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 decline with time.

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

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

How to build accurate and reliable machine learning models?
Methodology

Weather forecast system

REAL-TIME DATA

SCADA

HISTORIC ARCHIVE

TRAINING

HISTORIC ARCHIVE

input layer

hidden layer 1

hidden layer 2

output layer
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.

<table>
<thead>
<tr>
<th></th>
<th>cubic</th>
<th>linear</th>
<th>nearest</th>
</tr>
</thead>
<tbody>
<tr>
<td>gfs</td>
<td>0.832</td>
<td>0.835</td>
<td>0.832</td>
</tr>
<tr>
<td>ncep</td>
<td>0.823</td>
<td>0.825</td>
<td>0.826</td>
</tr>
<tr>
<td>ecmwf</td>
<td>0.837</td>
<td>0.836</td>
<td>0.834</td>
</tr>
<tr>
<td>arpege</td>
<td>0.841</td>
<td>0.842</td>
<td>0.841</td>
</tr>
<tr>
<td>arpege_eu</td>
<td>0.829</td>
<td>0.831</td>
<td>0.828</td>
</tr>
<tr>
<td>arome</td>
<td>0.813</td>
<td>0.814</td>
<td>0.811</td>
</tr>
<tr>
<td>arome_hd</td>
<td>0.786</td>
<td>0.785</td>
<td>0.786</td>
</tr>
</tbody>
</table>
Multiple weather forecast data sources

### GFS_025
- **Resolution**: 0.25° x 0.25°
- **Runs per day**: 4
- **Time step**: 1 hour
- **Height**: 10, 20, 30, 40, 50, 80, 100

### ARPEGE
- **Resolution**: 0.25° x 0.25°
- **Runs per day**: 4
- **Time step**: 1 hour
- **Height**: 10, 20, 35, 50, 75, 100, 150, 200, 250

### ARPEGE_EU
- **Resolution**: 0.1° x 0.1°
- **Runs per day**: 4 (8)
- **Time step**: 1 hour
- **Height**: 10, 20, 35, 50, 75, 100, 150, 200, 250

### AROME
- **Resolution**: 0.025° x 0.025°
- **Runs per day**: 4 (8)
- **Time step**: 1 hour
- **Height**: 10, 20, 35, 50, 75, 100, 150, 200, 250

### AROME_HD
- **Resolution**: 0.01° x 0.01° (1 x 1 km)
- **Runs per day**: 4 (8)
- **Time step**: 1 hour
- **Height**: 10, 20, 50, 100
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.

<table>
<thead>
<tr>
<th></th>
<th>cubic</th>
<th>linear</th>
<th>nearest</th>
</tr>
</thead>
<tbody>
<tr>
<td>gfs</td>
<td>0.832</td>
<td>0.835</td>
<td>0.832</td>
</tr>
<tr>
<td>ncepfs</td>
<td>0.823</td>
<td>0.825</td>
<td>0.826</td>
</tr>
<tr>
<td>ecmwf</td>
<td>0.837</td>
<td>0.836</td>
<td>0.834</td>
</tr>
<tr>
<td>arpege</td>
<td>0.841</td>
<td>0.842</td>
<td>0.841</td>
</tr>
<tr>
<td>arpege_eu</td>
<td>0.829</td>
<td>0.831</td>
<td>0.828</td>
</tr>
<tr>
<td>arome</td>
<td>0.813</td>
<td>0.814</td>
<td>0.811</td>
</tr>
<tr>
<td>arome HD</td>
<td>0.786</td>
<td>0.785</td>
<td>0.786</td>
</tr>
</tbody>
</table>
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.
Site description

- 8 x 900kW turbines
- Hub height: 50 m
- Elevation: 230 - 260 m (turbine location)
Different machine learning architectures offer different hyperparameters, i.e., 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.

<table>
<thead>
<tr>
<th>Hyperparameter</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning rate</td>
<td>0.01</td>
</tr>
<tr>
<td>Batch size</td>
<td>32, 128, 256</td>
</tr>
<tr>
<td>Layer configuration</td>
<td>[8], [16], [32], [32, 32], [64, 64], [32, 64, 8]</td>
</tr>
<tr>
<td>Dropout</td>
<td>0, 0.05, 0.25</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Hyperparameter</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning rate</td>
<td>0.01</td>
</tr>
<tr>
<td>max_depth</td>
<td>2, 3, 4, 5, -1 (no limit)</td>
</tr>
<tr>
<td>min_child_samples</td>
<td>200, 500, 1000</td>
</tr>
<tr>
<td>colsample_bytree</td>
<td>0.01, 0.2, 0.4, 0.9</td>
</tr>
</tbody>
</table>

*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, i.e., 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.

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

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%.
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