# STRATEGIES FOR WIDE-SCALE SHORT-TERM PV FORECASTING IN ENERGY COMMUNITIES

Nikolaus Houben, Lennard Visser, Wilfried van Sark, Hans Auer, Amela Ajanovic, Reinhard Haas

IAEE European Conference – Sept, 2022, Athens

## Research Objective

**Quantify the economic impact of various data and modelling strategies on short-term rooftop PV power forecasting from the perspective of an aggregator**

## Methods

### Phyical Model

– Predefined Parameterized Mapping of Irradiance to Power

– Used Model: PV Watts Model[1] (pvlib python package)

$G_{poaeff}$

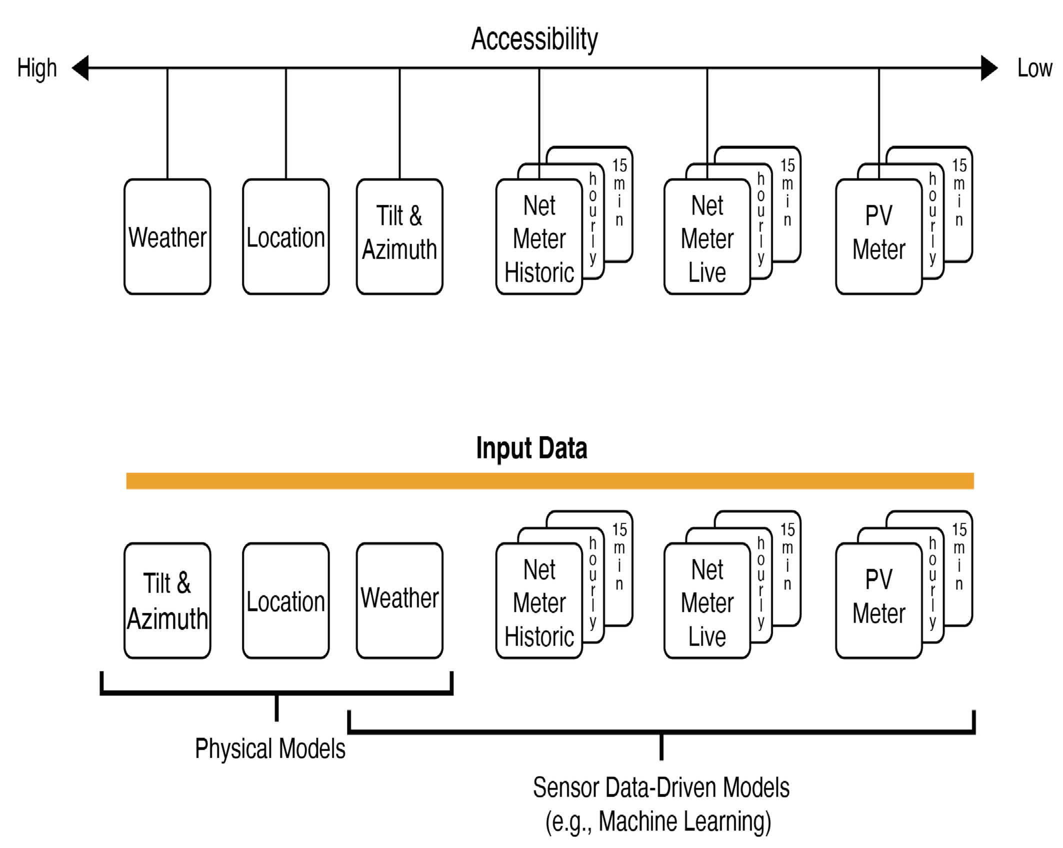
$$P_{dc} = \frac{G_{poaeff}}{1000} P_{dc0}(1 + \gamma_{pdc}(T_{cell} - T_{ref}))$$

$$\eta = \frac{\eta_{nom}}{\eta_{ref}}(-0.0162\zeta - \frac{0.0059}{\zeta} + 0.9858)$$

$$P_{ac} = \min(\eta P_{dc}, P_{ac0})$$

| Input | Unit |
|---|---|
| System Size (Pdc0) | kW (DC) |
| Module Type | Standard, Premium, Thin film |
| System Losses | % |
| Tilt Angle | Degrees |
| Azimuth Angle | Degrees |

**Table 1: PVWatts Model Input Parameters**




### Machine Learning Models

– Function Approximation based on Input and Output Data

– Finding the parameters of a function that makes the data most likely

– Used Models: Support Vector Machine, Random Forest, XGBoost, Multi-Layer-Perceptron

– Features: Global Horizonal Irradiance, Diffuse Horizontal Irradiance, Direct Normal Irradiance (Autoregressive Measurements)

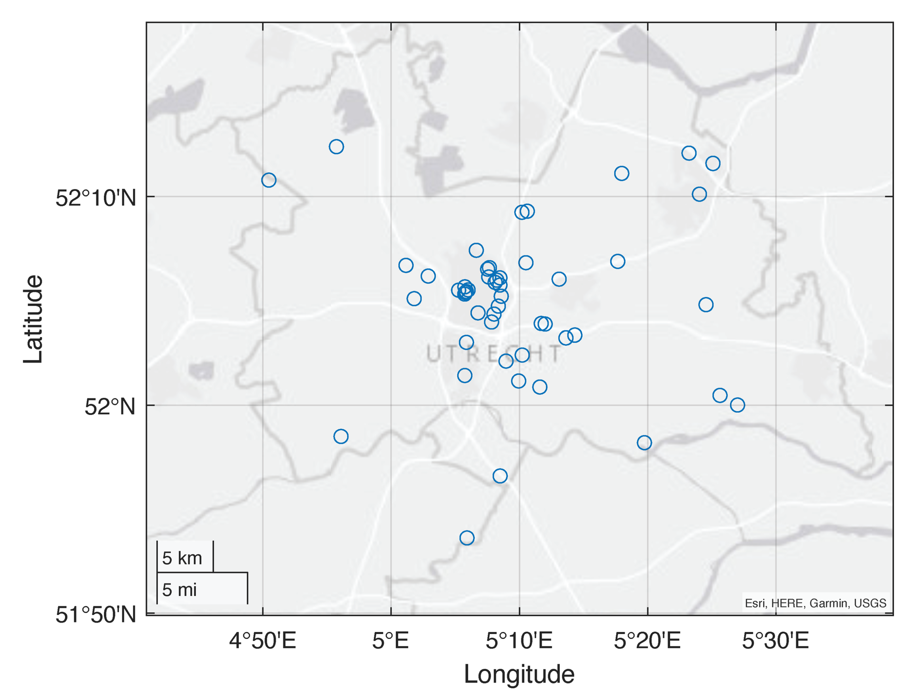
## Case Study & Data[2]

### Data Description

– Photovoltaic Data:
- 50 Rooftop PV systems in Utrecht, NL
- Power Data: 30s resolution, 2015-2017
- Meta Data: Location, Tilt, Azimuth, Capacity

– Weather Data:
- measured* Historic DNI, DHI, GHI, air temperature, 2015-2016

– Economic Data:
- Imbalance Prices for NL from ENTSOE
- dual Imbalance Prices, 2015-2021
- multiplied by the absolute error accordingly



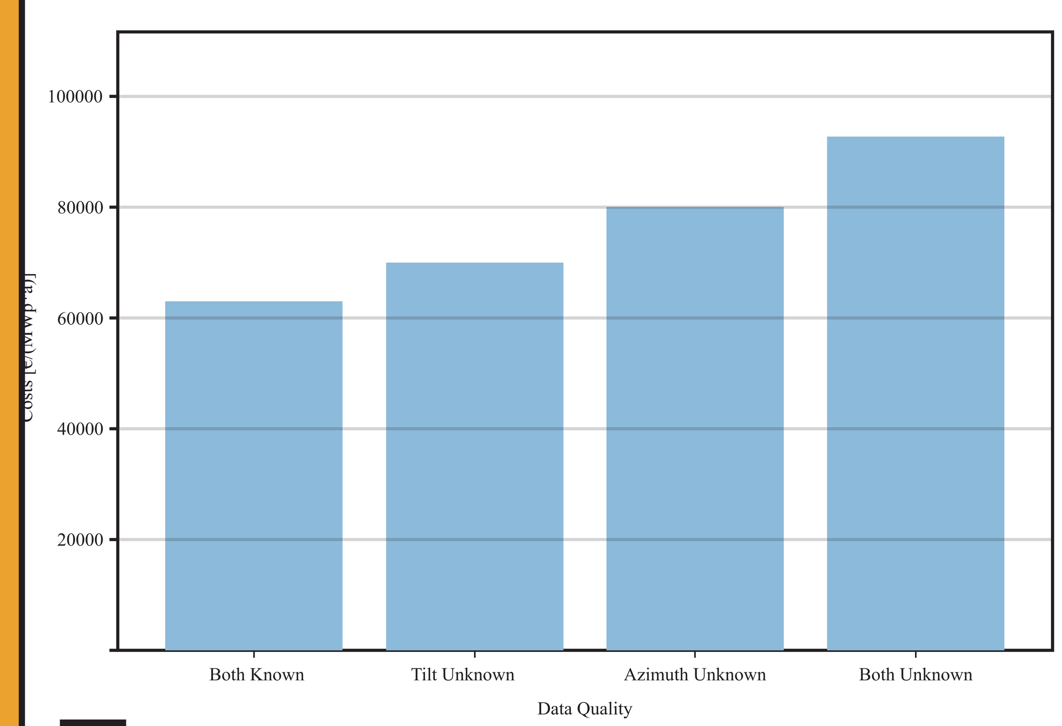

## Key Results from the Case Study

### #1: Quality of Meta Data

– Situation: No measurement data, limited meta data, so the aggregator has to estimate

– Idea: Randomize Tilt angle, Azimuth Angle or both for all systems

– Results: Knowing both angles saves up to 20% in costs



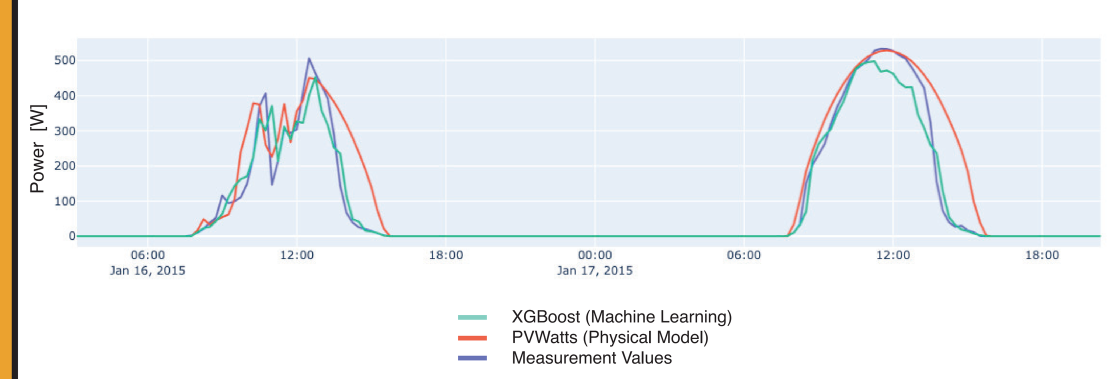

### #2: (Machine) Learning Shading Losses

– Key Idea: Input (Features) and Output (Target) Variables provide implicit information of the developing or present losses of the PV System

– Visual Selection of Days with apparent Shading (see below)

– Comparison of the Physical and best Machine Learning Model

– Result: Machine Learning models significantly outperform physical models on days with shading



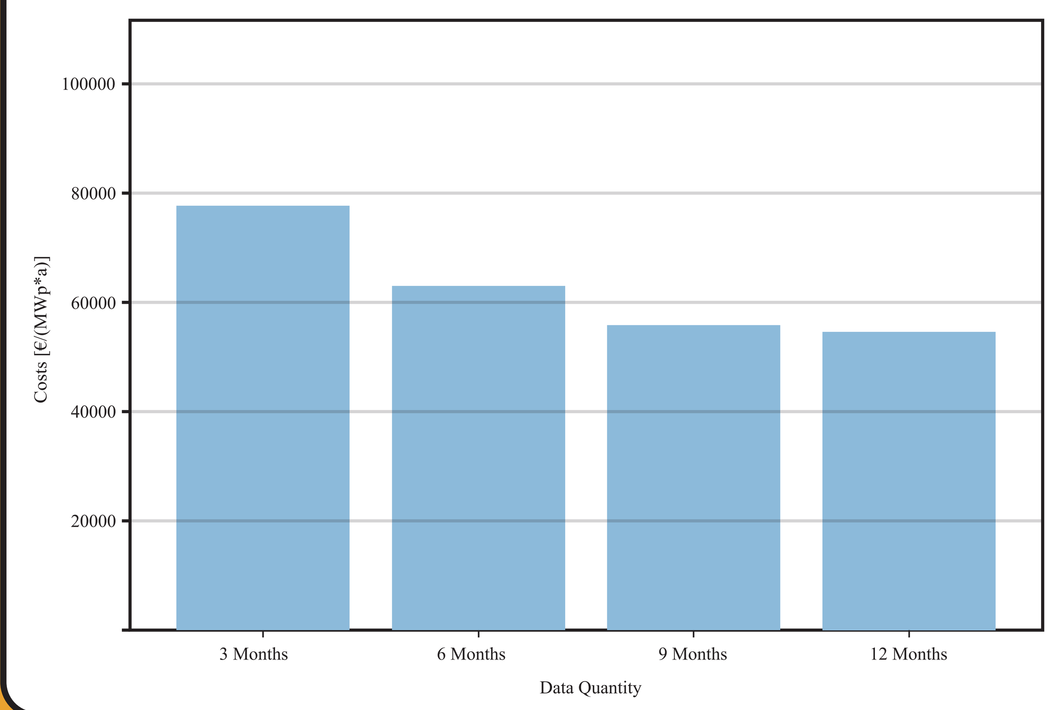

### #3: Quantity of Training Data

– Situation: varying availability of data for a PV System

– Key Idea: Incrementally Increase the size of the Training Data

– The validation period was fixed!

– Results show a 20% decrease in costs (and errors) when a full year of training data is available compared to 3 months




## Conclusion

1

When sensor meter data is unavailable, collecting metadata on system azimuth & tilt angles can significantly increase forecast accuracy and reduce costs

2

Having information on both correct azimuth and tilt angles has a synergistic effect regarding cost savings

3

Machine learning methods are generally superior to physical methods as they learn losses, however limited quantity of data can deteriorate accuracy substantially.

[1] Dobos, A. P. (2014). PVWatts version 5 manual (No. NREL/TP-6A20-62641). National Renewable Energy Lab.(NREL), Golden, CO (United States).
[2] Visser, L., Elsinga, B., Alskaif, T., & van Sark, W. G. (2022). Open-source quality control routine and multi-year power generation data of 175 PV systems. Journal of Renewable and Sustainable Energy.