Parabolic trough solar field condition assessment using raw time series data and artificial intelligence

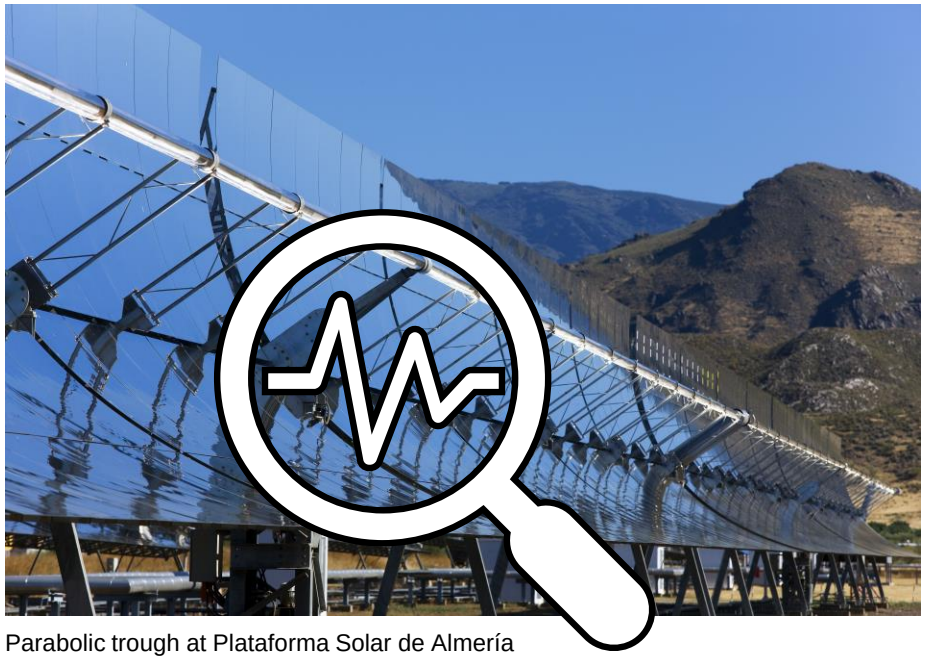
Alex Brenner, Tim Kotzab, Tobias Hirsch

SolarPACES 2025

September 25, 2025

Agenda

- Solar field condition
  - Deterministic models vs. data-driven approaches
  - Which data describes the current plant condition?
- Methods
  - Raw time series data
  - Learning representations from raw time series data
- Results
  - Visualization of learned representations
  - Application: Representations used for anomaly detection
  - Application: Representations compared to soiling values
- Conclusions
  - High potential for data-driven models as universal pre-processing for raw time-series data



Parabolic trough at Plataforma Solar de Almería
(Owned by the Spanish research center CIEMAT)

Solar field condition
Deterministic models vs. data-driven approaches

Power plant layout

Meteo data

Losses

Additional data …

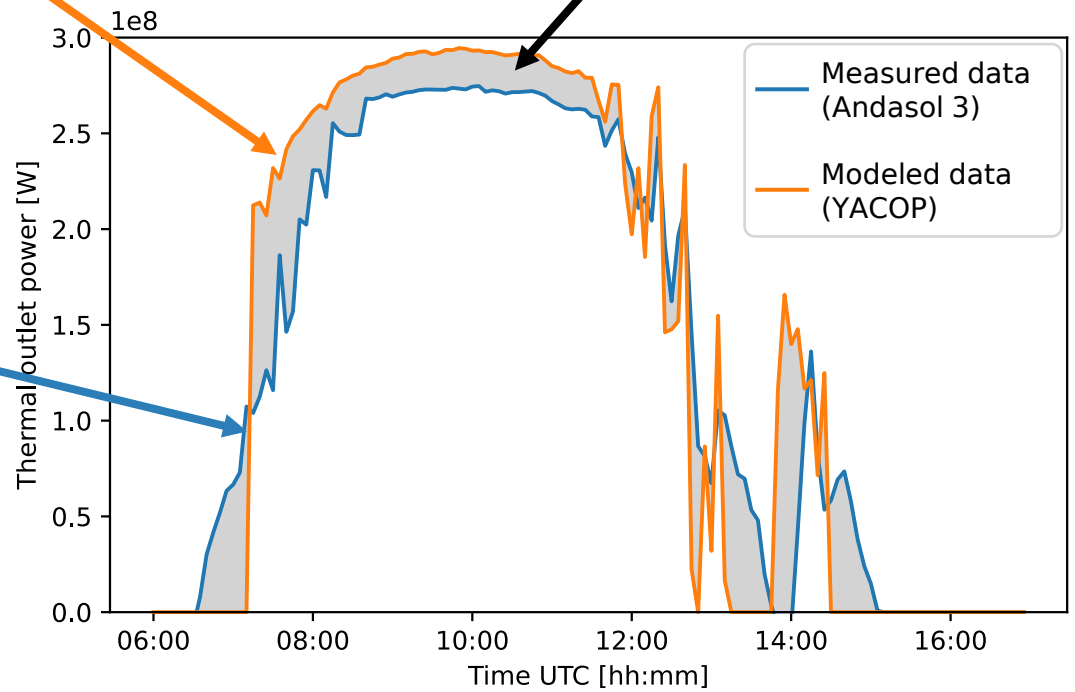
Deterministic power plant model, e.g. YACOP [1]

Gap between is model error + **missing information about solar field condition**

- Option 1: Add more condition measurements
- **Option 2: Use existing time series data + AI to extract the condition**



The real world: Andasol 3

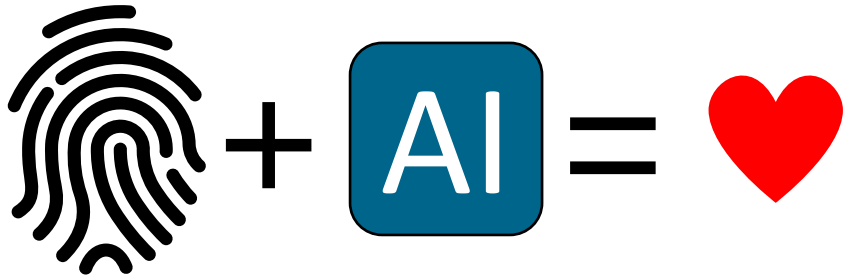


Solar field condition
Which data describes the current plant condition?

Importance

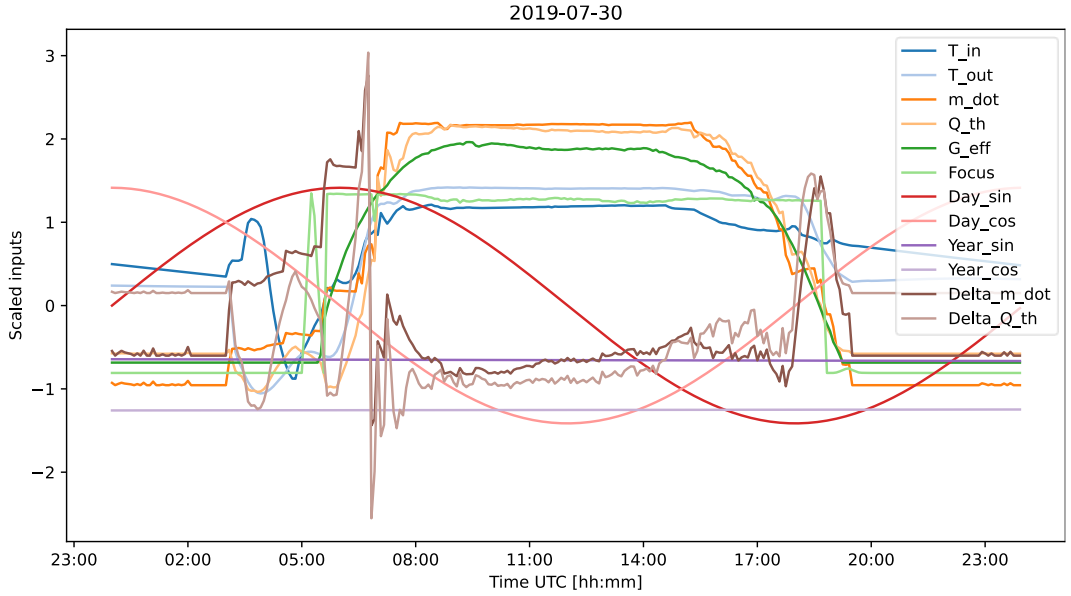
Effect	Operation data
Soiling	▪ Thermal outlet power reduced $\dot{Q}_{th,out}$ ▪ Mass flow and HTF temperature affected
Dumping	▪ Thermal outlet power reduced $\dot{Q}_{th,out}$ ▪ Reduced mass flow, reduced focus
Loop under maintenance	▪ Thermal outlet power reduced $\dot{Q}_{th,out}$ ▪ Very low HTF temperature ▪ No tracking
Heat loss	▪ Thermal outlet power reduced $\dot{Q}_{th,out}$ ▪ Reduced HTF temperature ▪ Possibly mass flow affected ▪ Fast cooling during off-sun periods
Faulty tracking	▪ Thermal outlet power reduced $\dot{Q}_{th,out}$ ▪ Reduced HTF temperature ▪ Possibly mass flow affected

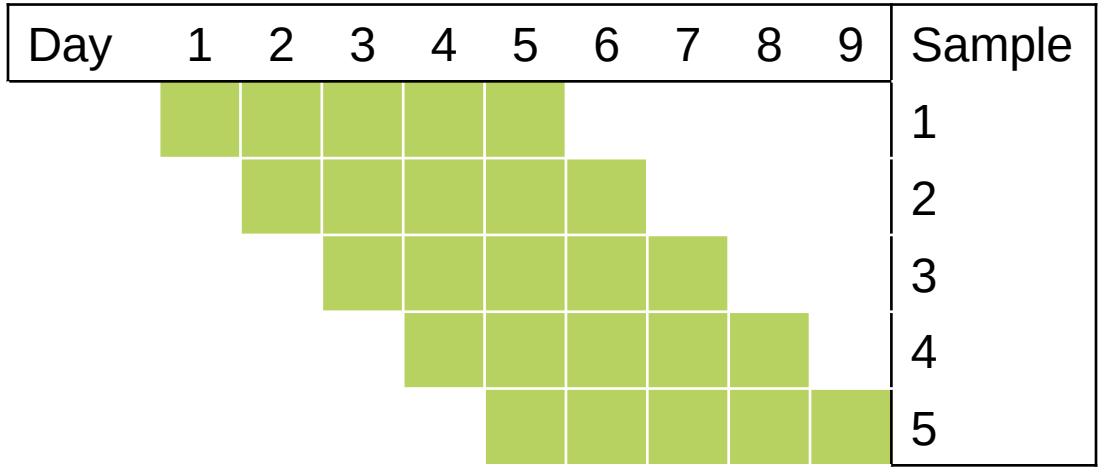
All effects
- Reduce thermal outlet power
- Are affected by the irradiance
- Occur in different forms
- **Leave patterns in the measurement data**



Methods
Raw time series data

- Raw time series data from Andasol 3 power plant (2013 to 2021)
  - Solar field inlet, outlet temperature
  - Inlet mass flow
  - Thermal outlet power
  - Average focus of solar field
  - Effective irradiance
  - Time of day and time of year
- Difference to deterministic model YACOP
- Scaled with mean and std. deviation
- Sliding window dataset with 5 consecutive days
- Split into train, validation and test dataset

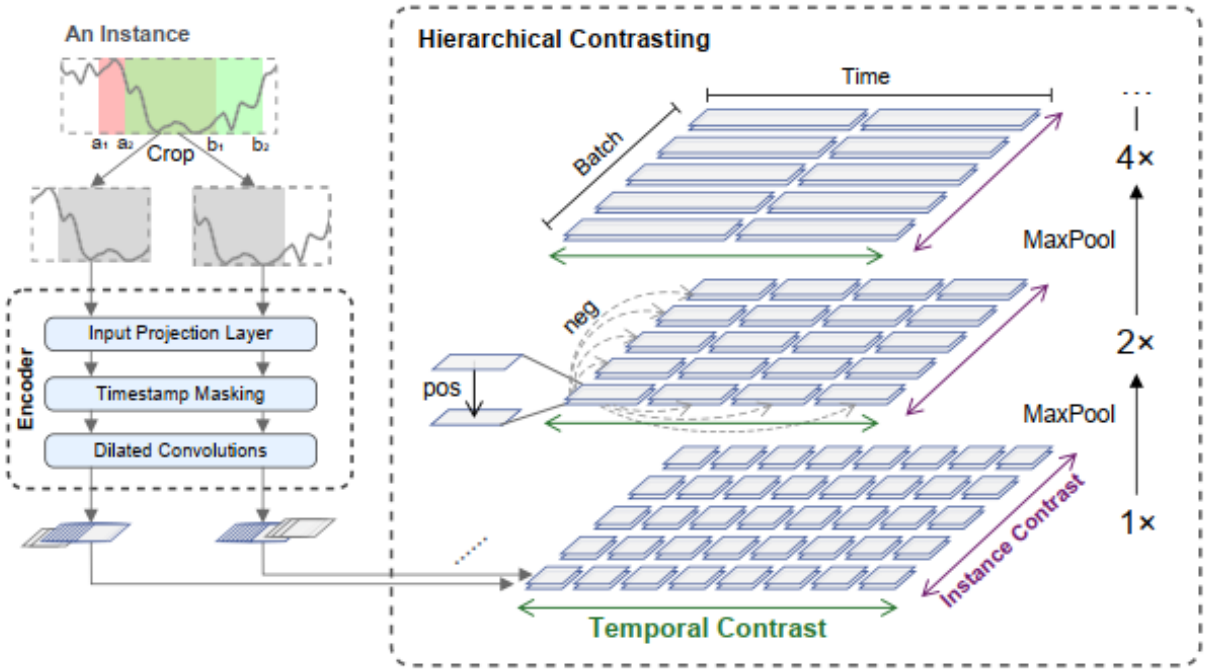




Methods
Learning representations from raw time series data

- Autoencoder model
  - Long-short-term-memory (LSTM) model
  - Self-supervised
  - Single representation for each day
  - Slow training, hard to train

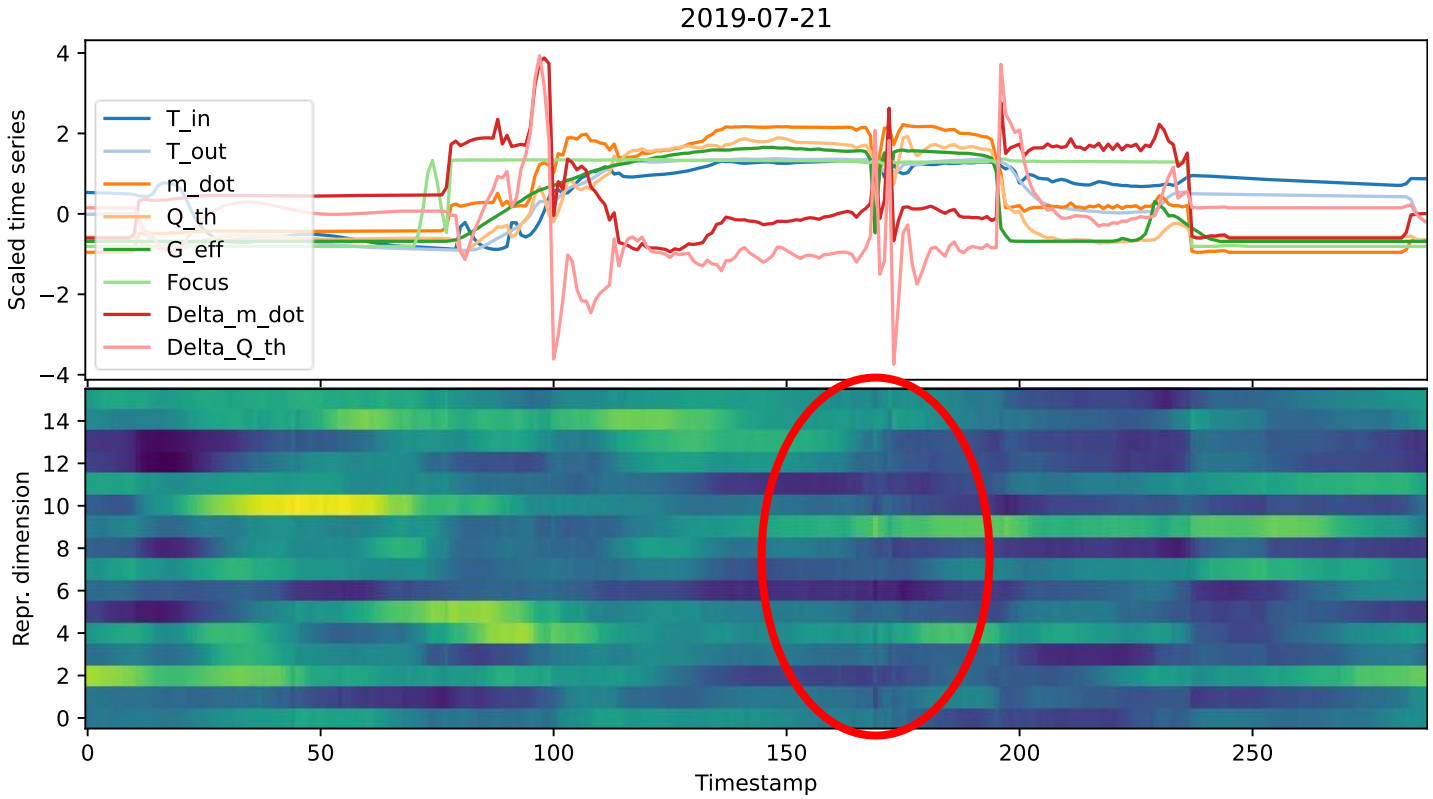
- TS2Vec model [2]
  - Self-supervised
  - Contrastive learning
  - Flexible representations (e.g. for each time step, or entire sample)
  - Fast training even with large dataset (years of data)



from Yue et al. 2021 [2]

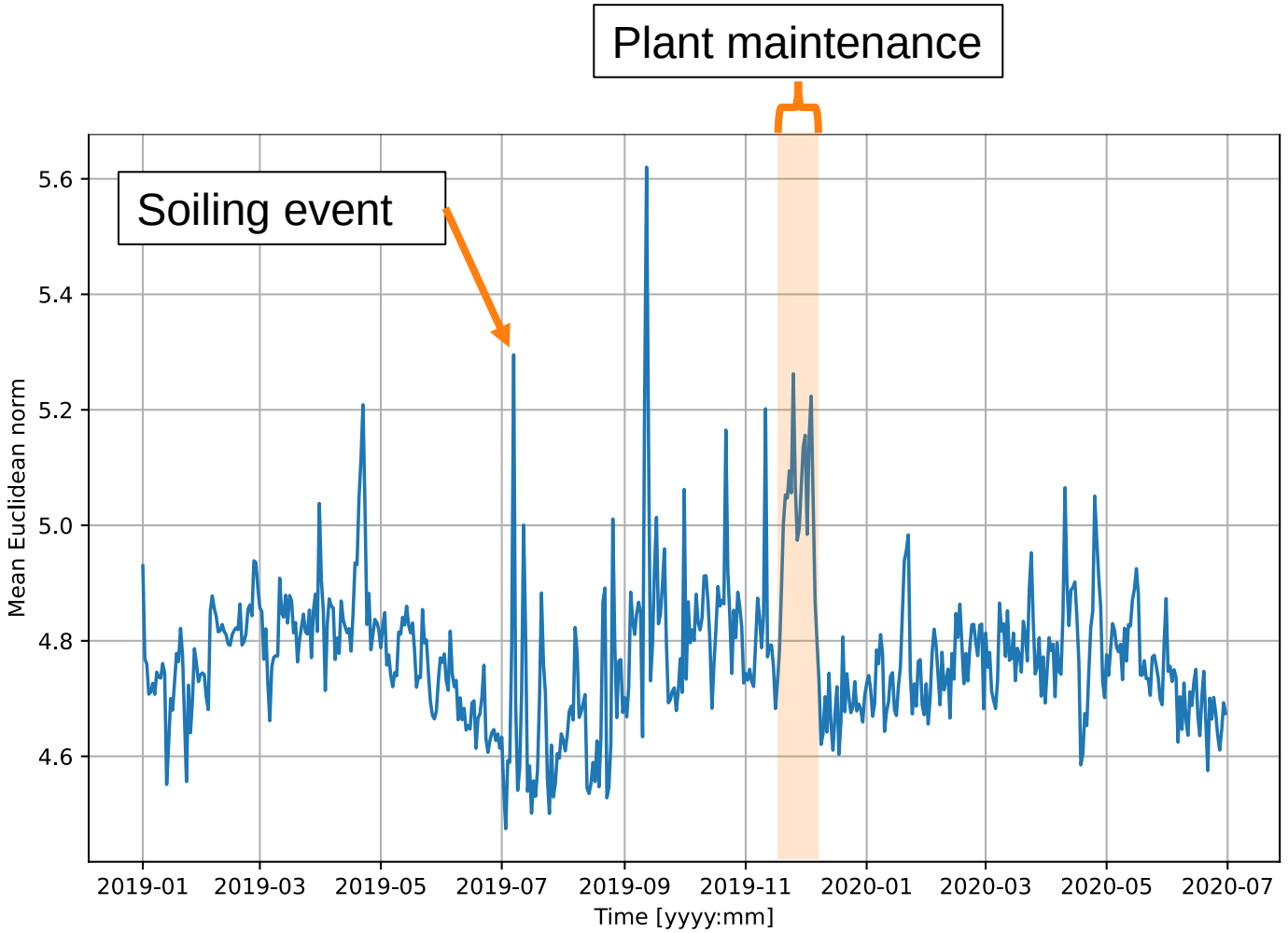
Results
Visualization of learned representations

- Selection of representations with max. variance
- Start-up and cool-down are visible in the representations
- Interruptions in the irradiance are visible as well
- **The major benefit of the representations comes from other downstream tasks**
  - Anomaly detection
  - Classification
  - Input for other models



Results
Application: Representations used for anomaly detection

- Similarity between different representations can be measured with the Euclidean norm
  - $EN(x_{d,t}) = \sqrt{\sum_{i=1}^{n}(x_{d,t,f})^2}$
  - Similar operation days → similar representations → similar Euclidean norms
- High peaks in the Euclidean norm correspond to unusual evens
  - Soiling event with high dust loads
  - Plant maintenance
- **High potential for an anomaly detection use-case**

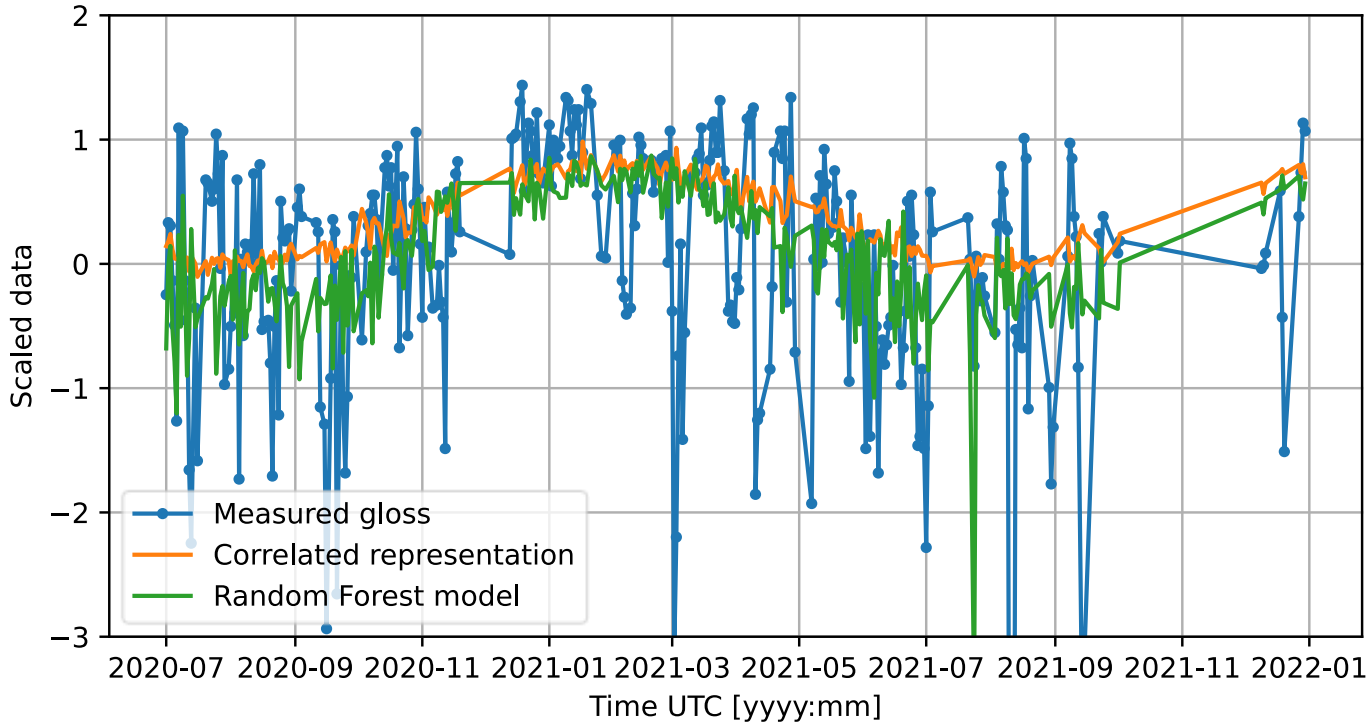


Results
Application: Representations compared to soiling values

- Gloss values are taken in the Andasol 3 plant as measure for soiling
- High seasonality of soiling is covered by the representations
- A simple Random Forest model using the five most correlated representations further improves the prediction

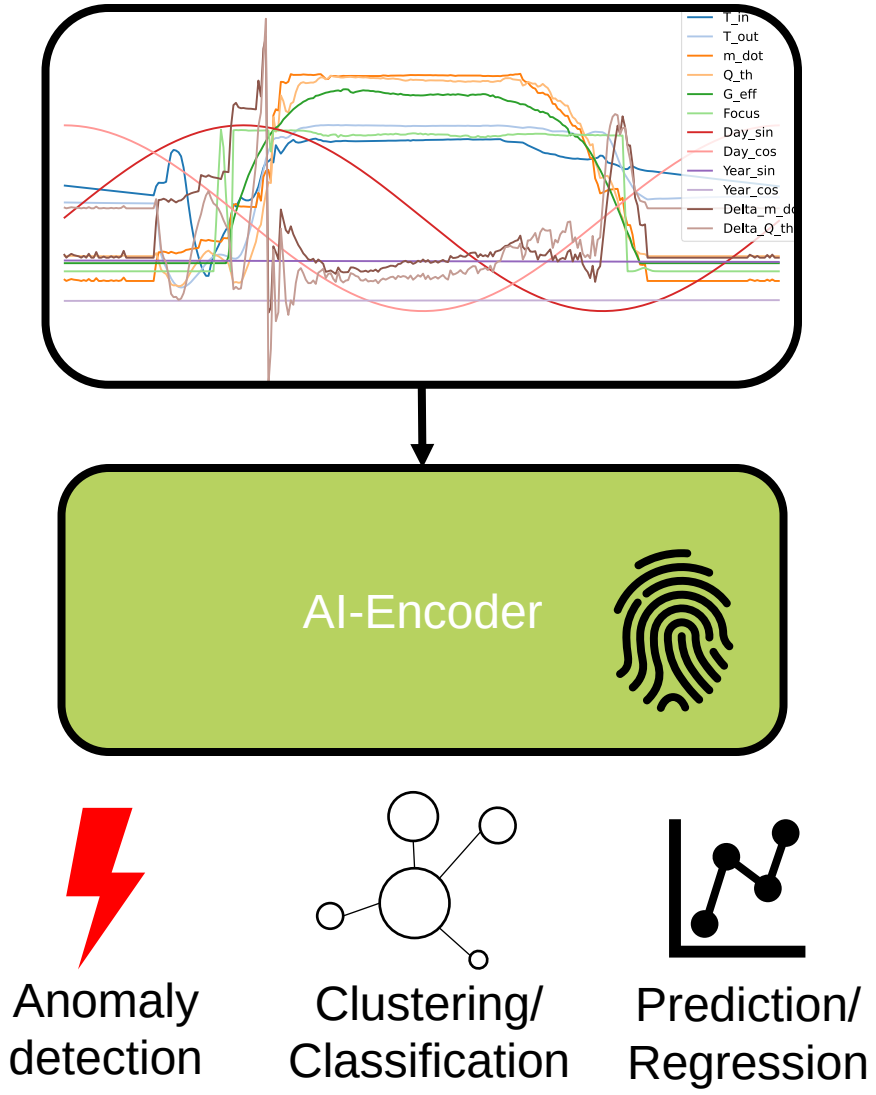


[3]



Conclusions

- Data-driven approaches have a high potential as universal pre-processing step for raw time series data
- Downstream tasks can benefit from the representations of the pre-processing step
- Paves the way for other data-driven models
  - Anomaly detection
  - Time series classification/clustering
  - Prediction models

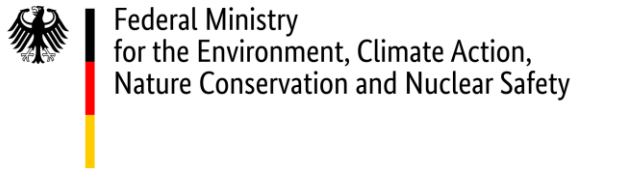


More project results will be presented in the ATA Insights webinar **“Digital solutions for monitoring and optimizing CSP and PV plants”** on Wednesday, **November 19, 3–5 p.m. CET**

THANK YOU FOR YOUR ATTENTION

Alex Brenner | alex.brenner@dlr.de

The authors would like to thank Marquesado Solar for providing the data as part of the AuSeSol-AI project.



DLR

References

[1] Loevenich et al., YACOP - A Fully Flexible Modeling Environment for the Techno-Economic Assessment of Energy Systems. Kölner Sonnenkolloquium 2024, 2024-06-26, Jülich, Deutschland. https://elib.dlr.de/206463/

[2] Yue et al., TS2Vec: Towards Universal Representation of Time Series, AAAI 2021. http://arxiv.org/pdf/2106.10466v4

[3] Gloss measurement with Zehntner ZGM 1110, Source Zehntner GmbH