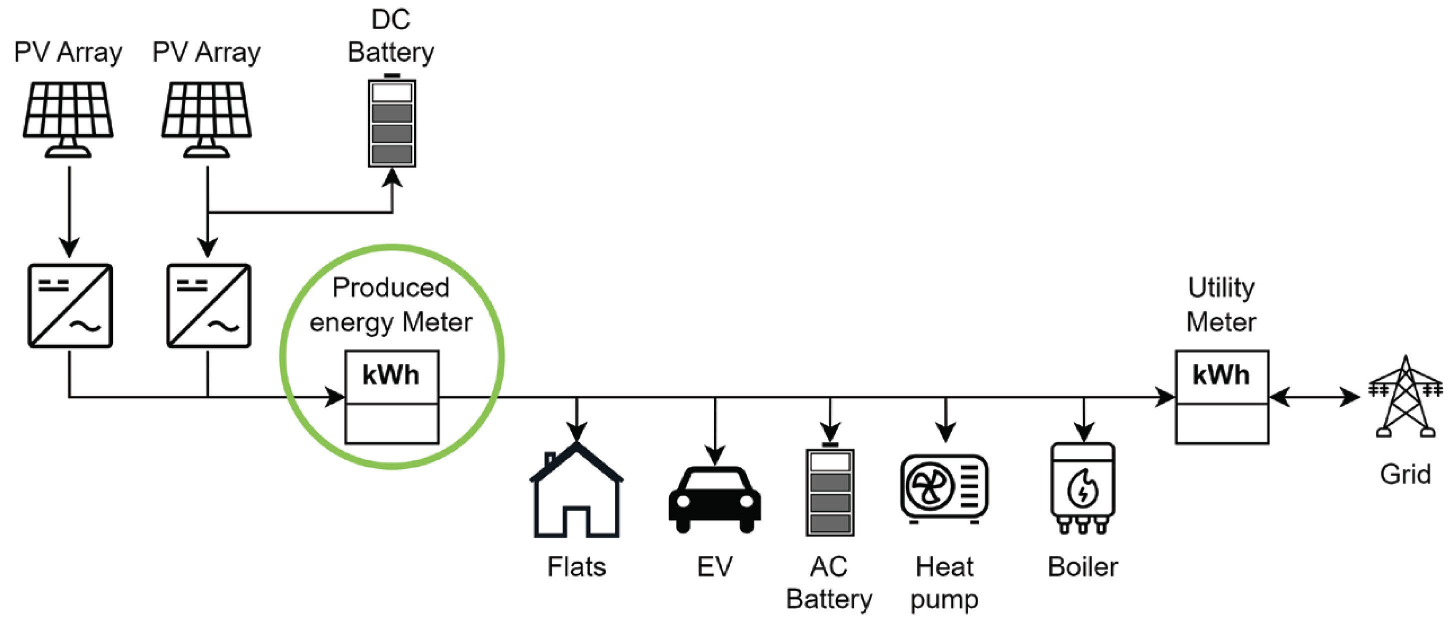
Promising Results:
- Estimate daily production with a Mean Absolute Percentage Error (MAPE) of 4.35%.
- Detect simulated anomalies with 97.4% accuracy.

INTRODUCTION
- The number of residential PV systems in Switzerland is rapidly increasing.
- **Detecting anomalies is crucial** to reduce Mean Down Time (MDT), increase solar production, and minimize economic losses.
- Each residential PV system **varies in configuration and available data**.
- However, daily energy production is a **commonly available metric**.

OBJECTIVES
- Develop a **universal** method to **detect and classify anomalies** in PV systems, using **only energy production** measurements.

TARGET GROUP
- Companies monitoring residential/small-scale PV systems.

USED DATA
- **Daily energy production** data from 326 PV systems in Switzerland over approximately 400 days.
- **Characteristics** of each system (location, orientation, size).



Detecting Anomalies in Solar Panels?

Look no further than their **Neighbors !**

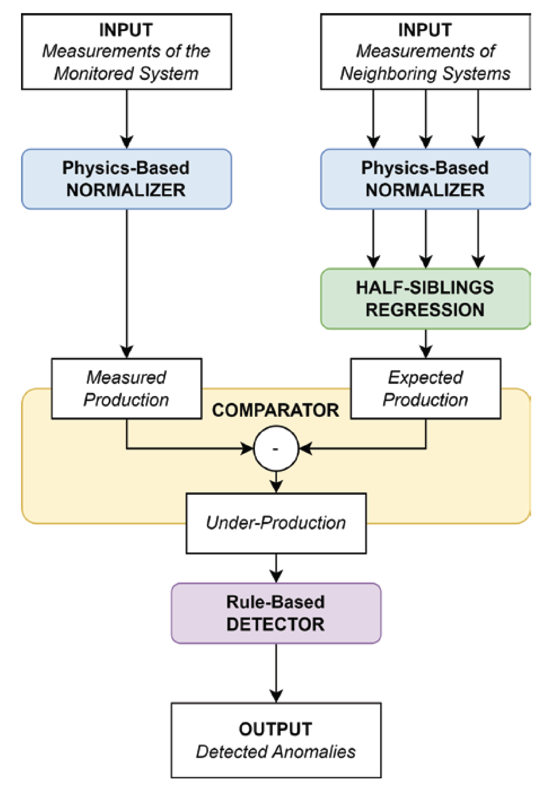
READ MORE

Bern University of Applied Sciences

PRiOT
PROFESSIONAL IOT SOLUTIONS

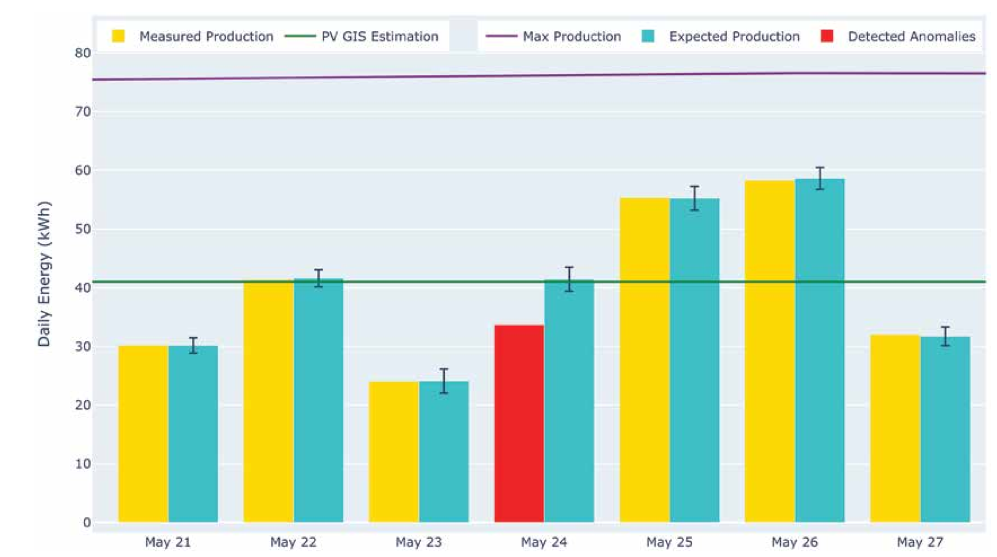
METHODOLOGY
The production of a PV system is influenced by three main types of factors:
1. **System-Specific Factors** (location, orientation, size, losses, ...)
2. **Regional Factors** (primarily weather conditions)
3. **Anomalous Factors** (which we aim to detect)

Anomalies are detected by isolating their impact from system-specific and regional factors through:
- **Normalizer**: A physics-based simulator that accounts for System-Specific Factors. It normalizes data to enable the comparisons between neighboring systems with different characteristics.
- **Half-Sibling Regression**: A Machine Learning model that account for Regional Factors. It estimates the expected production, by comparing production from neighboring systems.
- **Comparator**: Compares expected production with actual production to estimate underp-production.
- **Detector**: Use statistics and predefined rules to flags days with abnormal underproduction as anomalies.



RESULTS
Daily, our monitoring tool:
- **Estimates Expected Production** (with a MAPE of 4.35%).
- **Detects nomalies** (with 97.4% accuracy on simulated anomalies).
- Provides the **Maximum Daily Production** and the **System Performance**.

To do this, our monitoring tool needs:
- At least **2 neighboring** systems within a 3 km radius.
- At least **7 days** of historical data.



LIMITATIONS
- **We use only daily data**, limiting our method accuracy and restricting monitoring to a daily frequency.
- **No classification** of detected anomalies is made, requiring its manual analysis to determine the cause.

FUTURE DEVELOPMENT
To put our method into practice, we recommend the following:
- Develop a **prototype** in real-world conditions.
- Gather **user feedback** for continuous improvement.
- **Enhance detection** algorithms based on the feedbacks.
- **Label anomaly** types as they occur, to enable the development of automatic classification.
- Develop an automatic **anomaly classification**.