Day-ahead prediction of wind power production with multiple numerical weather prediction data and machine learning algorithms

The International Conference Energy & Meteorology (ICEM) 2023, Padova, June 27-29, 2023

Gang HUANG, Minh-Thang DO

1 BACKGROUND AND MOTIVATION

2 METHODOLOGY

3 CASE STUDY AND INSIGHTS

4 CONCLUSION

Need for more accurate wind power forecasting

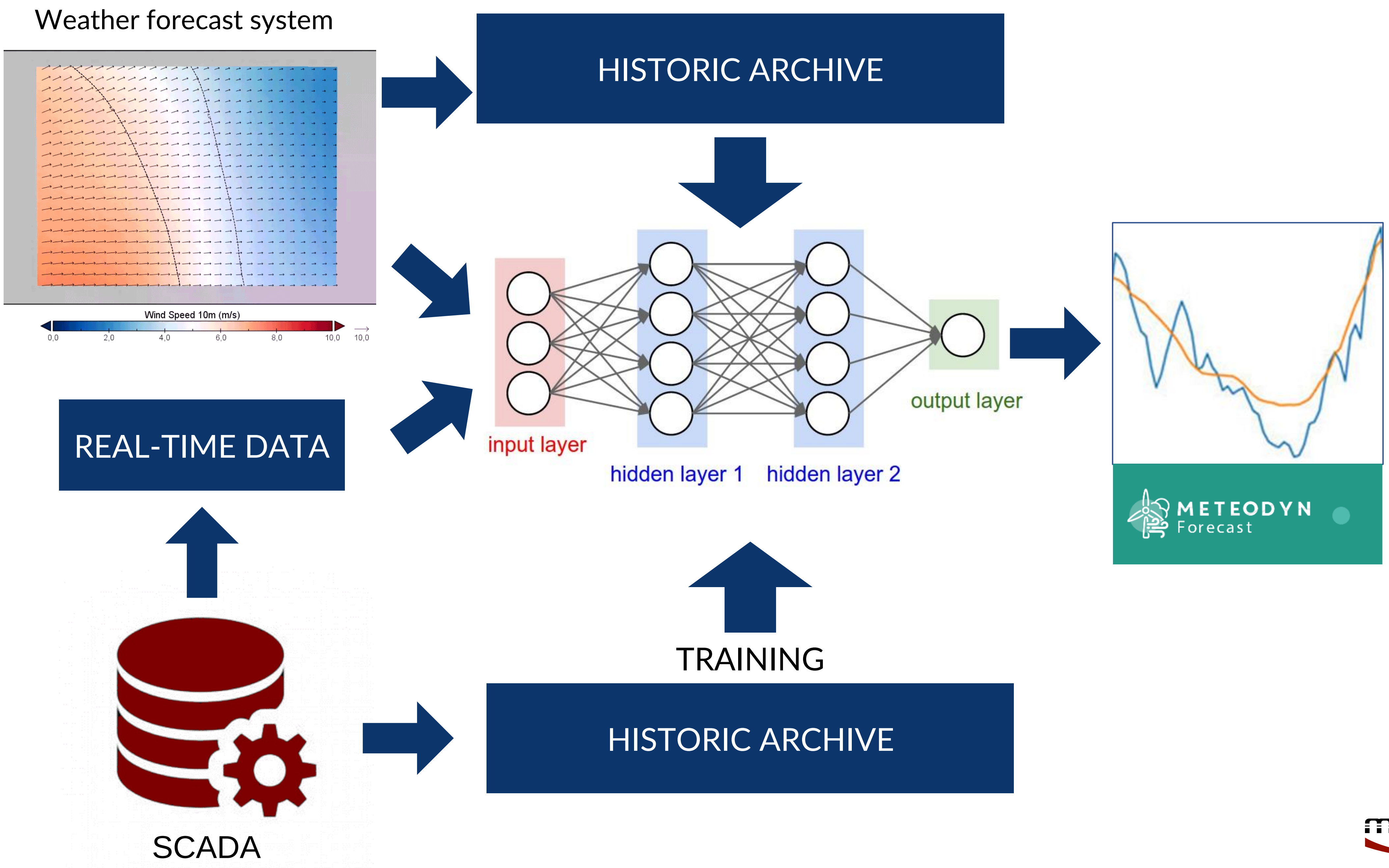
- Growing demand on forecast of wind farm production for several days ahead.
- Need for better forecasts to mitigate uncertainty.
- Model performance may decade with time.

How to leverage **weather forecast data** in wind power forecasting process?

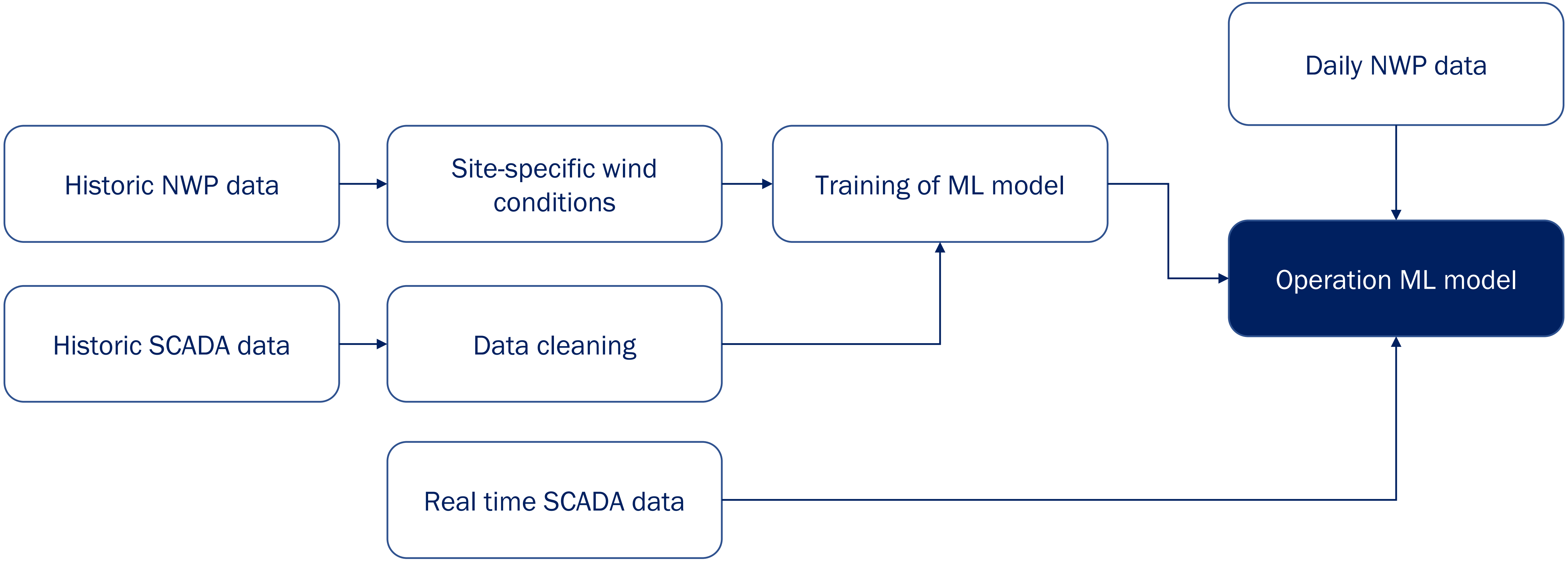
How to build **accurate and reliable machine learning models**?

► Need to a consistent method to build, validate and evaluate machine learning models

Methodology



Workflow

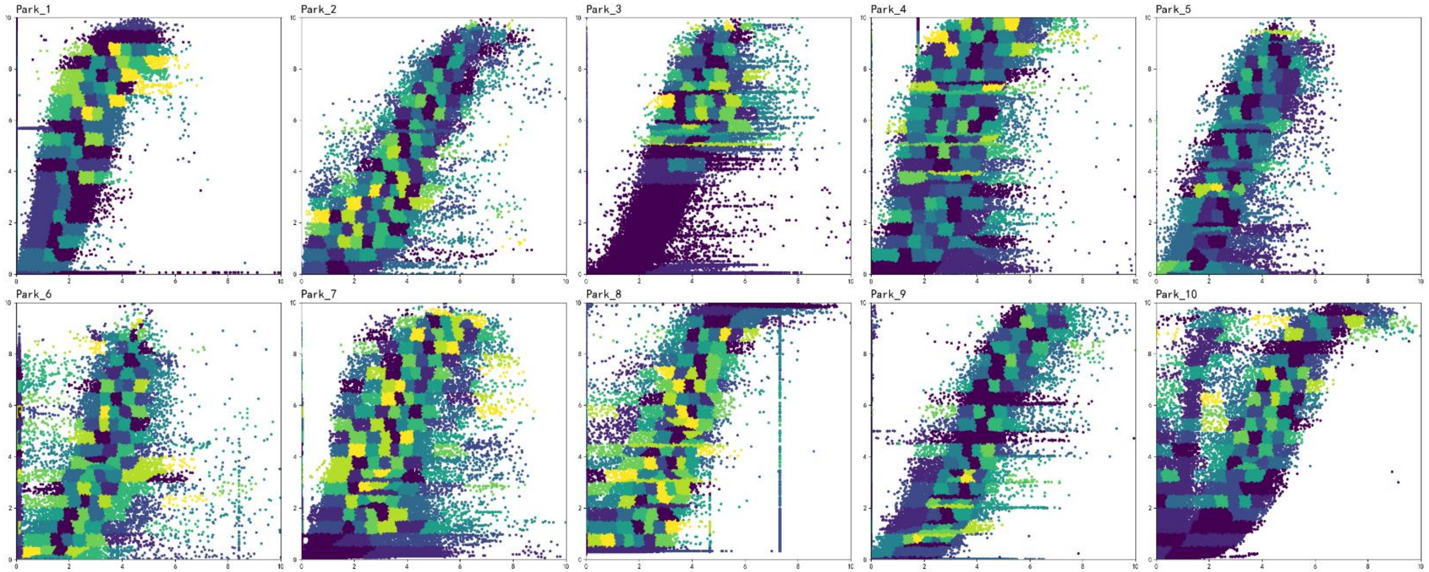


- Site specific wind data need to be extracted from historic NWP data from regional or worldwide models.
- Historic SCADA data from the wind farm need to be cleaned to remove operational anomalies such as curtailment.

Data preparation

"Power curve of wind farm fitting base on Gaussian mixture distribution and S-curve"

Wind Europe 2022 Workshop (Brussels)



Correlation

Feature: wind speed at 100m.

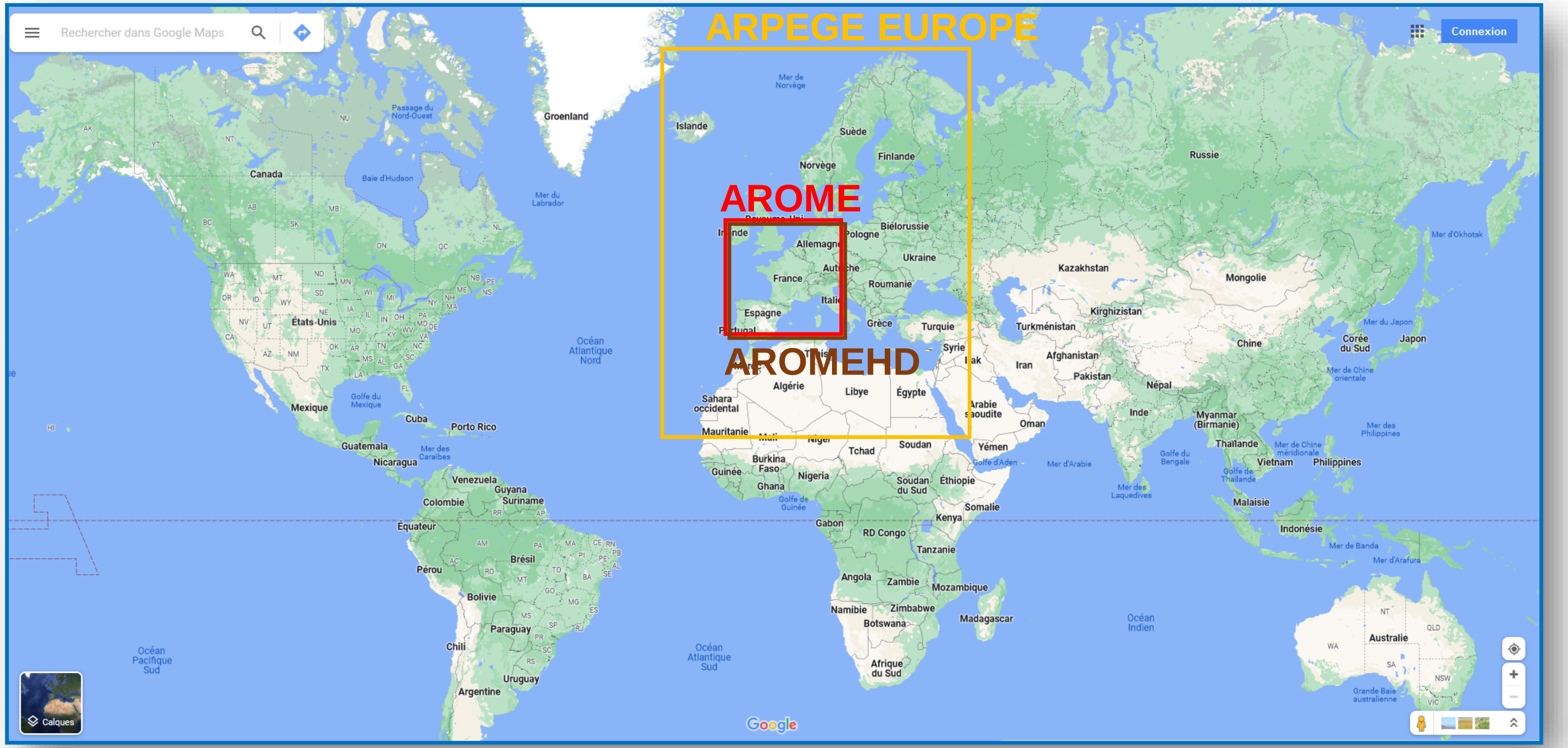
	Correlation		
	cubic	linear	nearest
gfs	0.832	0.835	0.832
ncepgfs	0.823	0.825	0.826
ecmwf	0.837	0.836	0.834
arpege	0.841	0.842	0.841
arpege_eu	0.829	0.831	0.828
arome	0.813	0.814	0.811
aromehd	0.786	0.785	0.786

meteodyn
a CLS Group Company

Multiple weather forecast data sources

GFS, ECMWF, ARPEGE

	GFS_050	ECMWF
Resolution	0.5° x 0.5°	0.1° x 0.1°
Runs per day	4	1 - 2
Time step	3 hour	3 hour
Height	10, 100	10, 100



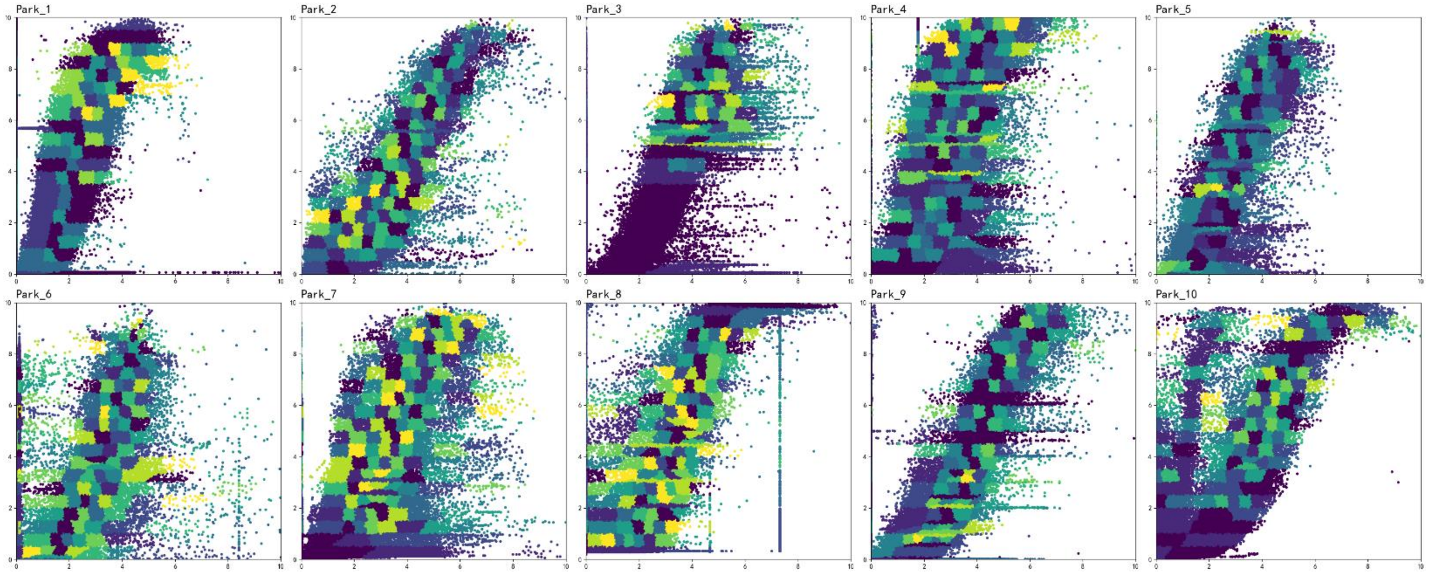
	GFS_025	ARPEGE	ARPEGE_EU	AROME	AROME_HD
Resolution	0.25° x 0.25°	0.25° x 0.25°	0.1° x 0.1°	0.025° x 0.025°	0.01° x 0.01° (1 x 1 km)
Runs per day	4	4	4	4 (8)	4 (8)
Time step	1 hour	1 hour	1 hour	1 hour	1 hour
Height	10, 20, 30, 40, 50, 80, 100	10, 20, 35, 50, 75, 100, 150, 200, 250	10, 20, 35, 50, 75, 100, 150, 200, 250	10, 20, 35, 50, 75, 100, 150, 200, 250	10,20,50,100

a CLS Group Company

Data preparation

“Power curve of wind farm fitting base on Gaussian mixture distribution and S-curve”

Wind Europe 2022 Workshop (Brussels)



Correlation

Feature: wind speed at 100m.

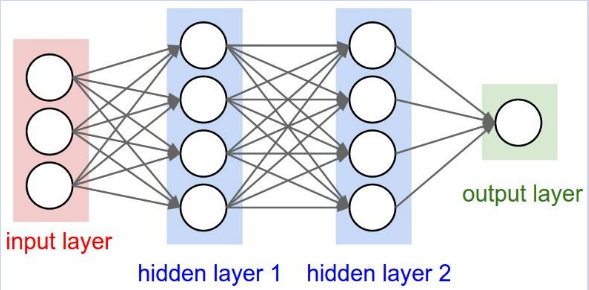
	Correlation		
	cubic	linear	nearest
gfs	0.832	0.835	0.832
ncepgfs	0.823	0.825	0.826
ecmwf	0.837	0.836	0.834
arpege	0.841	0.842	0.841
arpege_eu	0.829	0.831	0.828
arome	0.813	0.814	0.811
aromehd	0.786	0.785	0.786

meteodyn
a CLS Group Company

Machine learning architectures

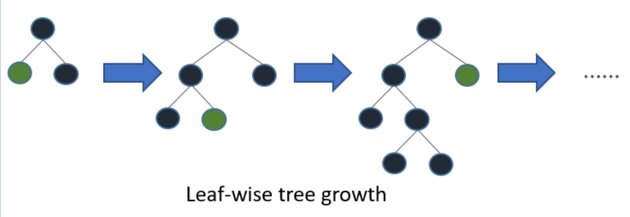
Neural network

- Capable of working on images, videos and language.
- MLP can deduce complex, non-linear relationship between input and output.
- Needs rigorous validation during training to avoid overfitting.



Gradient boosting decision trees

- Work well on tabular data.
- Decision tree: piecewise constant approximation of output by each input.
- Needs rigorous validation during training to avoid overfitting.

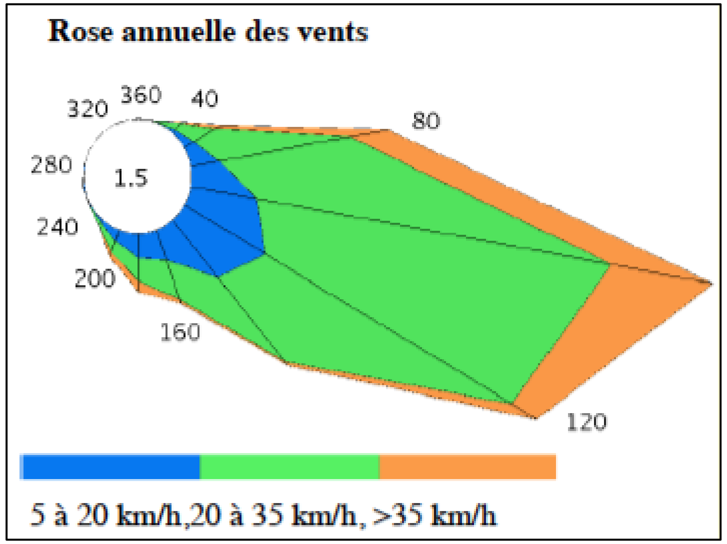


- Extra effort needs to be made in the domain of timeseries forecasting to avoid leakage from future data.
- Gradient boosting tree methods seem to be more light-weight and their performance is often proved in ML competitions.

a CLS Group Company

Site description

- 8 x 900kW turbines
- Hub height: 50 m
- Elevation:
  230 - 260 m (turbine location)



meteodyn
a CLS Group Company

Choice of hyperparameters

- Different machine learning architectures offer different hyperparameters, ie, parameters that characterize a machine learning model but are independent of input data.
- An iterative process is often needed in order to refine the searching range of hyperparameters.

Hyperparameter	Values
Learning rate	0.01
Batch size	32, 128, 256
Layer configuration	[8], [16], [32], [32, 32], [64, 64], [32, 64, 8]
Dropout	0, 0.05, 0.25

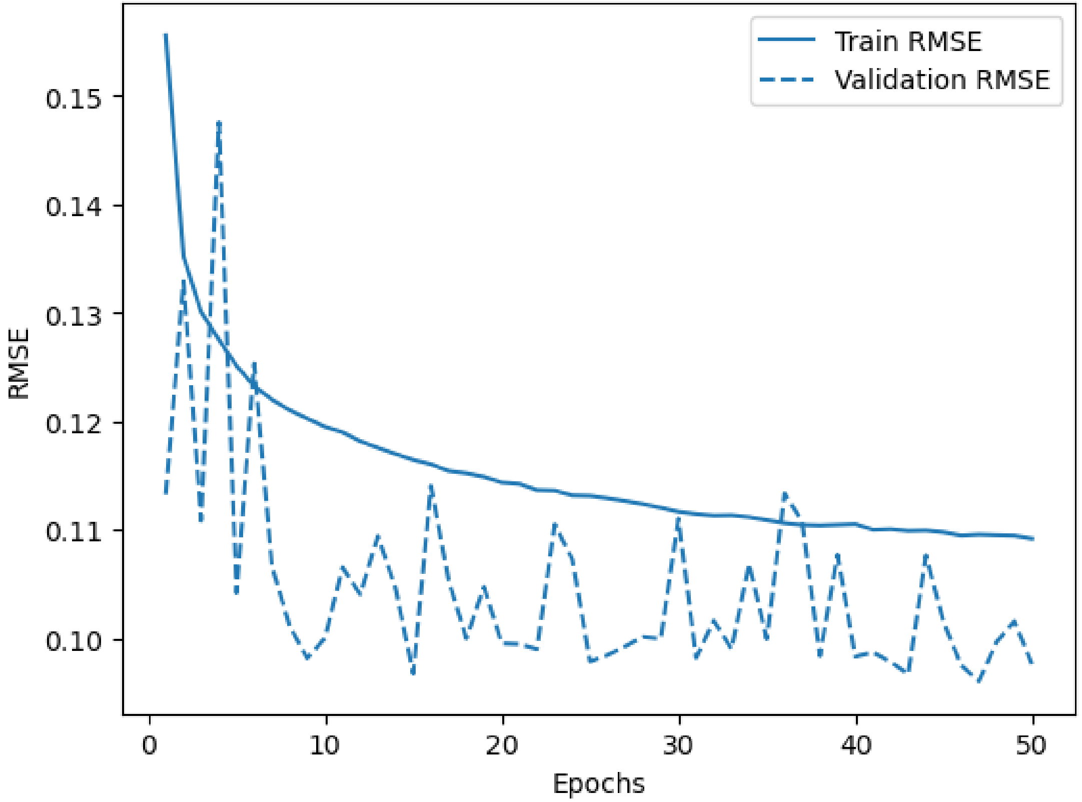
Table 1 Hyperparameters used for NN models. Higher batch size and dropout with simpler layer configurations help with preventing overfitting.

Hyperparameter	Values
Learning rate	0.01
max_depth	2, 3, 4, 5, -1 (no limit)
min_child_samples	200, 500, 1000
colsample_bytree	0.01, 0.2, 0.4, 0.9

Table 2 Hyperparameters used for LightGBM models. Lower "colsample_bytree, "max_depth" and higher "min_child_samples" help with preventing overfitting.

Training of ML models

- Different machine learning architectures offer different hyperparameters, ie, parameters that characterize a machine learning model but are independent of input data.
- An iterative process is often needed in order to refine the searching range of hyperparameters.



Results - MLP (Multilayer perceptron)

- By complexifying the NN, the performance improves up to a limit.
- Simple models have more scatter.

$$\text{Score} = \left(1 - \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{p_i - \hat{p}_i}{p_{\text{rated}}}\right)^2}\right) \times 100\%$$

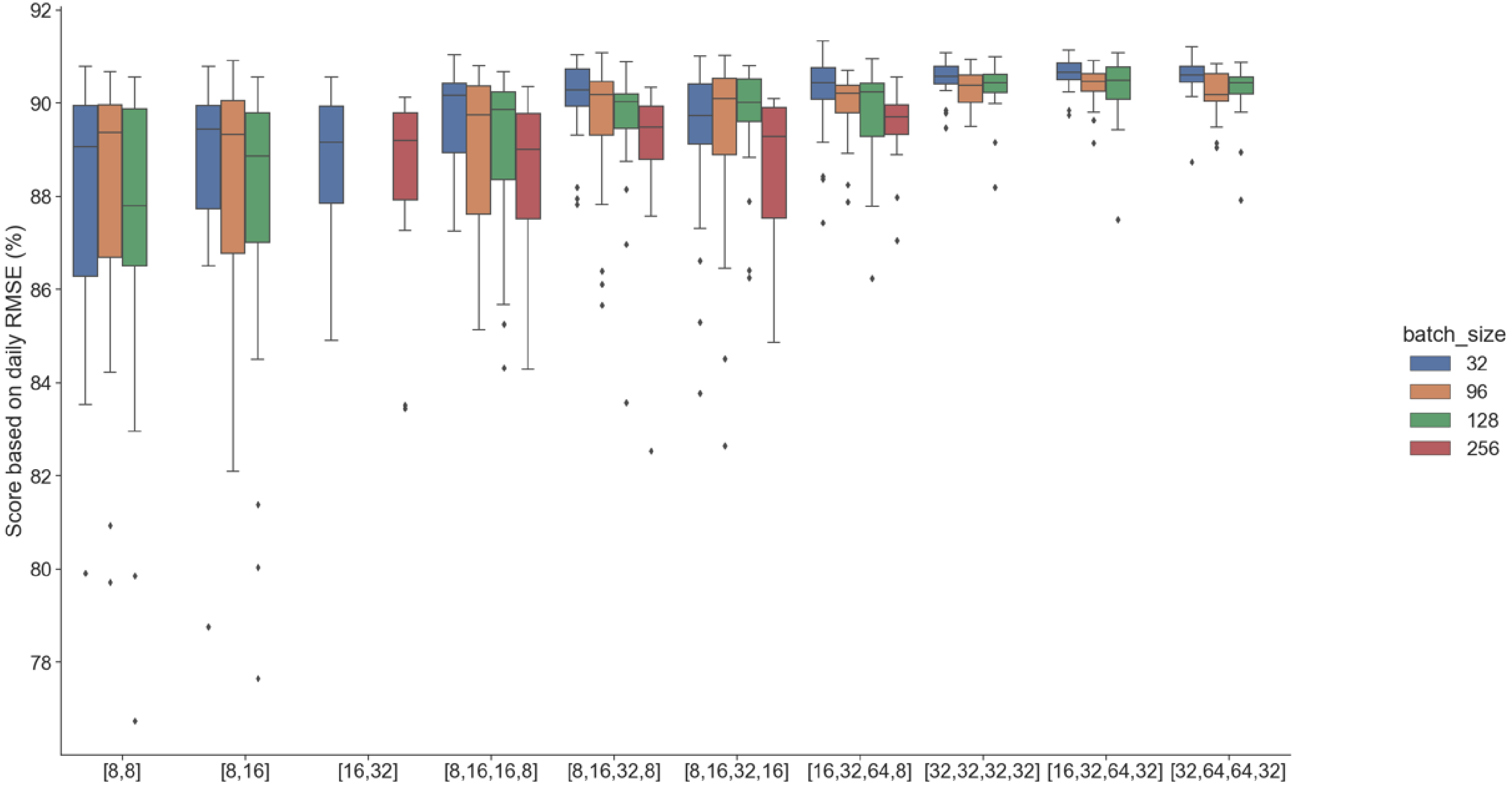


Figure 2 NN forecast score as a function of layer configuration (horizontal axis) and batch size (colors).

Results - LightGBM

- An optimal choice of hyperparameters can be found for an intermediately complex model.
- Less scatter and better performance.

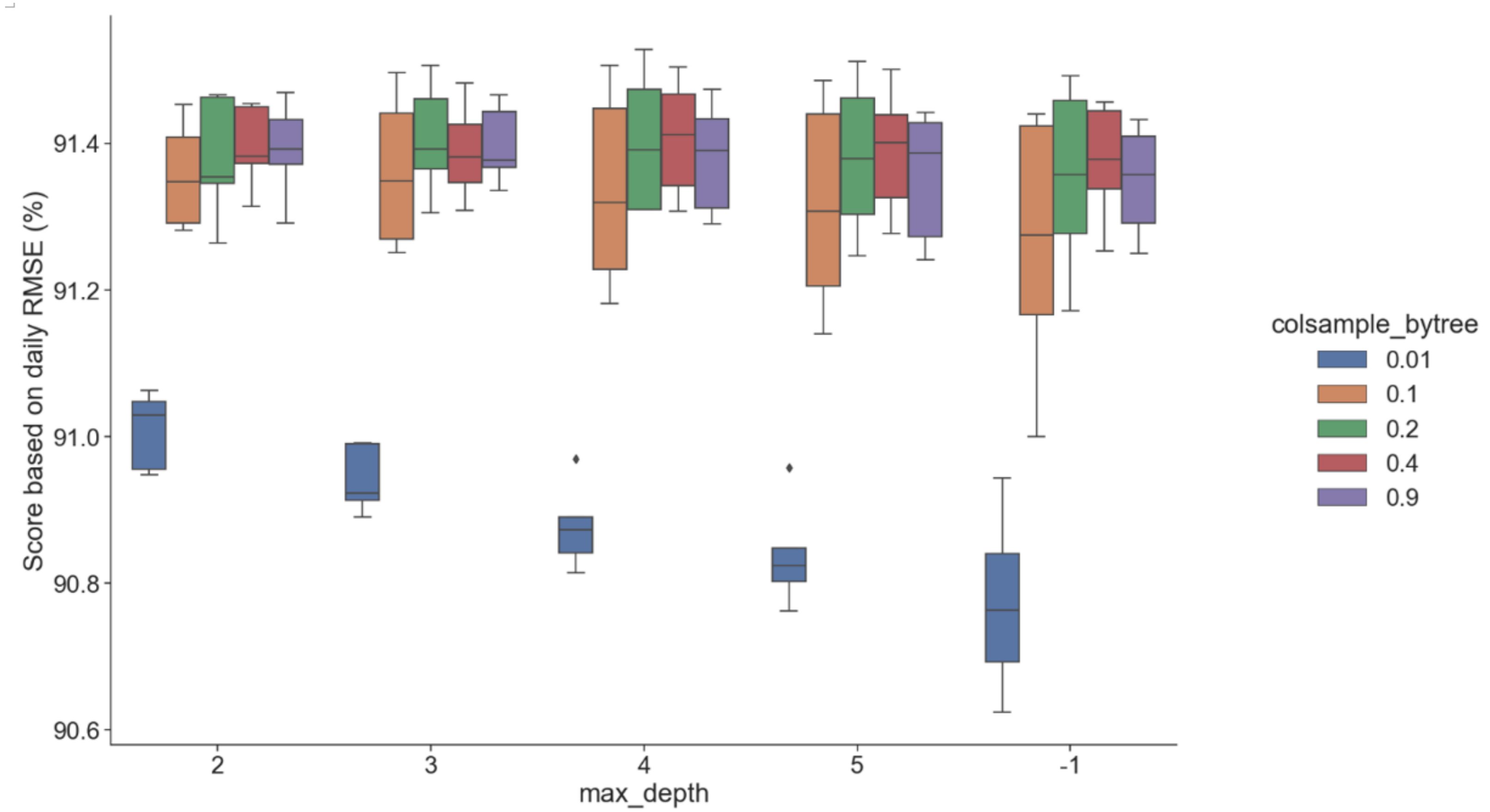


Figure 3 LightGBM forecast score as a function of the hyperparameters max_depth (horizontal axis) and colsample_bytree (colors). For comparison, the best mean score in Figure 2 for NN models is around 90.6%.

a CLS Group Company

Summary of current study

Combined use of multiple NWP data improve wind power forecasting

Gradient-boosted tree models have the advantage of being robust to noise in data

Validations of machine learning models reduce the uncertainty of wind power forecasting

Remaining challenge: How to downscale NWP data to better predict wind conditions at wind farm level

Thank you!

Gang HUANG

Research Engineer, Meteodyn

gang.huang@meteodyn.com

+33 2 40 71 05 05