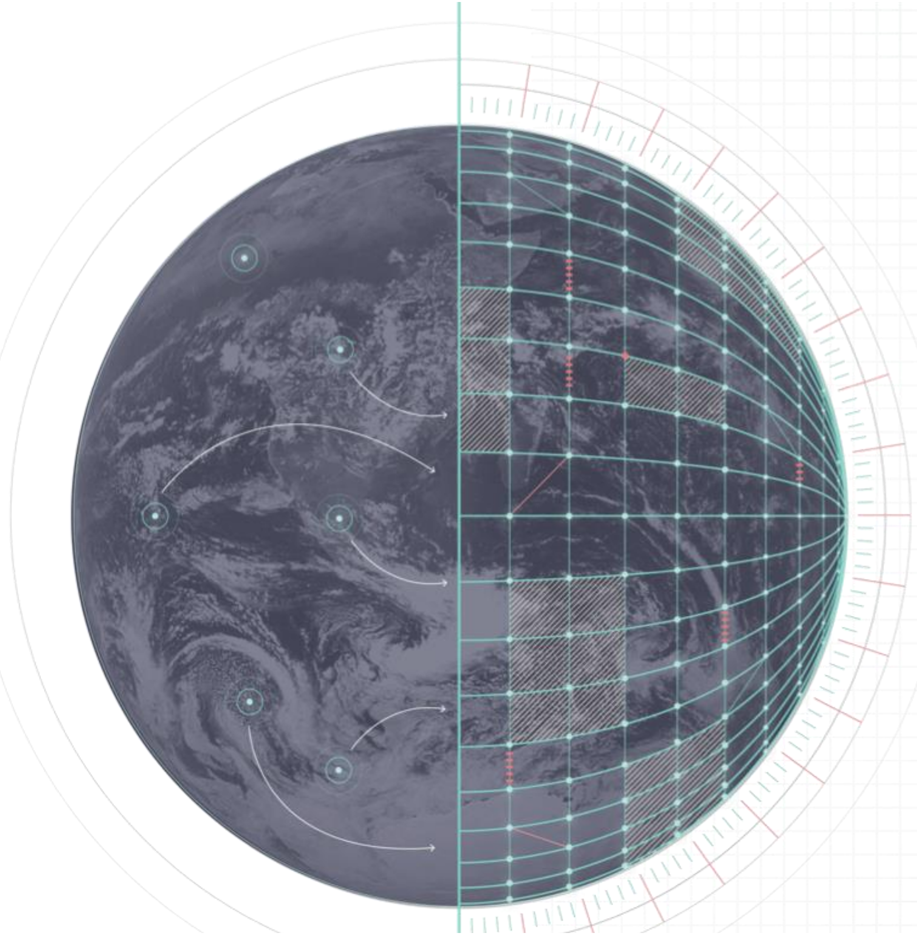


DESTINATION EARTH

Extremes in wind speed and power: post-processing approaches in DE_330 for known and unknown wind farm configurations

Post-processing Austria and Croatia
(Irene Schicker, Petrina Papazek, Pascal Gfäller, Iris Odak Plenkovic, Ivan Vujec, Kristian Horvath)

User Meeting Bonn, 13. & 14 November, 2023, DE_330

Funded by the European Union Destination Earth implemented by ECMWF esa EUMETSAT

Extremes, wind energy, post-processing – How To?

- What extremes do we consider?
  - High persistent wind speed exceeding defined thresholds
  - High persistent wind speed exceeding defined thresholds and large spatial coverage
  - Gustiness (EFI index for wind gusts)
  - Ramping

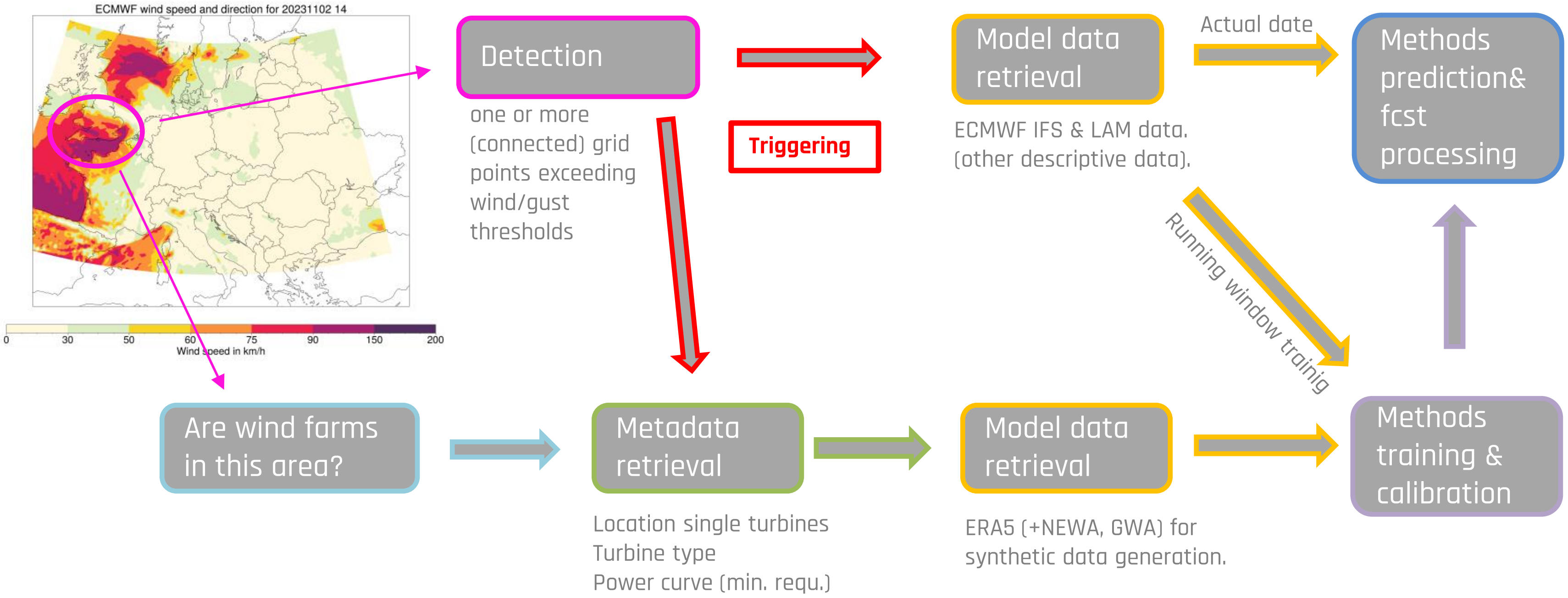
- What do we need for wind energy post-processing?

- How does the planned triggering workflow work?

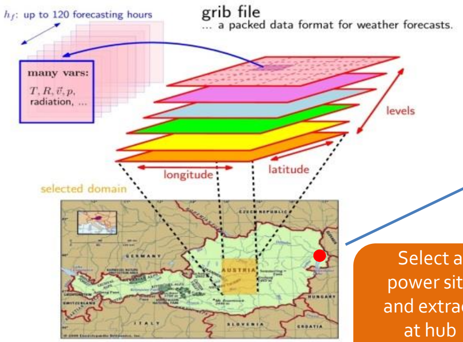
- Use cases, lessons learned, user integration & outlook

Workflow post-processing in DE_330- How To?

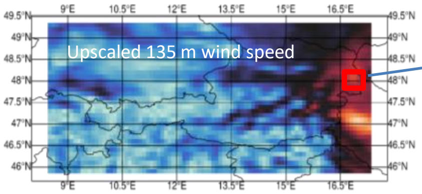
Storm Ciarán, forecast ECMWF IFS 1.11.2023, 00UTC, for 2.11.2023 13UTC

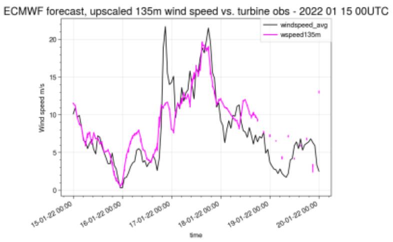


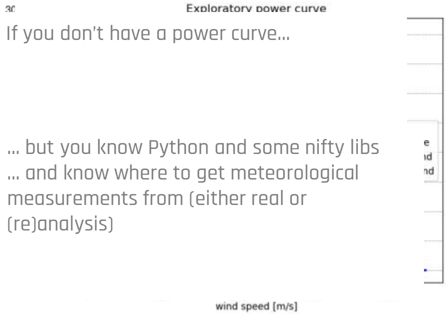
Importance of metadata – short



Select a power site and extract at hub height

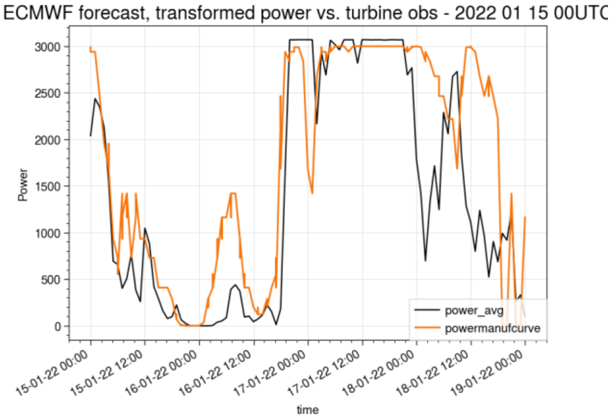


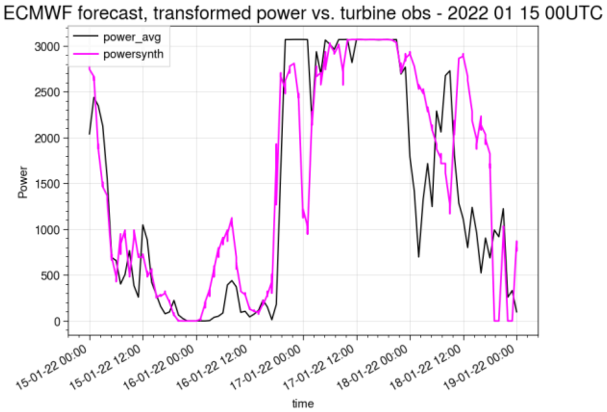


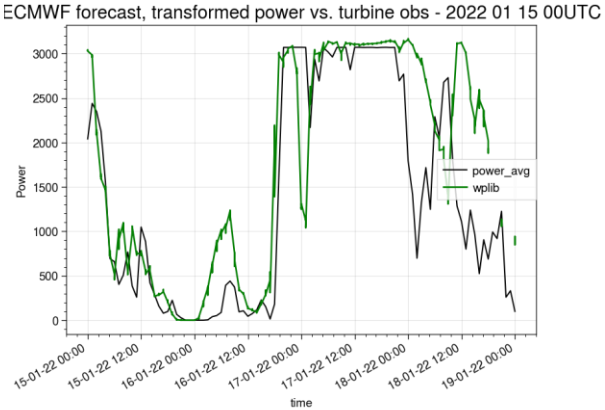


If you don't have a power curve…

… but you know Python and some nifty libs
… and know where to get meteorological measurements from (either real or (re)analysis)







Methods – statistical post-processing

Approach

- The analog method: requires a **relatively moderate-size dataset (usually a year-long)**.
- The Kalman filter algorithm: the recursive one, **so the training period can be relatively short (e.g., a few weeks). It requires solid-quality almost real-time measurements**.

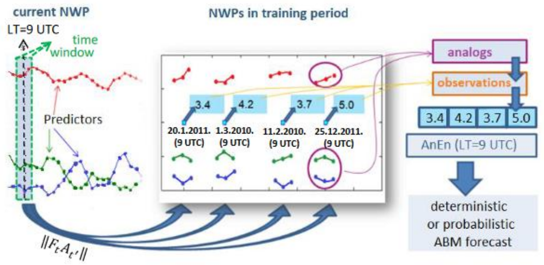
- Input data:
  - NWP model output data
  - **measurements of target variable**

- Target: wind speed/power data

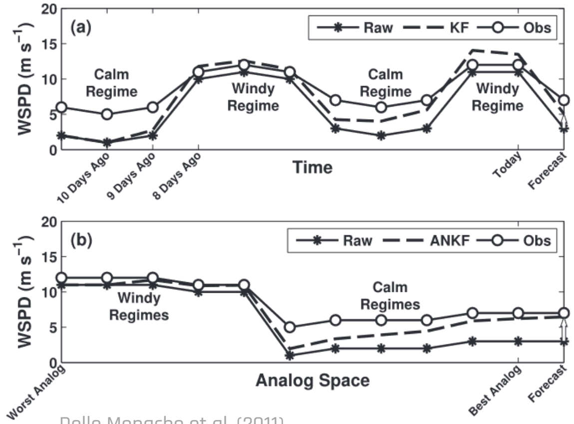
Methods

- **Analog method**: based on the search for the most similar forecasts in the training period. The verified observations of analogs are used to generate new forecast. Additional improvements upon the starting analog-based forecast:
  - weight optimization
  - correction for high wind speed.

- **The algorithm inspired by the Kalman filter:** uses recent past forecasts and observations to improve current raw forecast. By optimizing some of the algorithm parameters, additional forecast improvements can be shifted to more extreme events.



Odak Plenković (2020)



Delle Monache et al. (2011)

ECMWF

Methods – machine learning based post-processing

Approach:

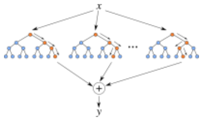
- Running window: using the last 45 – 90 days to train a model
- *Future, generalized : (semi-synthetic) historical production data, multi-year with attention which will be re-fitted with every use-case/triggered case*
- Features:
  - produced power, synthetic (real) of the past day (2-hours)
  - NWP model data: ECMWF-IFS and/or LAM
  - Non-linear feature generation
- Target: (semi-)synthetic (real) power data, wind speed at hub height if needed/wanted

Methods – multi-model ensemble:

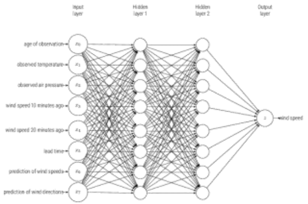
- Multi-linear regression with non-linear features

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \varepsilon$$

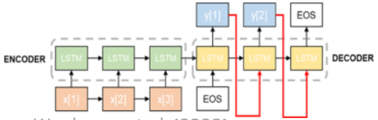
- Random forest



- XGBoost regression
- Fully connected multilayer perceptron neural network



- Currently evaluated: sequence-to-sequence LSTM model



Woohyun et al. (2020)

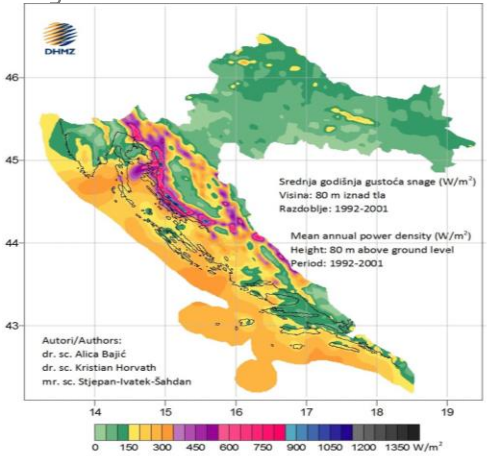
Use case 1 and 2 - Croatia

The area

Complex coastal terrain prone to downslope bora windstorms - the main wind energy resource in the area

Exceptional temporal and spatial bora wind varibility reaching hurricane 5 scale gusts

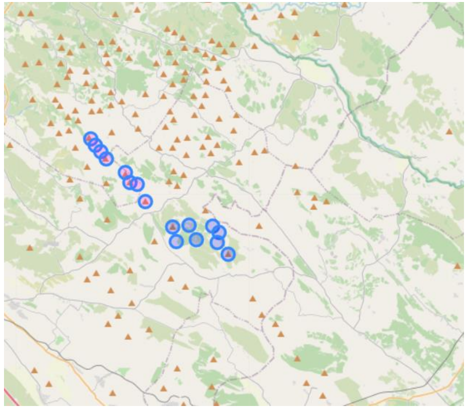
Excellent test area for VHR modelling experiments & demonstrations related to strong winds



Case wind energy - location

Next step application of methods to synthetic wind farm data (wind speed, power)

Description Croatia target region – results to come

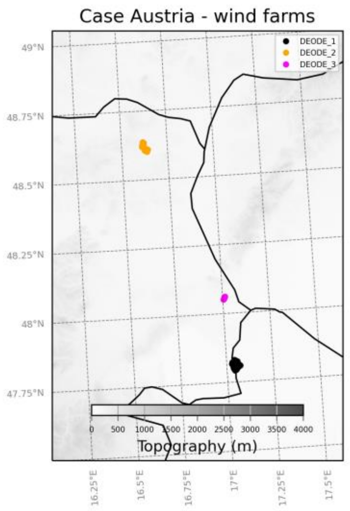


Name	# turbines	type	hub height	Ø	Cut-in/cut-out [m/s]
CRO LOC 1	8	Siemens SWT 2.3	80	93	3.5/25.0
CRO LOC 2	8	Siemens SWT 2.3	100	101	3.0/20.0

ECMWF

Use case 2 – (Northeastern) Austria

Triggering – event + wind farms in region of event



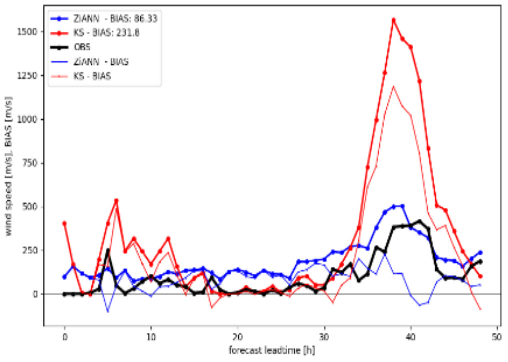
Model data retrieval

ECMWF IFS & LAM data. (other descriptive data).

ERA5 (+NEWA, GWA) for synthetic training data generation.

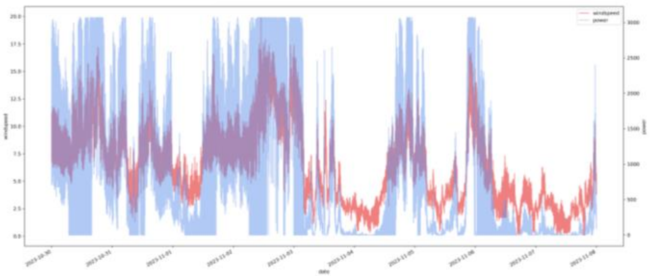
Model pipeline:

- pre-processing input data (feature generation), save in run_init file for training and prediction
- training the multi-model ensemble on all synthetic data types
- Predict multi-model ensemble



Name	# turbines	type	hub height	Ø	Cut-in/cut-out [m/s]
DEODE 1	38	Enercon E101	135	101	2.0/29.1
DEODE 2	14	Vestas V90	105	90	2.5/28.0
DEODE 3	8	DeWind D6	92	62.3	2.5/23.0

DEODE, real data



ECMWF

Lessons use cases and what is missing

Lessons learned:

- Metadata retrieval, i.e. location and wind turbine specifications, is a hart nut to crack esp. if no direct connection to owner/maintainer exists.
- Synthetic data generat … sting service) for local (hardware speaking) ge …

Workflow needs to be faster for shorter warning periods!

- For fast and on-the fly predictions, methods need to be able to either generalize or be „simple“ methods able to work with only a few training data.

Missing:

- Coupling with c_p/c_t curve estimation developed by FMI.
- Wakes in wind farms: add „parametrization“ in post-processing for wakes.
- Fast access to ERA5 data for the generation of synthetic data.
- Fast access to historic ECMWF/LAM forecasts for model training.

User integration

What are the specific needs of the users, how can we engage communication to not end up in our scientific bubble?

- How much in advance to we have to detect/trigger? → This defines our forecast horizon
- What temporal frequency is needed? 10min, 15min,..
- How do we overcome the metadata issue?
- What's the incentive to engage? Do we need to better highlight the benefits of this work?
- Feedback on predictive skills/usefulness of the (possible) users: how valuable is the model? What is the added value?
- Forecasting (workflow) game?

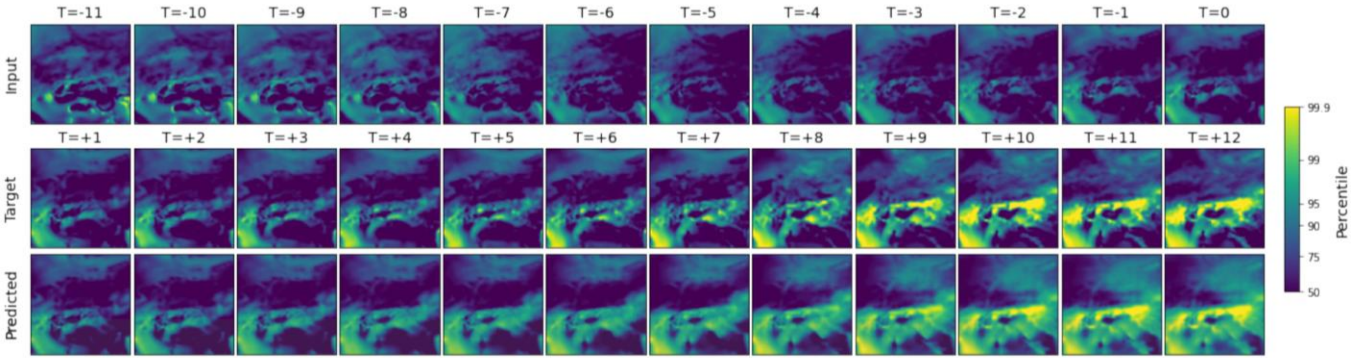
Outlook and next steps

- Integration of wakes in synthetic data generator (should enable the post-processing forecasting models to „learn“)

- Implementation of a spatio-temporal-attention transformer and bayesian neural network

- „Transferable model“: model trained on huge data set of (semi-synthetic) historical production data across Europe, multi-year with attention (location, type, hub height, etc.) which will be re-fitted with every use-case/triggered case, transferable also to unknown sites

- Spatial nowcasting-to-intra-day model (ConvLSTM) for appr. hub height wind speed, longer lead times



Scheepens et al. (2023)