**DTU Library**

# Using OR + AI to predict the optimal production of offshore wind parks: a preliminary study

**Fischetti, Martina**

*Publication date:*
2017

*Document Version*
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

# Using OR + AI to predict the optimal production of offshore wind parks: a preliminary study

Martina Fischetti and Marco Fraccaro

$P = \frac{1}{2}\rho A v^3 C_p$

$\sqrt{17}$

$e^{i\pi} = -1$

{2.718281828284}
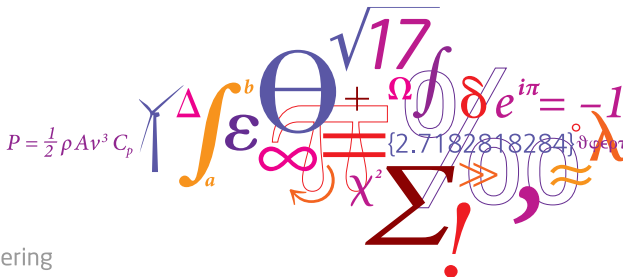
# Research question

**Research question:** Given a bunch of optimized instances for a given problem, can a machine predict the *value* of the optimized solution for a new instance?
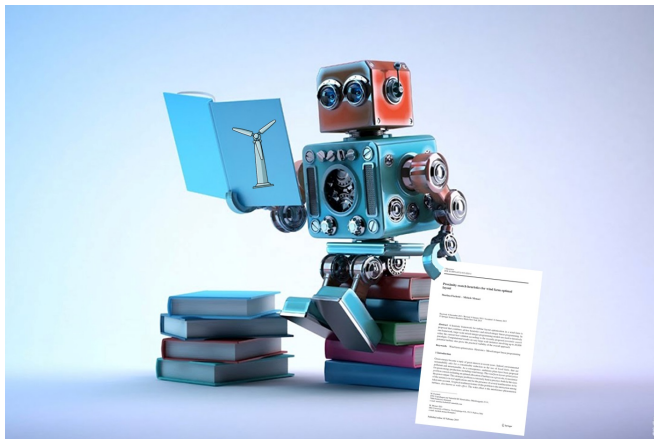
**Research question:** Given a bunch of optimized *wind farm layouts*, can a machine predict the *production value* of the optimized solution for a new *site*?

## Wind Farm Layout Optimization

Determine:
a feasible optimal allocation of turbines that maximizes power production.
Given:

- a site (offshore)

- characteristics of the turbines to build

- measurements of the wind in the site
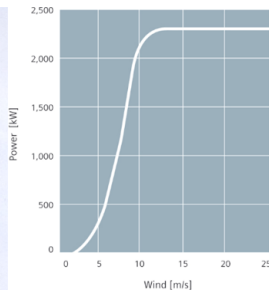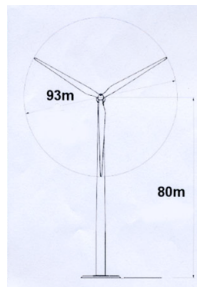
## Wind Farm Layout Optimization

Taking into account:

- proximity constraint

- minimum/maximum number of turbines

- wake effect

## Wind Farm Layout Optimization

Example of solution for a 3000x3000 (m) offshore area. The interference (colors in the background) is the average interference on real-world wind data from Vattenfall

## MILP models
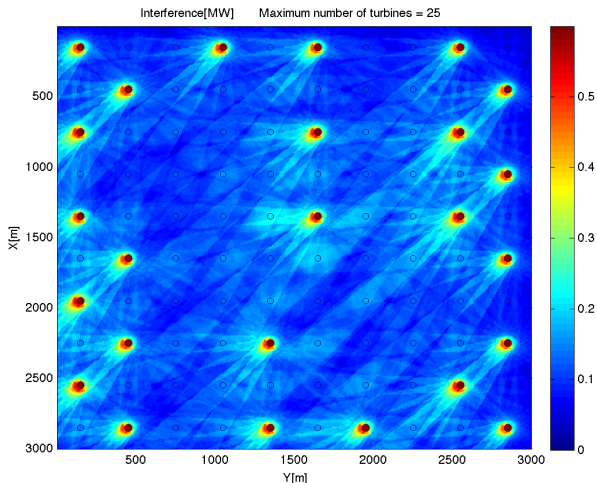
The layout problem can be formulated as a MILP problem.
Variables:

$$x_i = \begin{cases} 1 & \text{if a turbine is built at site } i \in V; \\ 0 & \text{otherwise} \end{cases} \qquad (i \in V)$$

where V is the set of potential turbine positions.
Let $I_{i,j}$ be the interference (production loss) on $j$ because of $i$

## MILP models

The objective function (to be maximized)

$$\sum_{i \in V} P_i x_i - \sum_{i \in V} \left( \sum_{j \in V} I_{ij} x_j \right) x_i \qquad (1)$$

is restated as

$$\sum_{i \in V} (P_i x_i - w_i) \qquad (2)$$

where

$$w_i := \left( \sum_{j \in V} I_{ij} x_j \right) x_i = \begin{cases} \sum_{j \in V} I_{ij} x_j & \text{if } x_i = 1; \\ 0 & \text{if } x_i = 0. \end{cases}$$

denotes the total interference caused by a turbine in position $i$.

## MILP models

$$\max \quad z = \sum_{i \in V}(P_i x_i - w_i)$$

$$\text{s.t.} \quad N_{MIN} \le \sum_{i \in V} x_i \le N_{MAX}$$

$$x_i + x_j \le 1 \qquad \forall \text{incompatible } i, j \in V, i < j$$

$$\sum_{j \in V} I_{ij} x_j \le w_i + M_i(1 - x_i) \quad \forall i \in V$$

$$x_i \in \{0, 1\} \qquad \forall i \in V$$

$$w_i \ge 0 \qquad \forall i \in V$$

where $P_i$ and $I_{ij}$ are average values over a large number of wind scenarios, and $M_i >> 0$

## Proximity Search

To solve instances with 10000+ possible positions

- 1-opt, 2-opt
- Proximity Search on the MILP model

Given an initial (heuristic) solution search for a better solution in the neighbourhood by using the MILP solver as a black box

This algorithm works very well and it is now used inside the company, running the tool overnight

For more info:

- M. Fischetti, M. Monaci, **Proximity search heuristics for wind farm optimal layout**, Journal of Heuristics 22 (4) (2016) 459–474.

## Machine Learning

Let us suppose now that the company experts could build a new park wherever they like (i.e. the site is not given) –> they should evaluate the potential of a large number of possible sites

Due to the large number of sites, a full optimization (8h+ for each site) is infeasible

**Research question:** Given a bunch of optimized *wind farm layouts*, can a machine predict the *production value* of the optimized solution for a new *site*?

## Building training/test set

We artificially created different sites by generating sets of possible points on a regular grid (10m point-to-point distance) inside **rectangles** of different dimensions (all possible combinations of edge sizes 6000, 7000, 8000, 9000, 10000, 11000, 12000, 13000 and 14000m).

## Building training/test set

- Adwen 8 MW, with a rotor diameter of 180m

- Vestas 8.4 MW, with a rotor diameter of 164m

- Siemens 7 MW, with a rotor diameter of 154m

- Vestas 8 MW, with a rotor diameter of 164m

- Siemens 3.2 MW, with a rotor diameter of 113m

- Siemens 2.3 MW, with a rotor diameter of 101m

# Building training/test set

**Real-world wind statistics** from the real offshore wind parks (Borssele 1 and 2, Borssele 3 and 4, Danish Krigers Flak and Ormonde).
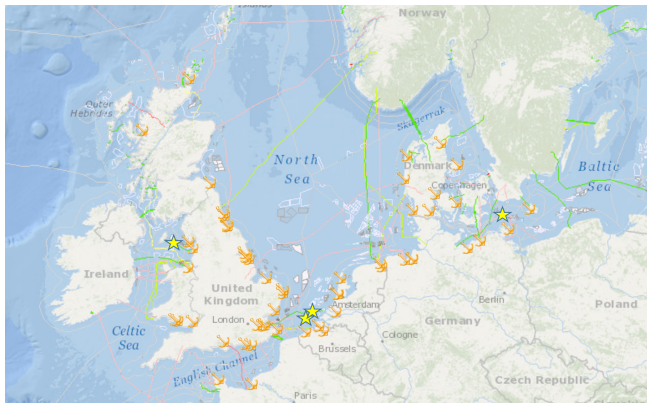


We imposed that a fixed number of 50 turbines needs to be located in the site, minimum distance 5 rotor diameters → we obtained about 2000+ instances

## Building training/test set

**1** the so-called *gross production*, i.e., the power production of the optimized solution neglecting the interference factor

**2** the optimized layout and its power production

## Building training/test set

❶ the so-called *gross production*, i.e., the power production of the optimized
solution neglecting the interference factor
$\rightarrow$ this will be used as input information (feature)

❷ the optimized layout and its power production
$\rightarrow$ **this is what we want to predict**

## Building training/test set

❶ the so-called *gross production*, i.e., the power production of the optimized solution neglecting the interference factor)
$\rightarrow$ this will be used as input information (feature)

❷ the optimized layout and its power production
$\rightarrow$ **this is what we want to predict**

1) requires short computing time and can be calculated in a pre-processing step.
Optimization for the difficult case 2) was obtained through our MILP-based heuristic [1], with a time limit of 1 hour on a standard PC using IBM ILOG CPLEX 12.6.

*[1] M. Fischetti, M. Monaci,Proximity search heuristics for wind farm optimal layout, Journal of Heuristics 22 (4) (2016) 459–474.*

## Building training/test set

Finally, instead of directly estimating the optimized production of a site, we estimate its normalized difference from the gross production, defined as

$$\text{reduction} = \frac{\text{gross production - optimized production}}{\text{gross production}} \qquad (3)$$
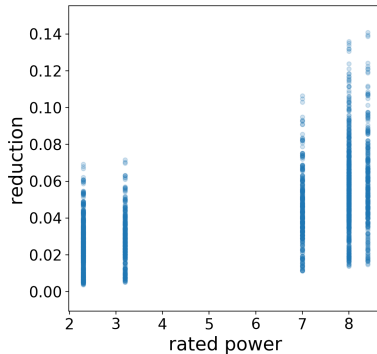
This is a value between 0 and 1 that can easily be compared between instances with production of different scales.

## Feature selection

In order for our ML models to capture the wind park problem, it is very important to describe its characteristics in a meaningful way.

Selected features:



- **rated power
  for the turbine model [MW]**

## Feature selection

In order for our ML models to capture the wind park problem, it is very important to describe its characteristics in a meaningful way.
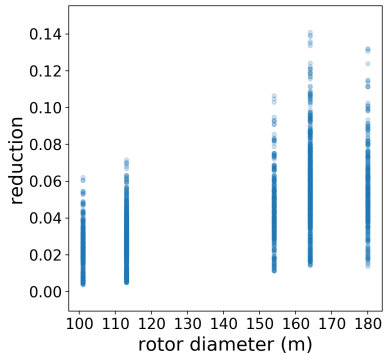
Selected features:



- rated
  power for the turbine model [MW]

- **rotor diameter
  for the turbine model [m]**

## Feature selection

In order for our ML models to capture the wind park problem, it is very important to describe its characteristics in a meaningful way.
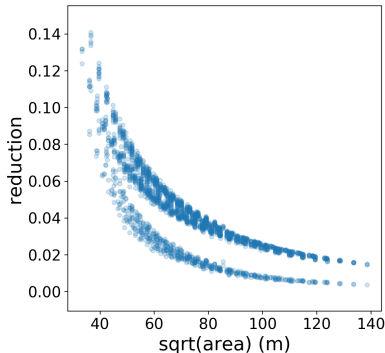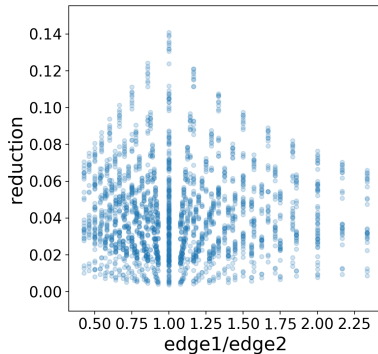
Selected features:



- rated
  power for the turbine model [MW]

- rotor
  diameter for the turbine model [m]

- **the square root of the
  area of the site [rotor diameters]**

## Feature selection

In order for our ML models to capture the wind park problem, it is very important to describe its characteristics in a meaningful way.

Selected features:



- rated
  power for the turbine model [MW]

- rotor
  diameter for the turbine model [m]

- the square root of
  the area of the site [rotor diameters]

- **the ratio between
  the two edges of the rectangle**

# Feature selection

In order for our ML models to capture the wind park problem, it is very important to describe its characteristics in a meaningful way.
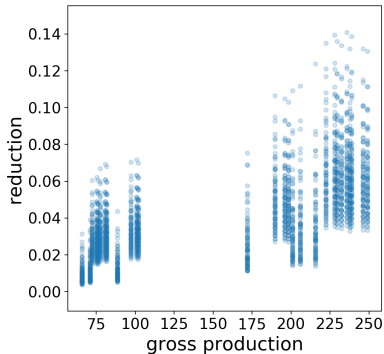
Selected features:



- rated
  power for the turbine model [MW]

- rotor
  diameter for the turbine model [m]

- the square root of
  the area of the site [rotor diameters]

- the ratio between
  the two edges of the rectangle

- **the production
  without interference [MW]**

## Machine Learning

Two different ML models to estimate the reduction in power production due to the interference:

• Linear Regression

• Neural Networks (NNs)

In addition, we also defined a simple baseline model (*Mean Value*), that regardless of its input always predicts the mean reduction of the training set. This last model mimics what is normally done by humans, and is used for comparison.

## Machine Learning

2268 instances

–> **training set** randomly choosing 60% of the instances;

–> **test set** the remaining 40%

$\rightarrow$ hyperparameters of the models are chosen using the *scikit-learn* [1] function *GridSearchCV* (5-fold cross-validated on the training set)

---

[1] a ML library for python

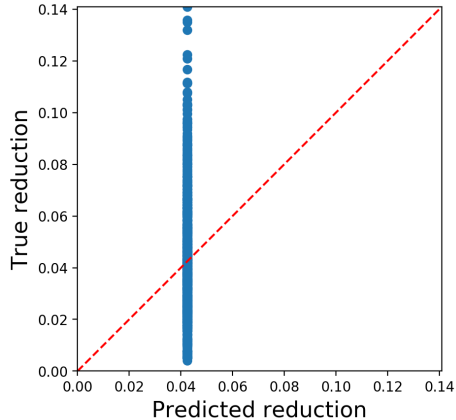## Results

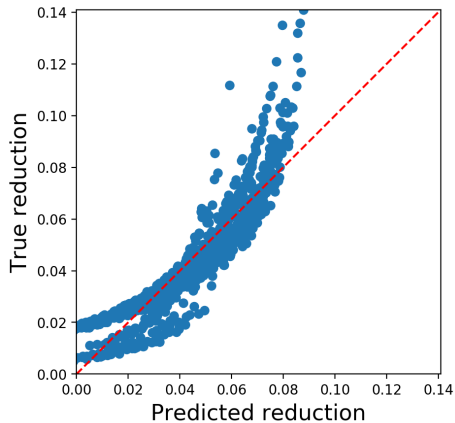What a human would do: reduction mean value



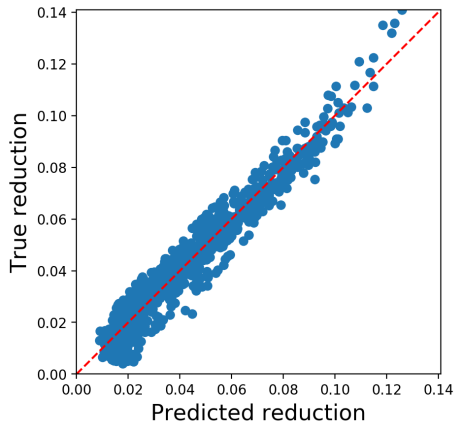Figure: Mean Value

Figure: Linear Regression

Figure: Neural Network

**Conclusions**

This preliminary work shows

- the relevance of using MO and ML techniques together;

- that ML techniques (NNs in particular), trained on a large number of optimized solutions, could well predict the optimal value of new instances of the same (wind park layout) problem

- that the ML estimate highly outperforms the human estimate (Mean Value model)

## Recent work

We have further worked on this project.

- we increased the number of instances including additional sites

## Recent work

We have further worked on this project.

- we increased the number of instances including additional sites

- instead of the gross production we used the production (with wake effect) of a regular layout on a grid (both as feature and as banchmark)

## Recent work

We have further worked on this project.

- we increased the number of instances including additional sites

- instead of the gross production we used the production (with wake effect) of a regular layout on a grid (both as feature and as banchmark)

- we better defined our training/test set (not randomly selected but depending on the wind in the site)

## Recent work

We have further worked on this project.

- we increased the number of instances including additional sites

- instead of the gross production we used the production (with wake effect) of a regular layout on a grid (both as feature and as banchmark)

- we better defined our training/test set (not randomly selected but depending on the wind in the site)

- we designed new ML models (i.e. Support Vector Regression)
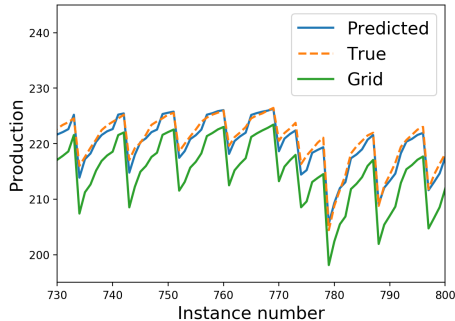
# Future work

Figure: Support Vector Regression

For more info:

- M. Fischetti, M. Fraccaro, **Machine Learning meets Mathematical Optimization to predict the optimal production of offshore wind parks**(submitted).