PV Forecasting Case Studies

Petrina Papazek, Pascal Gfäller, Irene Schicker
GeoSphere Austria – Analysis and Model Development/Post-Processing, Hohe Warte 38, Vienna

correspondence: Petrina.papazek@geosphere.at

Funded by the European Union — **Destination Earth** — Implemented by ⬤ECMWF ⬤esa ⬤EUMETSAT

PV location: short / inconsistent solar power time-series ⚠️

➡️ **PP / semi-synthetic data:** continuous, coherent forecasts across all relevant parameters

➡️ **AI / learning:** post-processing and AI powered forecasting

today's contribution: put to test at DE_330 scenarios

Purpose & Goals

Energy/PV operators need reliable nowcasting to intra-day forecasts tailored to sites...

We study ...

AI based forecasts
- **objective:** PV nowcasting (short-range)
- **init-time:** UTC 6:00-16:00 (solar day), daily+hourly
- **testing:** 2023
- **training/validation** period for AI: 2010 – 2022
- data/model optimized to each location

using a ...

synthetic data generator
- **objective:** generate past time series at PV site
- **testing: 2022 (**2023)
- synthetic data generation: 2013-2017
- training period AI: 2021-2020
- input data sources: ERA-5, CAMS, (observation)
- output: AROME surface, PV

driven by various :

DATA SOURCES: analyses + forecasts

UNLIMITED
- **>> ERA5-land:** hourly reanalysis in high resolution (surface)
- **>> CAMS:** radiation time series of the ADS data store (satellite prod.)
- **>> OBS/TAWES:** observation obtained at close observation sites

REDUCED
- **>> AROME:** high-resolution NWP (numeric weather model) forecasts
- **>> IrradPhyDNet:** in-house grid based nowcasting for irradiances
- **>> PV:** observation at production site

2023 DE_330 CASES

📍 **Denmark (DK):** selected sites (#11)
📍 **Austria (AT):** trafo level sites (#8)

<u>AIM</u>: up to + 6h ahead, 15 min. update frequency
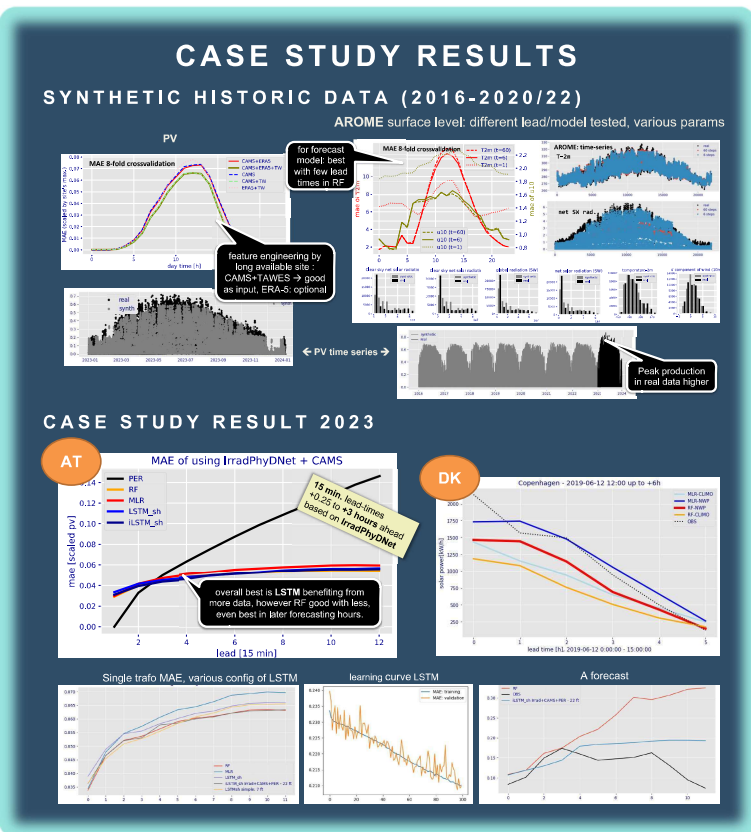
Outlook

semi-synthetic + location based predictions are a win in point optimized PV forecasts when utilizing AI/deep learning in complex terrain / difficult situations

... in DE_330 we aim for:
- **operationalization** / **triggered** for events + more locations
- including **completely synthetic scenarios** driven by meta-data as needed

... we need to improve:
- diurnal / seasonal **peak production**
- **forecast range** /extend

CASE STUDY RESULTS

SYNTHETIC HISTORIC DATA (2016-2020/22)

AROME surface level: different lead/model tested, various params

for forecast model: best with few lead times in RF

feature engineering by long available site : CAMS+TAWES → good as input, ERA-5: optional

← PV time series →

Peak production in real data higher

CASE STUDY RESULT 2023

AT — MAE of using IrradPhyDNet + CAMS

15 min. lead-times +0.25 to +3 hours ahead based on IrradPhyDNet

overall best is LSTM benefiting from more data, however RF good with less, even best in later forecasting hours.

DK — Copenhagen - 2019-06-12 12:00 up to +6h

Single trafo MAE, various config of LSTM | learning curve LSTM | A forecast

POST-PROCESSING METHODS

synthetic data generation | feature selection, clearance
interpolation+climatology+persistance | site tailored model optimization
computational performance / GPU | AI/deep learning forecasts

synthetic data generator :
a set of random forest models use time-series of related data sources (**X**) to predict data of a reduced source (**Y**)

TRANSFORMATION (X,Y):
set of predictions **P**
- >> consistent resolution
- >> matched lead time
- >> normalized

post-processing by AI/LSTM :
AI/ANNs, e.g.: sequence-to-sequence LSTM (long short-term memory) serve as a method learning diverse background forecast models **P** to give a PV forecast **Z**

ADDED VALUE
- ✓ tailored to single PV sites
- ✓ both real and synthetic PV
- ✓ fast computation by AI
- ✓ beneficial in complex terrain and events
- ✓ to be extended to various foecasting ranges, init time and forecasting range
- ✓ able to operate on top of IrradPhyDNet's forecasts
- ✓ triggering + on demand

(i) INPUT: AUXILIARY
- IrradPhyDNet
- CAMS rad. time-series
- ERA5/CERRA
- met. OBS (TAWES)

META INFOS
- location: lat, long, altitude
- module, peak pv
- azimuth
- declination
- time: doy, hour

(I.) >>DATA GENERATION: training

PROCESSING
- pvlib computation
- feature selection
- normalization: none, min-max, robust
- interpolation, imputation
- time step recursion

MODELS
- RF
- MLR
- ANN

(ii) TARGET: REDUCED
- **PV** output
- 15 minute updates

semi-synthetic data [for all target features]

(II.) >>DATA GENERATION: prediction

once our AI is trained can be quickly run for triggered events

Z: post-processed forecast

AI / LSTM

P$_I$:climatologies + P$_{II}$:interpolated NWP + P$_{III}$:persistence
IrradPhyDNet

GeoSphere Austria
Oct. 2024