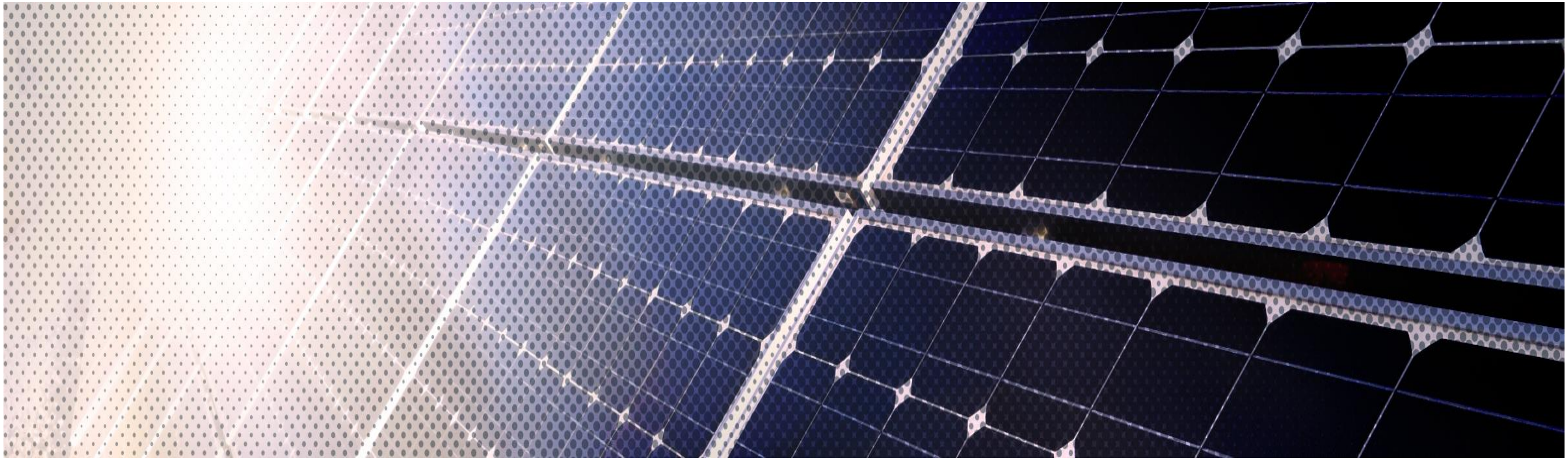


Performance evaluation of PV power predictive models for real-time monitoring

Andreas Livera[1], **Marios Theristis[1]**, George Makrides[1], Juergen Sutterlueti[2], Steve Ransome[3] and George E. Georghiou[1]
[1]PV Technology Laboratory, University of Cyprus, Nicosia, Cyprus
[2]Gantner Instruments GmbH, Schruns, Austria
[3]Steve Ransome Consulting Ltd, Kingston upon Thames, UK

Acknowledgement

Specific Objective: Development of an innovative condition monitoring platform for proactive and reactive O&M with enhanced data analytic functionalities

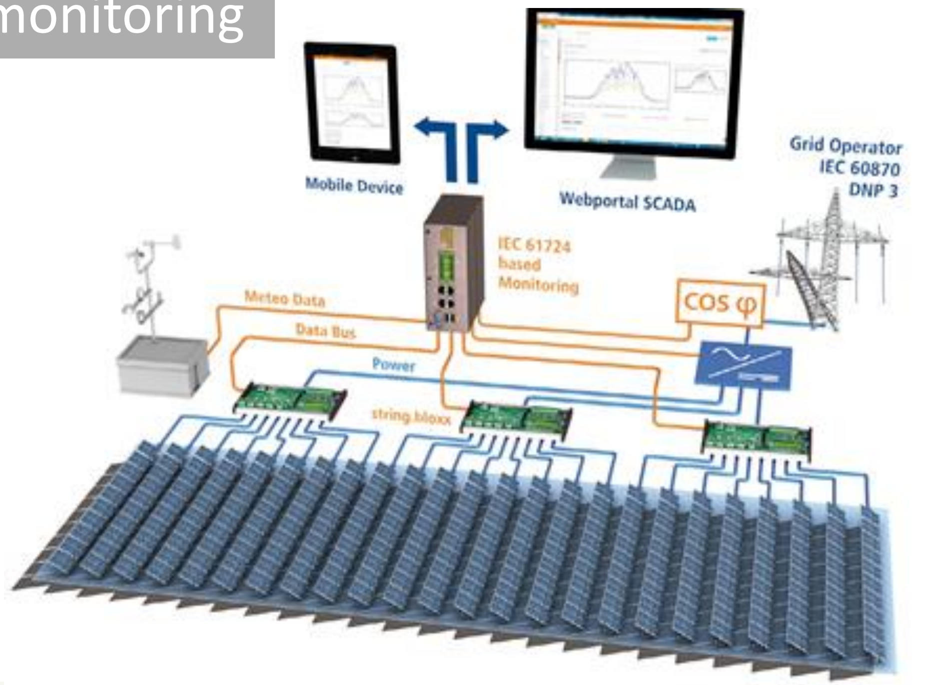
Advanced baseline condition monitoring solution to ensure operational quality and optimise energy production

Partners: GI and UCY
Project: Innovative Performance Monitoring System for Improved Reliability and Optimized Levelized Cost of Electricity IPERMON [Solar-ERA.net project]
Budget: €400,000
Duration: 36 Months (April 2016 – Sept 2019)
Weblink: http://www.pvtechnology.ucy.ac.cy/projects/ipermon/

Introduction

- Accurate output power prediction is crucial for PV performance assessment
- Predictive models are required for data-analytic features of advanced PV monitoring systems

Advanced PV system monitoring



Data-analytic features

- System health state
- Failure diagnosis

FOSS
Research Centre for Sustainable Energy

University of Cyprus
PV Technology

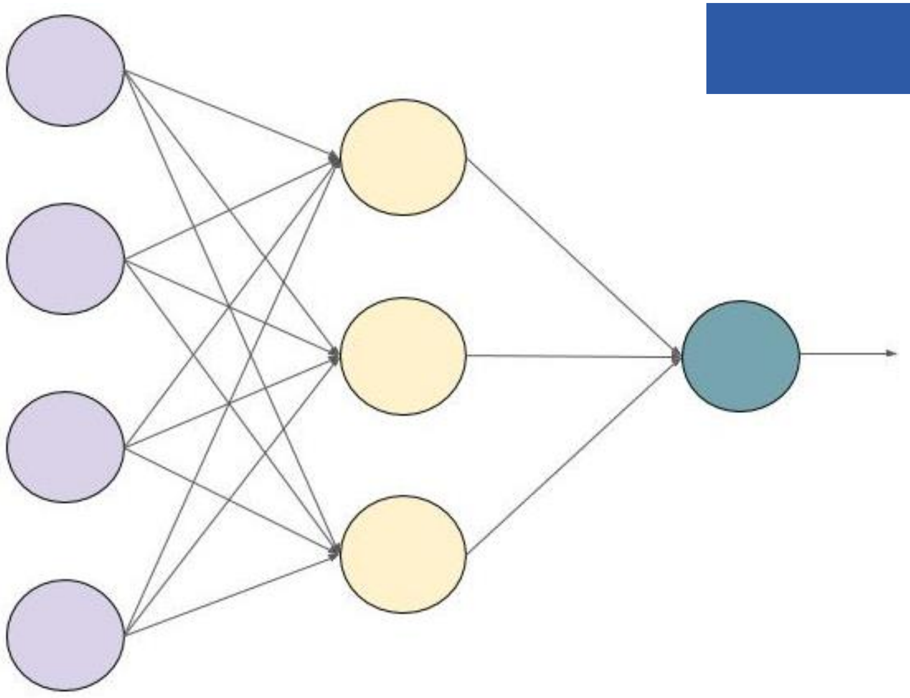
Gantner instruments

Objective

Development of an optimized location- and technology-independent predictive modeling methodology at **minimum requirements**

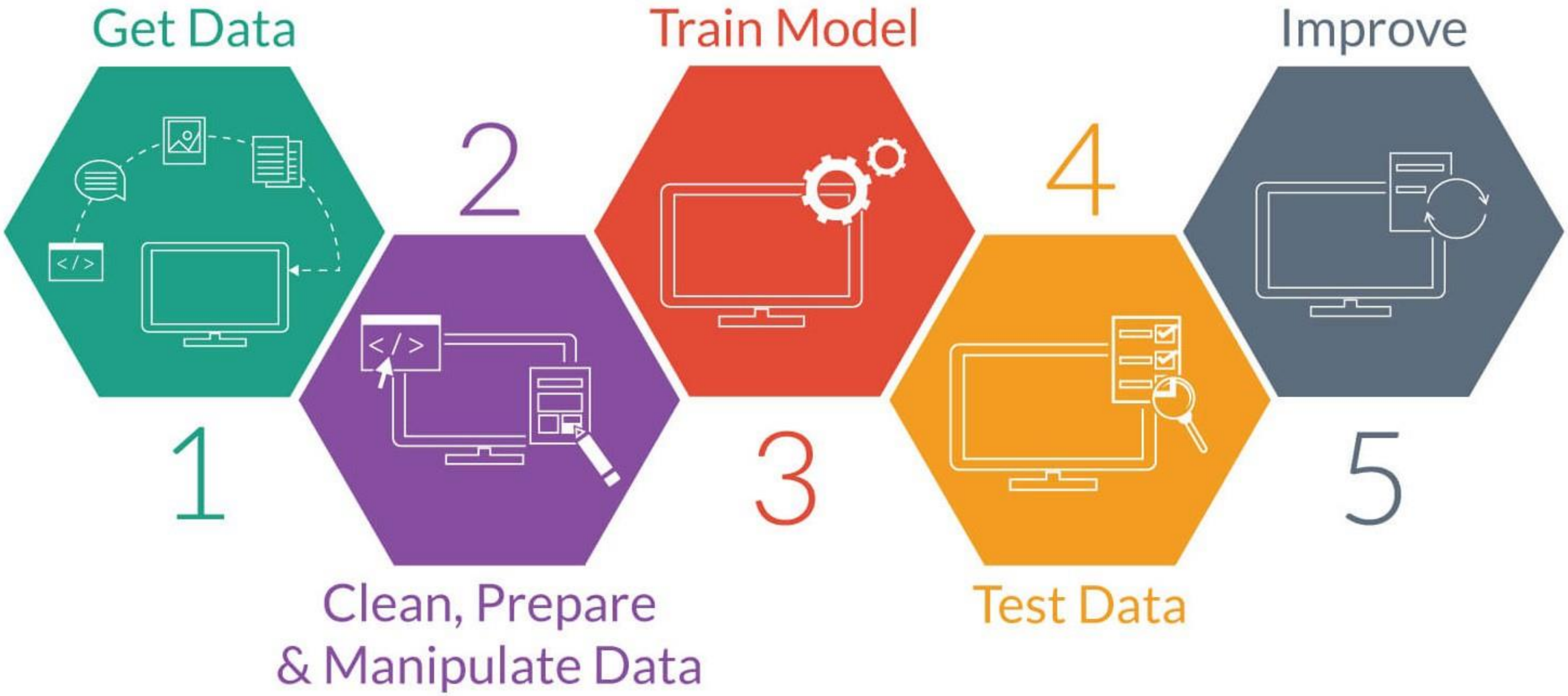
Input

- Features
- Dataset split method
- Dataset split partition
- Filtering stages
- Weather conditions



Output

Methodology – Approach

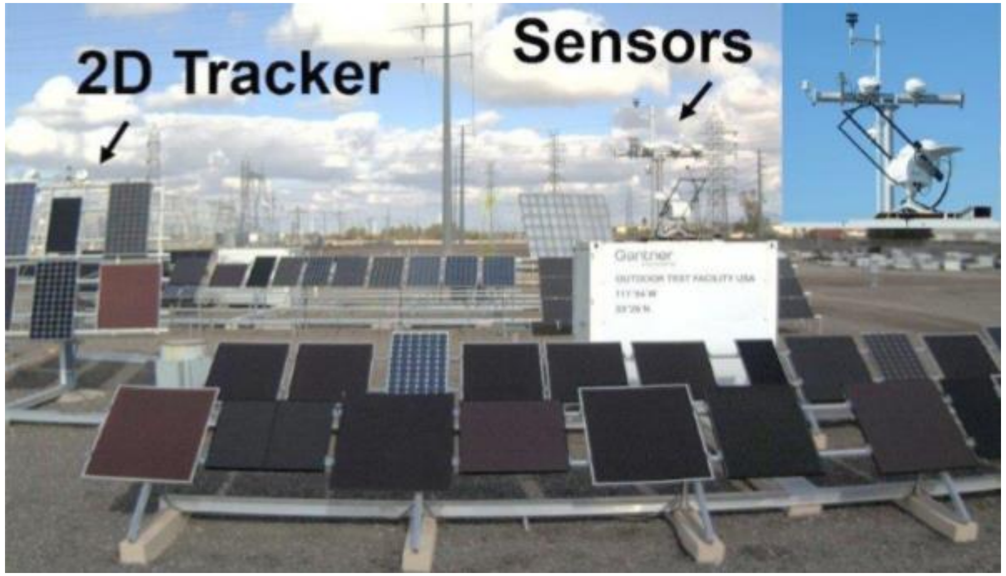


Methodology – Experimental setup

- Recording of meteorological and PV operational measurements (IEC 61724)
- Measurement resolution 1-sec and recording intervals 1-, 15-, 30- and 60-min

UCY OTF – Nicosia, Cyprus

PV String level

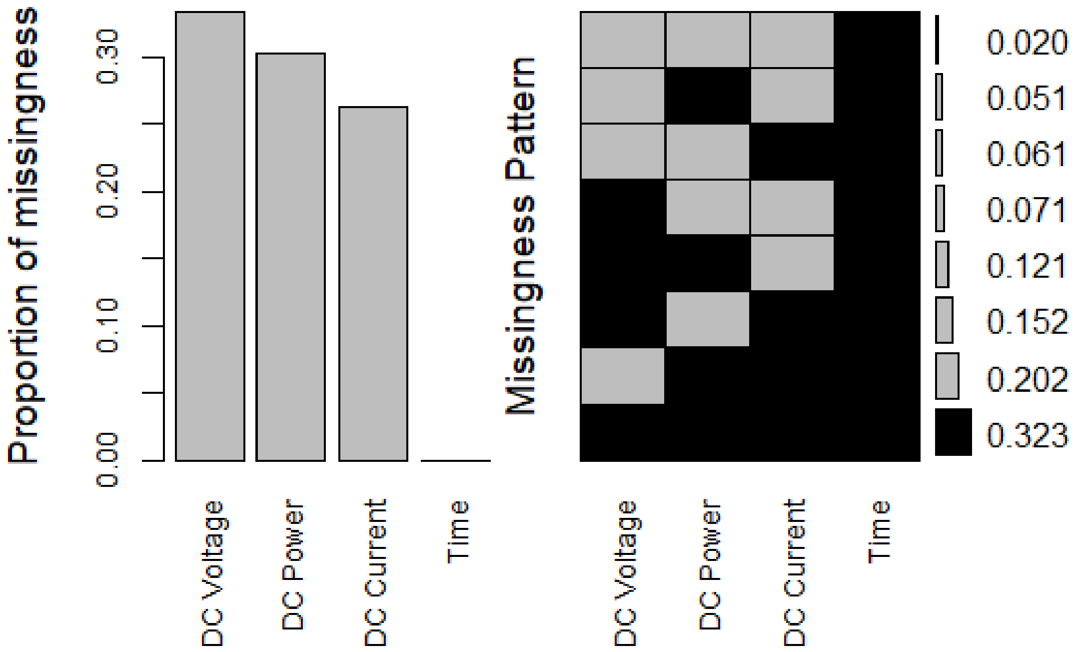


GI OTF –Arizona, USA

PV Module level

Methodology – Data quality routines (DQRs)

- Identification of repetitive data and duplicates
- Identification of missing or erroneous data, outliers and outages
- Correction of erroneous/missing data through data imputation techniques



Methodology – Data quality routines (DQRs)

1. Identification of duplicates
- Check timestamp measurements against known timestamp series
- Check for row measurement duplicates

2. Identification of missing data
- Search for NAN values from the dataset

3. Identification of erroneous data
- Set threshold ranges for:
  - $0 < \text{Irradiance} < 1300\ (W/m^2)$
  - 0 < DC Power < STC power x 1.3
  - 0 < DC Voltage < STC Voltage x 1.1
  - 0 < DC Current < STC Current x 1.25

4. Data filtering
- Night time effects can be removed (e.g. Irradiance < 50 W/m^2)

5. Data correction
- Data imputation techniques for handling erroneous or missing data

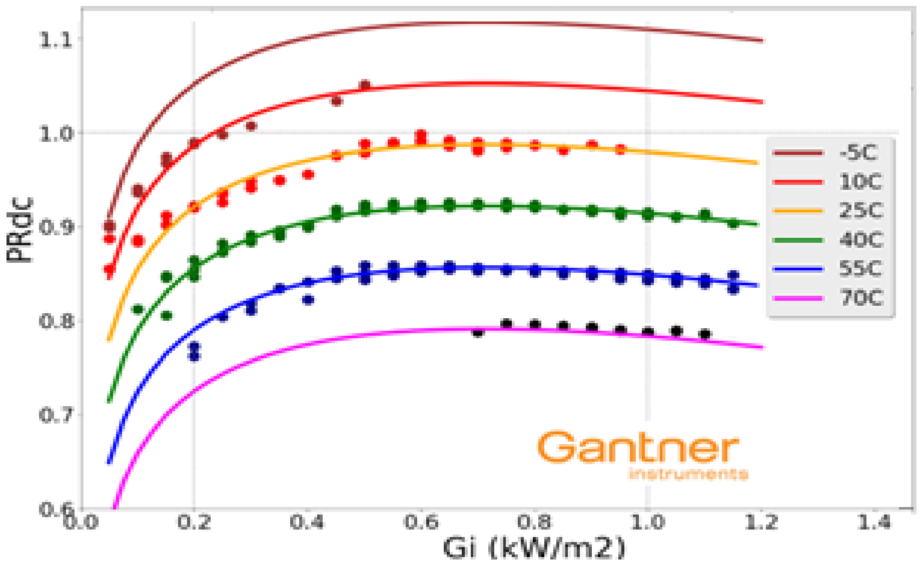
Methodology – Predictive model selection

Empirical

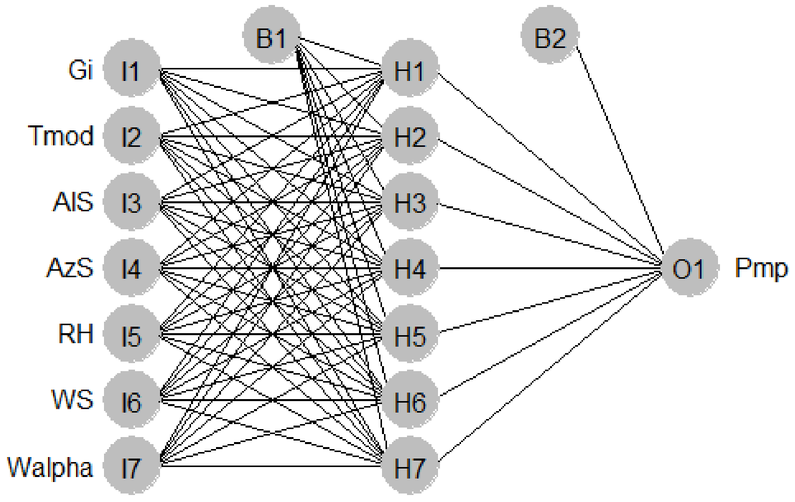
MECHANISTIC PERFORMANCE MODEL ‘MPM’

$$PR = (P_{MEAS}/P_{NOM}/G_I) = C_1 + C_2 * Tmod + C_3 * \log_{10}(Gi) + C_4 * Gi + C_5 * WS$$

P TOLERANCE	GAMMA	LLEC	RS	WIND
%	%/K	%@LIC	%@STC	%/(ms-1)



Machine Learning



Feed-Forward Neural Network (FFNN)

Methodology – Train model and test data

Dataset (1 year of hourly historical actual data)

Dataset

G_I	T_{mod}	*RH*	*WS*	W_{alpha}	*AzS*	*AlS*	*Pmp*

Measured Inputs Calculated Inputs Output

Dataset split method

- Continuous
- Random

Dataset split partition

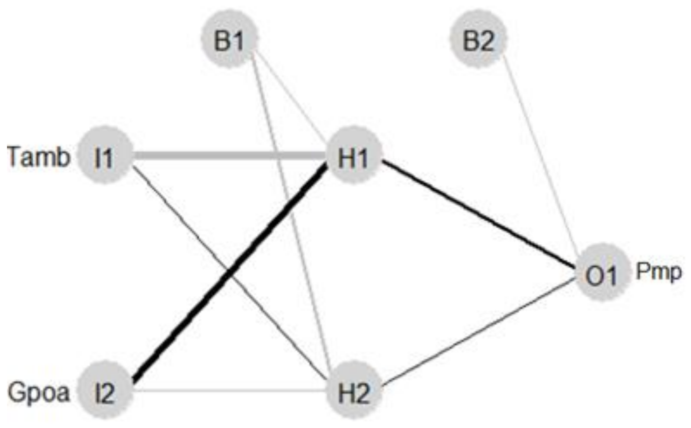
- 70:30% train and test set
- 30:30% train and test set
- 10:30% train and test set

Train set	Test set
Train set 70% (255 Days)	Test set 30% (110 Days)
Train set 30% (110 Days)	Test set 30% (110 Days)
Train set 10% (35 Days)	Test set 30% (110 Days)

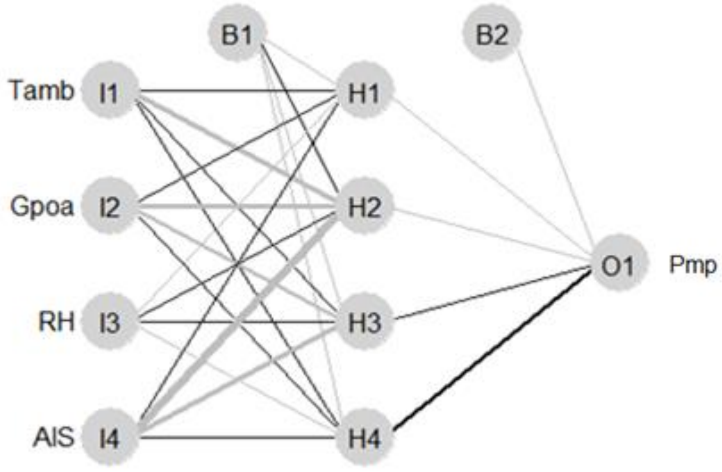
Results – Input features (Machine Learning)

- Machine learning model with measured and calculated features

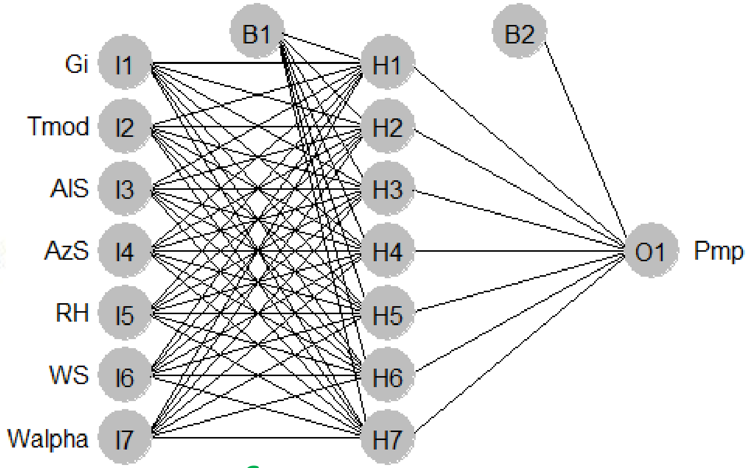
2 Inputs



4 Inputs



7 Inputs



Best performance FFNN

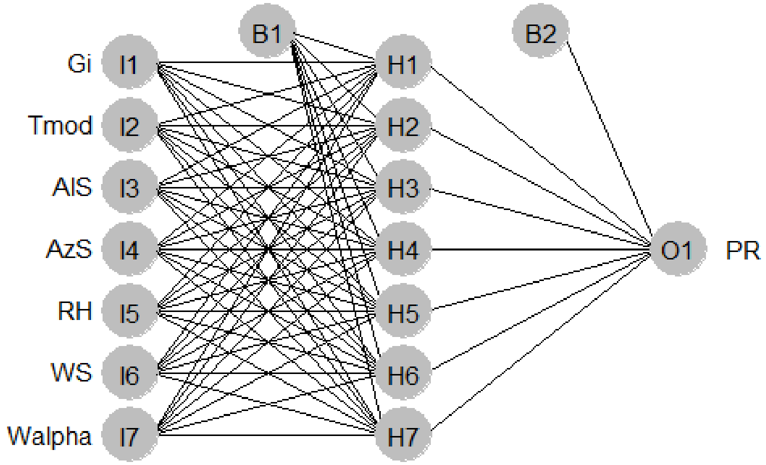
	2 Inputs	4 Inputs	7 Inputs
Random 70:30%	nRMSE 1.18%	nRMSE 1.12%	nRMSE 0.91%
Continuous 70:30%	nRMSE 1.33%	nRMSE 1.13%	nRMSE 0.93%

UCY OTF

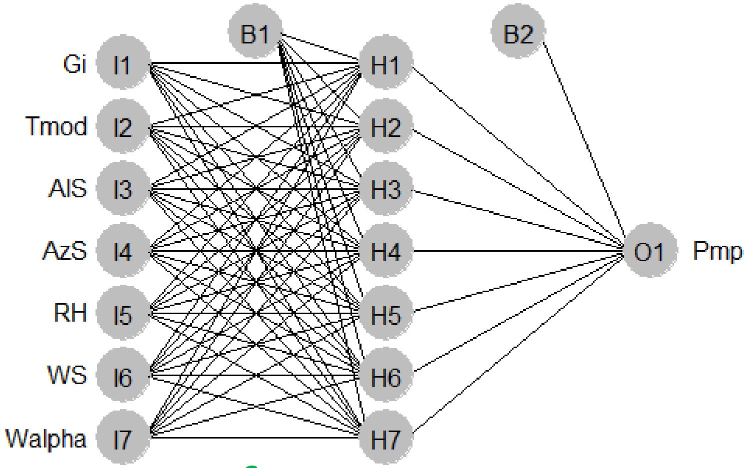
Results – Output features (Machine Learning)

- Machine learning model with measured and calculated features

7 Inputs – PR output



7 Inputs – P_{mp} output



Best performance FFNN

Random 70:30% — **nRMSE 1.30%** — **nRMSE 0.91%**

Continuous 70:30% — **nRMSE 1.33%** — **nRMSE 0.93%**

UCY OTF

Random – Recommended dataset split method

Results – Input features (Mechanistic)

- Mechanistic model with measured and meaningful, orthogonal, robust and normalized features

$$PR = (P_{MEAS}/P_{NOM}/G_I) = C_1 + C_2 * Tmod + C_3 * Log_{10}(Gi) + C_4 * Gi + C_5 * WS$$

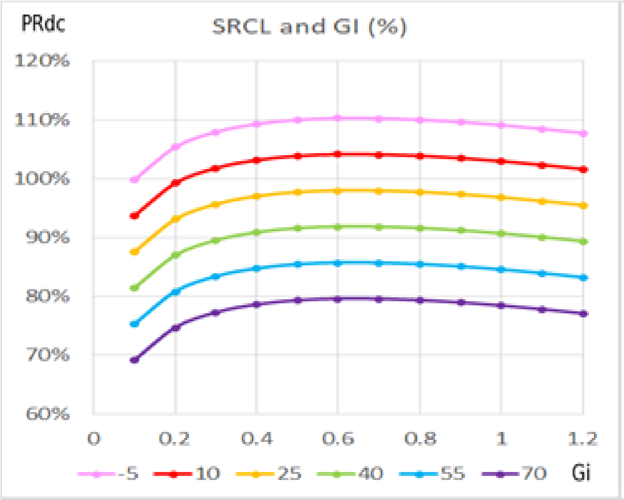
P TOLERANCE	GAMMA	LLEC	RS	WIND
%	%/K	%@LIC	%@STC	%/(ms-1)

Inputs:

- Module temperature (T_{mod})
- Global irradiance (G_I)
- Wind speed (WS)

Requirements for optimal devised model:

- Irradiance Filter ($G_I > 100\ W/m^2$)
- Time Filter (08:00 ≤ Time ≤ 17:00)

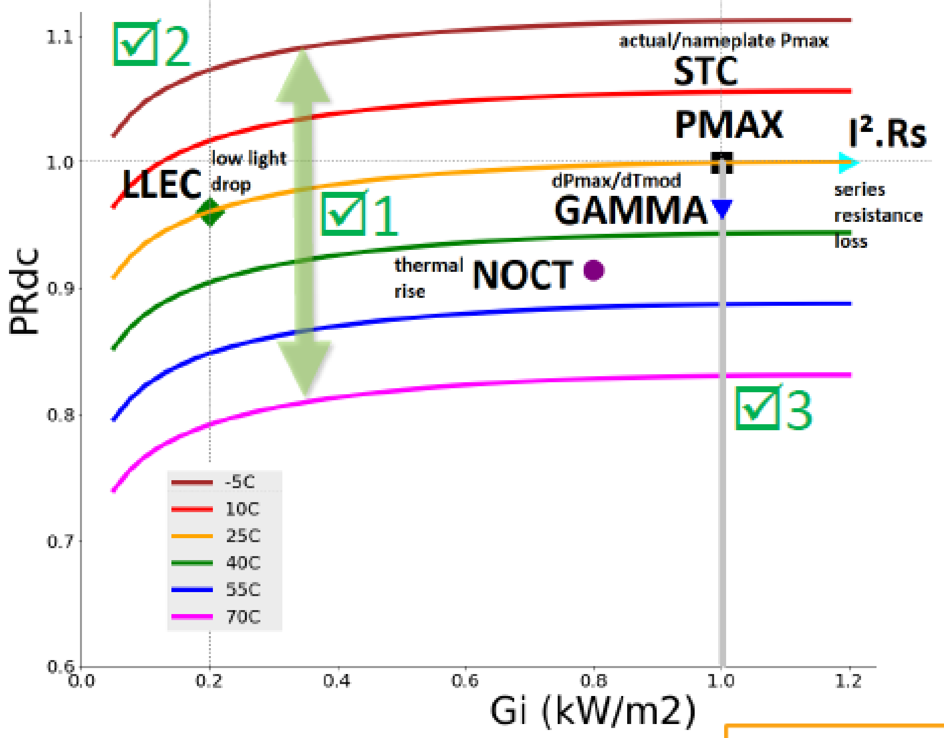


FOSS Research Centre for Sustainable Energy

University of Cyprus PV Technology

Gantner instruments

Results – Input features (Mechanistic)



Illustrating good PV Performance

Uniform vertical ☑1 separation means gamma = constant

Smooth behavipur at lowest ☑2 and highest ☑3 light levels

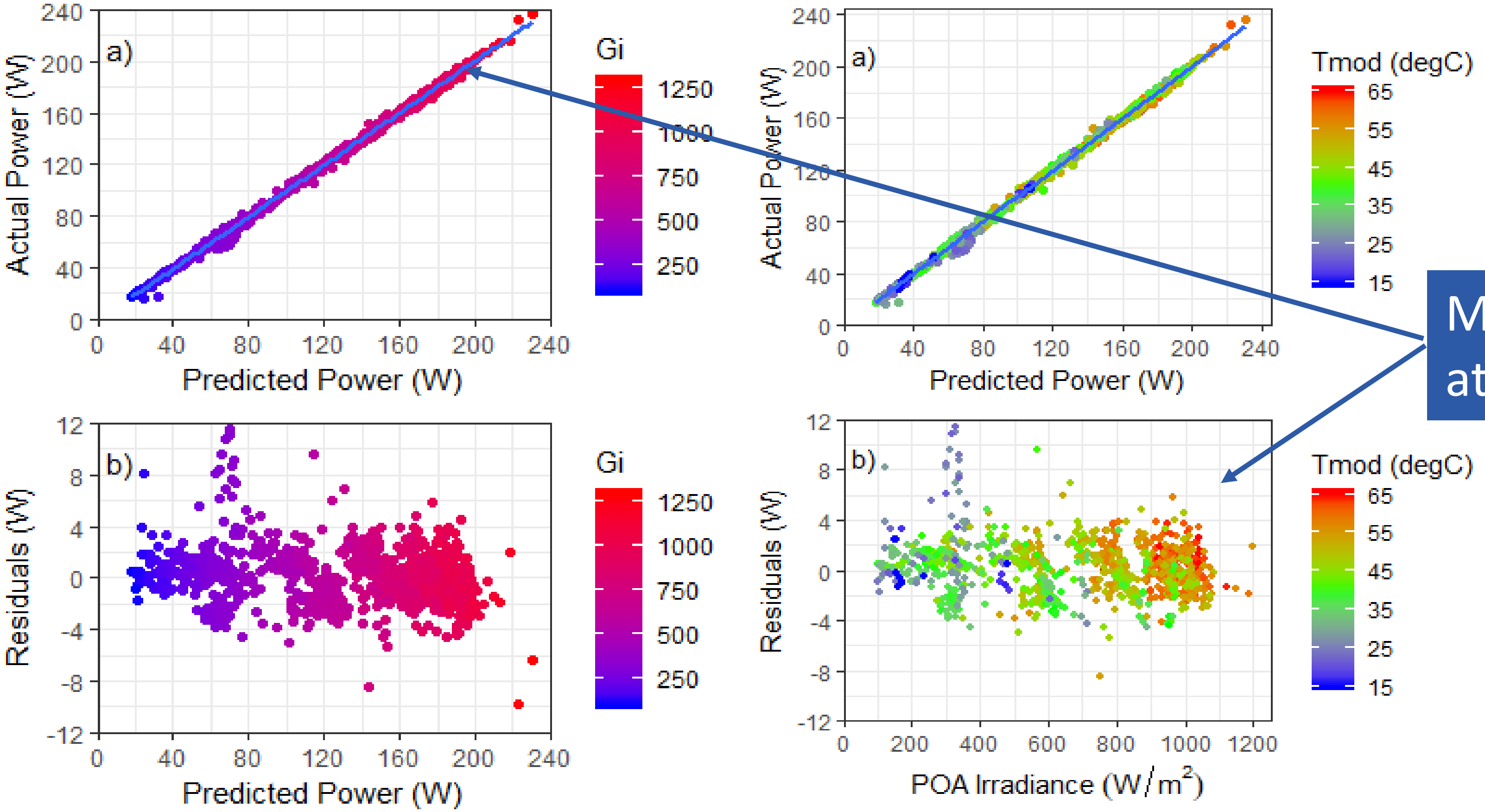
Random 70:30% - GI OTF

Coefficient	Value
C1 (%)	114.09
C2 (%/K)	-0.39
C3 (%)	25.05
C4 (%)	-17.87
C5 (%/ms^{-1})	0.08

$$PR = (P_{MEAS}/P_{NOM}/G_I) = C_1 + C_2 * Tmod + C_3 * Log_{10}(Gi) + C_4 * Gi + C_5 * WS$$

P TOLERANCE (%) | GAMMA (%/K) | LLEC (%@LIC) | RS (%@STC) | WIND (%/(ms-1))

Results – Influence of filtering (Mechanistic)

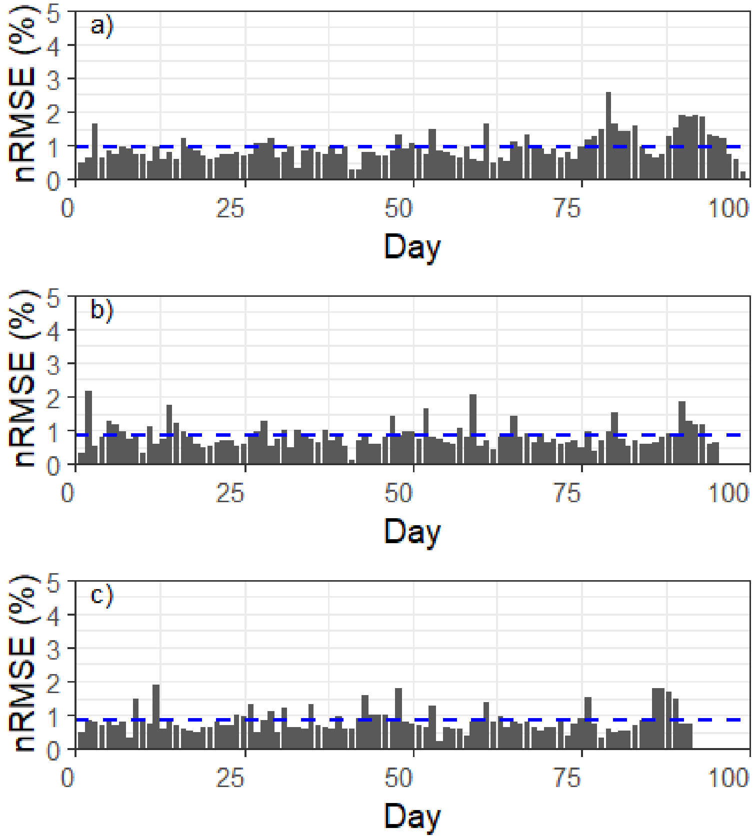


MPM – Improved performance at high irradiance levels

Random 70:30% - GI OTF

Results – Influence of filtering (Mechanistic)

- Filtering at $G_I > 100\ W/m^2$ (GI OTF)



$G_I > 100\ W/m^2$

nRMSE 1.03%

$G_I > 400\ W/m^2$

nRMSE 0.88%

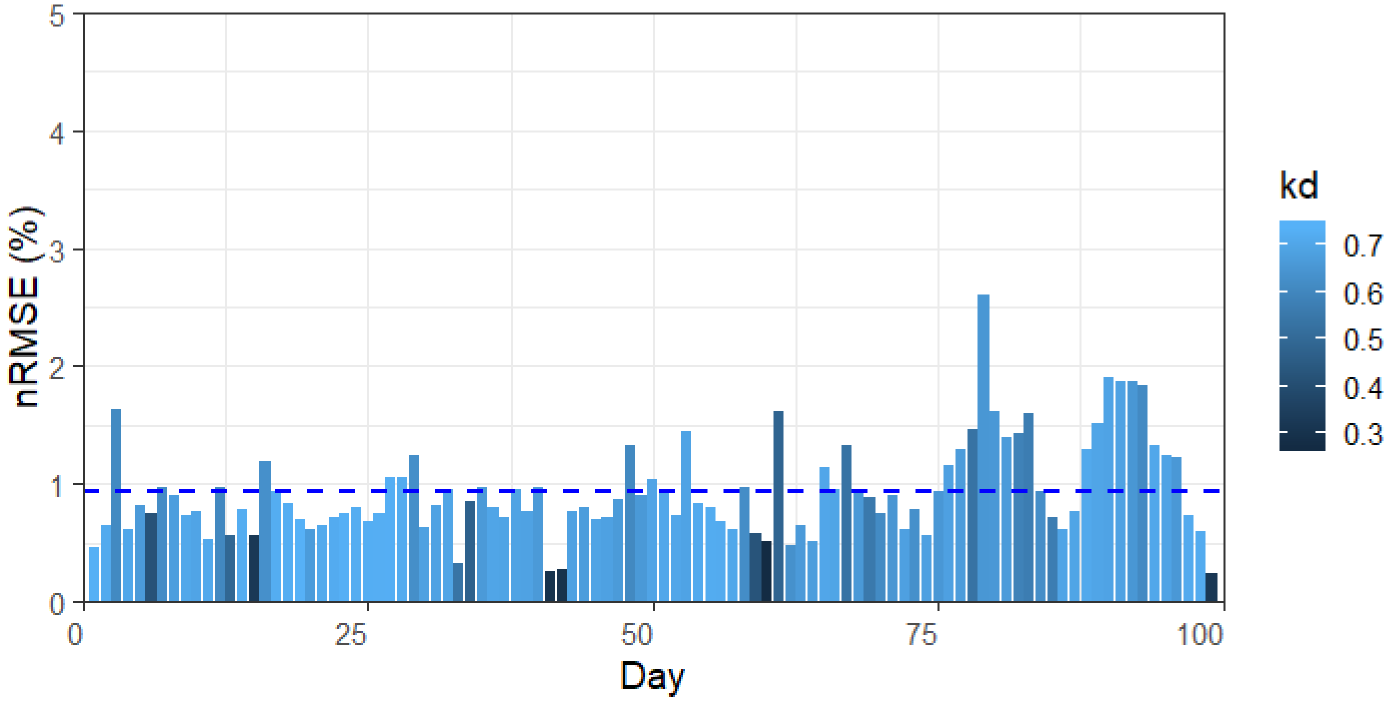
$G_I > 600\ W/m^2$

nRMSE 0.87%

MPM – Higher accuracy by applying irradiance filters (2.15% without any filter)

Results – Influence of filtering (Mechanistic)

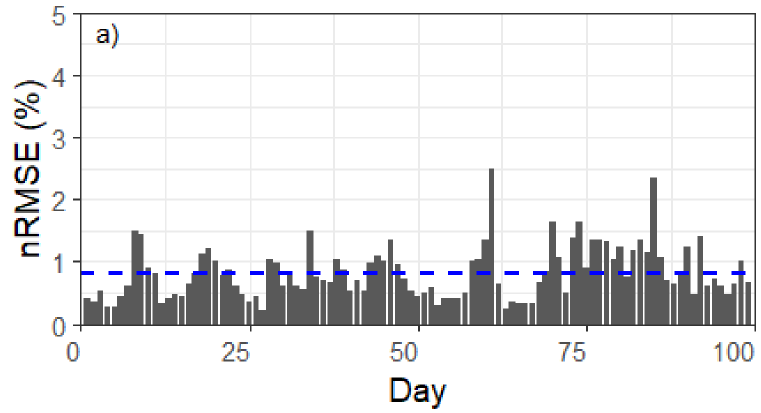
- Filtering at $G_I > 100\ W/m^2$ (GI OTF)



72% of days exhibiting daily nRMSE accuracies below 1% independent of the type of day (clearness index)

Results – Influence of filtering (Machine Learning)

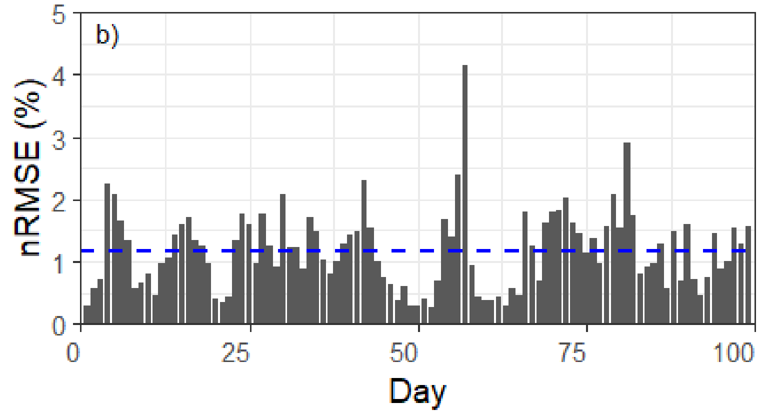
- Filtering at different irradiance levels (UCY OTF)



Without filter

nRMSE 0.91%

ML - Improved performance at increased data for training



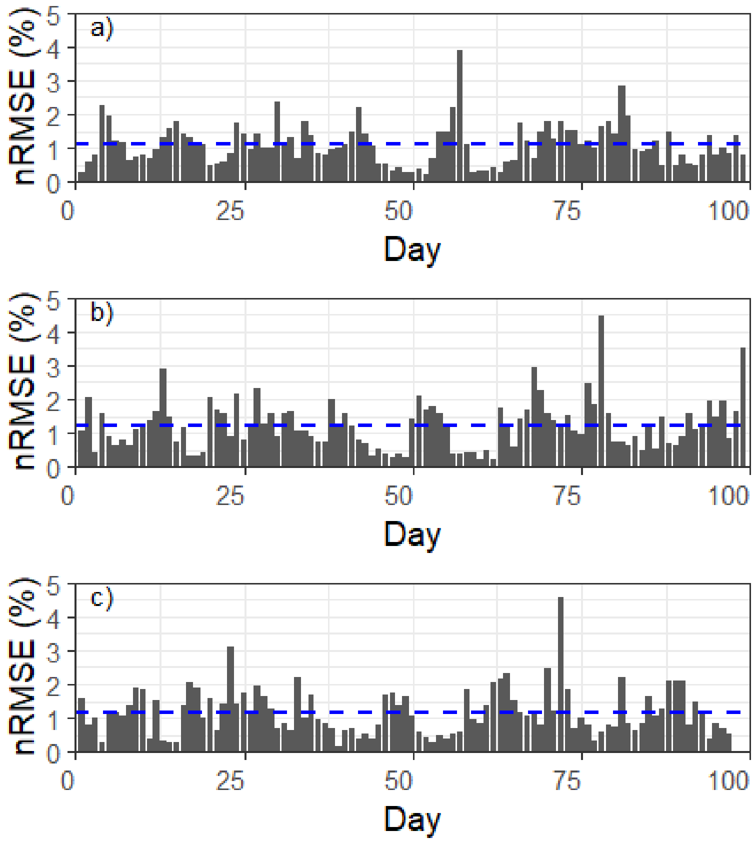
$G_I > 100\ W/m^2$

nRMSE 1.31%

ML – Accuracy not improved by applying irradiance filter

Results – Influence of filtering (Machine Learning)

- Filtering at $G_I > 100\ W/m^2$ (UCY OTF)



$G_I > 100\ W/m^2$

nRMSE 1.31%

$G_I > 400\ W/m^2$

nRMSE 1.36%

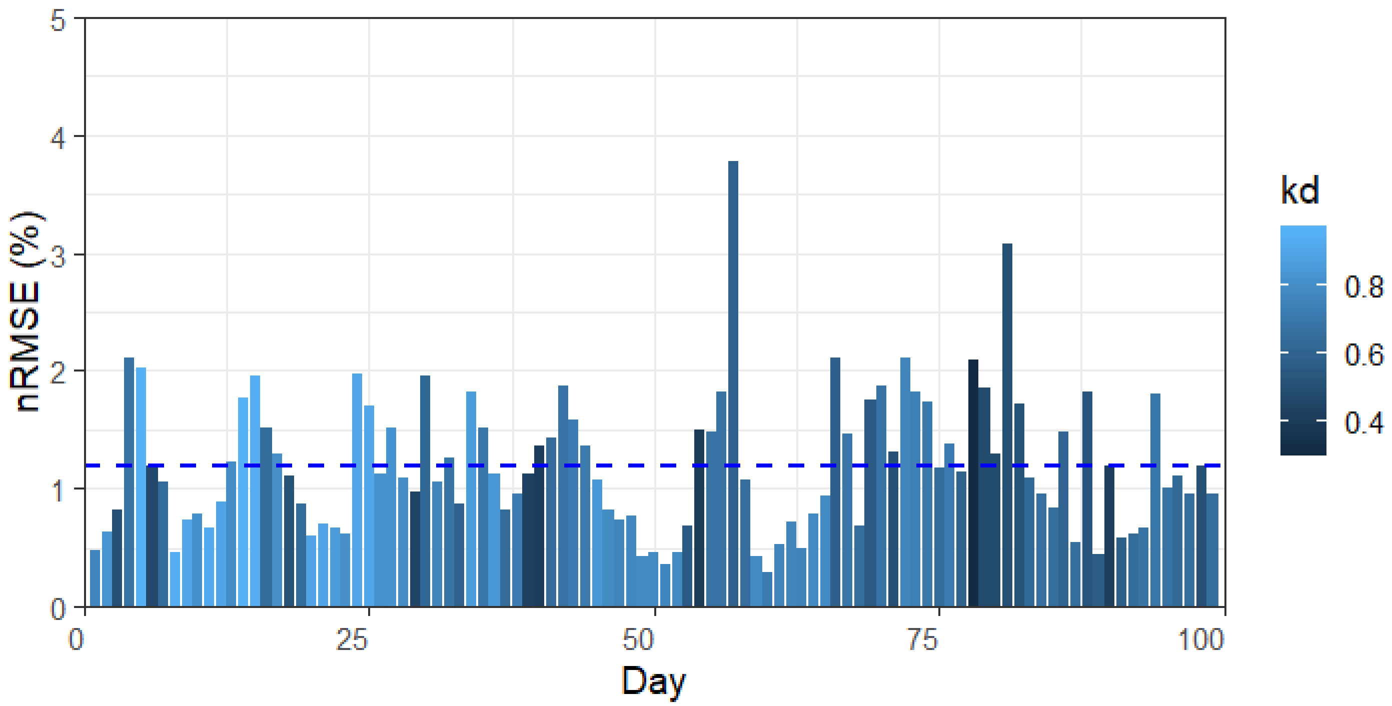
$G_I > 600\ W/m^2$

nRMSE 1.29%

ML – Accuracy not improved by applying irradiance filter

Results – Influence of filtering (Machine Learning)

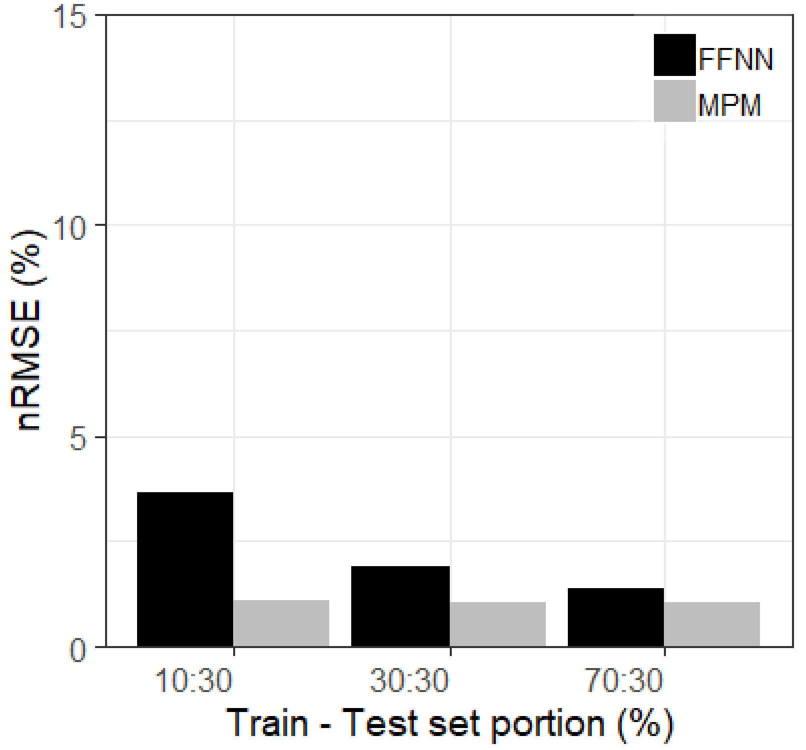
- Filtering at $G_I > 100\ W/m^2$ (UCY OTF)



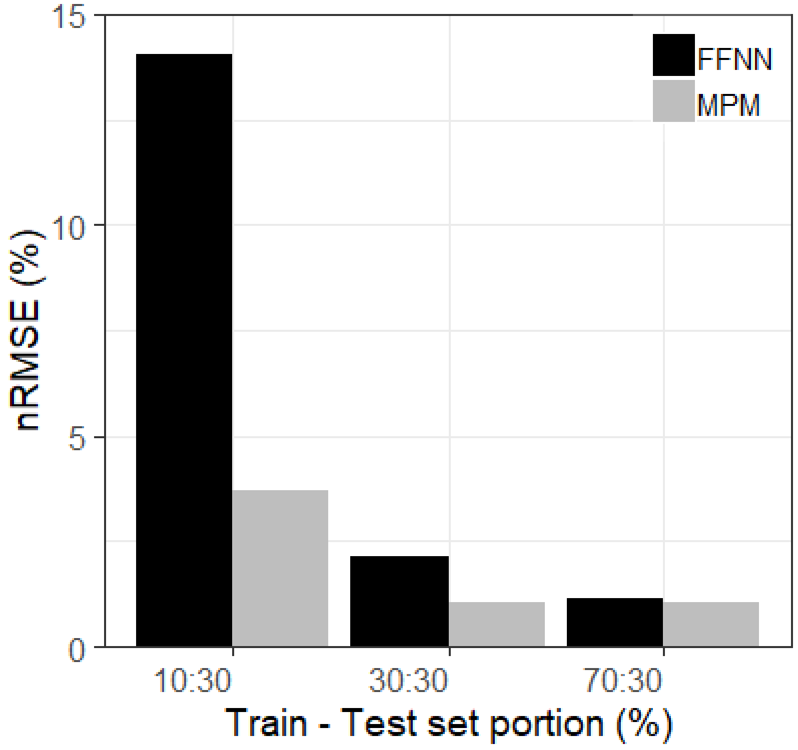
62% of days exhibiting daily nRMSE accuracies below 1.3% independent of the type of day (clearness index)

Results – Dataset split partitions

- Training at different dataset split partitions (10, 30 and 70% of yearly data)



Random GI OTF



Continuous GI OTF

Continuous training – Seasonal errors

Random training - Accurate predictions for both models even at small amount of training data partitions

Summary

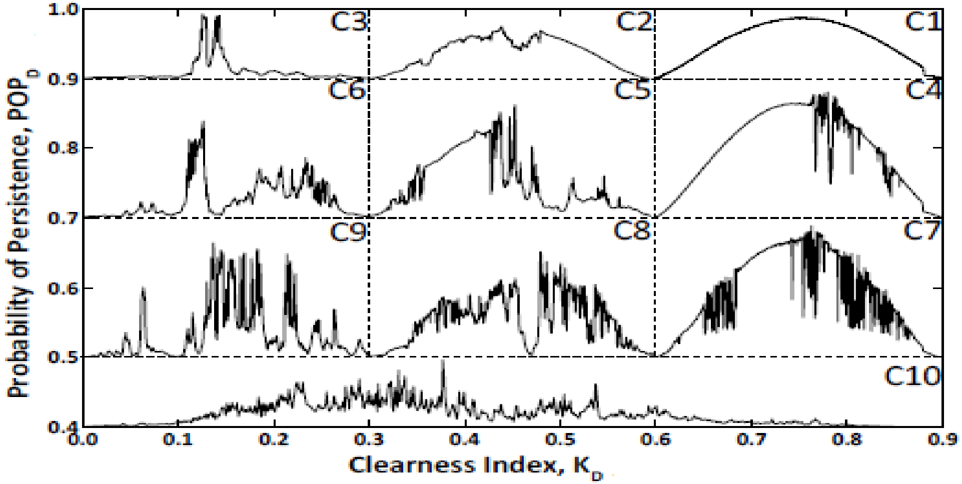
Mechanistic	Machine Learning
• Simple implementation (low complexity)	• Higher complexity for implementation
• Robustness at high irradiance conditions	• Robust at all irradiance conditions only after training at different data combinations
• Irradiance filter improves prediction accuracy	• No data filtering requirements
• Robust model at low duration data set partitions	• Higher training data partitions yield more accurate predictions
• Useful, physically meaningful coefficients	• No direct usable coefficients

Conclusions

- The MPM and the FFNN predictive models were compared in terms of input/output features (model complexity), filtering criteria, dataset split method and partition
- Optimal models: 7 inputs parameter FFNN compared with 5 inputs parameter MPM
- Application of irradiance filter yielded higher predictive accuracy only for the MPM
- Random dataset split method is recommended for both models
- FFNN - Lowest nRMSE of 0.91% for a random 70:30% train/test set approach (UCY OTF)
- MPM - Lowest nRMSE of 1.12% for a random 10:30% train/test set approach (GI OTF)

Next steps…

- Influence of irradiance profiles classification and establishment of minimum requirements for daily weather classification
- Further improvement of the MPM (spectral and AOI corrections) for more accurate predictions
- Benchmarking on several PV systems installed at different locations



Class	k_d	POP_d	Description
1	$k_d \geq 0.6$	$POP_d \geq 0.9$	High Quantity High Quality
2	$0.3 \leq k_d < 0.6$	$POP_d \geq 0.9$	Medium Quantity and High Quality
3	$k_d < 0.3$	$POP_d \geq 0.9$	Low Quantity High Quality
4	$k_d \geq 0.6$	$0.7 \leq POP_d < 0.9$	High Quantity Medium Quality
…	…	…	…
8	$0.3 \leq k_d < 0.6$	$0.5 \leq POP_d < 0.7$	Medium Quantity Low Quality
10	-	$POP_d < 0.5$	Very Low Quality

Thank you for your attention

Marios Theristis, PhD

PV Technology Laboratory
University of Cyprus
Email: theristis.marios@ucy.ac.cy

Website: www.pvtechnology.ucy.ac.cy
www.gi-cloud.io