2023 edition, sponsored by RTE
Configuration of offshore electrical substations to collect offshore power wind

KI025
P. Panciatici, M. Ruiz, C. Vessaire,
L. Baty, M. Brichet and A. Parmentier

1 Context and Needs

The energy transition requires the installation of more offshore wind farms in the near future. These wind farms produce energy that needs to be collected and connected to the main onshore grid. RTE (The French Transmission System Operator) is responsible for designing, maintaining and operating the offshore grid that performs this task.

The offshore grid is relatively simple, but it needs to be optimized for cost and reliability. This hackathon focuses on a simplified version of the problem of designing the offshore grid for a given wind farm. We will take into account investment costs, operational costs, and failure costs.

You have different options for each piece of equipment, with different costs and reliability characteristics. The design involves selecting the following elements: the number, size and type of substations; the number of wind turbines connected to each substation; the size and type of cable between the offshore substation and the onshore connection substation; the location, size and type of cables between the offshore substations.

2 Problem statement

Figure 1 illustrate an off-shore wind farm. Electricity produced by wind turbines is sent first to an offshore substation, then to the unique on-shore station. There are several offshore substations. Each wind turbine is connected to a single substation. Each substation is connected to the onshore station. We make the assumption that wind turbines and the onshore station never fail. On the contrary, the off-shore substations sometimes fail. Some cables are therefore built between a substation and some of the neighbor substation, and enable to evacuate power through these neighbors when the initial substation fails. The purpose of this Hackathon is to design the wind farm grid. The location of the single onshore station and of the wind turbines are known in advance. On the contrary, we must choose which substations to build on a set of given possible locations. We also have to choose which cables we install, and the rating[1] of the cables installed. Cables are directed: Power can be transmitted only in one direction along each cable. The aim is to minimize the sum of the construction costs and of the operating costs.

2.1 Grid

The power graph. Let v_0 be the onshore station, V^{s} be the set of possible locations for substations, and V^{t} be the set of wind turbines. Let $V = \{v_0\} \cup V^{\mathrm{s}} \cup V^{\mathrm{t}}$ be the set of vertices. Let E be the set of edges, i.e.,

[1] rating: quantity of power that can be produced or transported

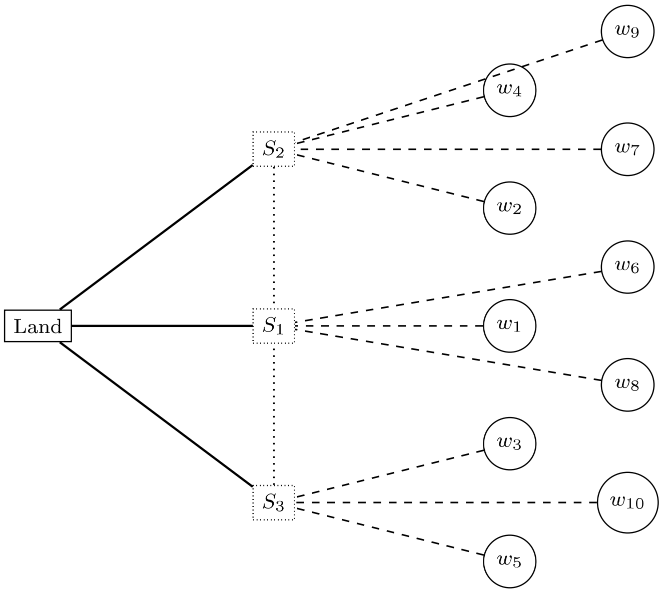

Figure 1: Illustration of a simple offshore wind farm and its connection to land. w_i are wind turbines. S_j are the possible offshore substations. Land is the onshore connection point; Dashed lines are cables from one wind turbine to a offshore substation; Dotted lines are cables between offshore substations. Plain lines are cables between offshore substation and the onshore connection point.

the set of unordered pair of vertices (u, v) in V such that we can build a cable between u and v. It contains:

- an edge (v_0, u) for each substation u in V^{s}; Denote by E^0 this set of edges.
- an edge (u, v) for each wind turbine $u \in V^{\mathrm{t}}$ and substation v in V^{s}; Denote by E^{t} this set of edges.
- an edge (u, v) between each pair of substations (u, v) in V^{s}; Denote by E^{s} this set of edges.

We have $E = E^0 \cup E^{\mathrm{s}} \cup E^{\mathrm{t}}$. We denote by ℓ_e the length (distance) of edge e.

Building substations. On each location in V^{s}, we can build a substation. Let us denote by S the set of different types of substations that can be built. We denote by x_{vs} the binary variable equal to 1 if we build a substation of type s in v and to 0 otherwise. The following constraint ensures that at most one substation is built in each location.

$$\sum_{s \in S} x_{vs} \leq 1, \quad \text{for all } v \in V^{\mathrm{s}} \tag{1}$$

Each type s of substation has a *rating* r_s, an expected *proportion of time spent in failed state* p_s, and a *construction cost* c_s.

Cables between land and substations. If a substation is built, then a single cable must be built between the substation and the onshore station. Let Q^0 be the set of types of cables available for this. Each of these cables have a *rating* r_q and an expected *proportion of time spent in failed state* p_q. We denote by c_{eq} the *cost* of installing a cable of type q along edge e. It is the sum of a fixed cost and a cost proportional to its length ℓ_e: $c_{eq} = c_q^{\mathrm{f}} + c_q^{\ell} \ell_e$. Let y_{eq} the binary variable equal to 1 if such a cable is installed and to 0 otherwise. We have the following constraint

$$\sum_{q \in Q^0} y_{eq} = \sum_{s \in S} x_{vs}, \quad \text{for all } v \in V^{\mathrm{s}}, \text{ and } e = (v_0, v) \tag{2}$$

Cables between substations and turbines. We make the hypothesis that the wind turbines all have the same *rating* $P^{\max}$. The cables between wind turbines and substations are identical. The *cost* c_e of installing a cable along edge e depends on its length ℓ_e: $c_e = c_{\mathrm{t}}^{\mathrm{f}} + c_{\mathrm{t}}^{\ell}\ell_e$. We denote by z_e the binary variable equal to 1 if a cable is built along e and 0 otherwise. A single cable is built between each wind turbine and substations:

$$\sum_{e=(v,t)\in E^t} z_e = 1, \quad \text{for all } t \in V^{\mathrm{t}}. \tag{3}$$

We assume that cables between wind turbines and substations never fail, and that their rating is greater than $P^{\max}$.

Cables between substations. Let us finally denote by Q^{s} the set of types of cables that we can install between stations. For each such cable q, we denote by $c_{eq} = c_q^{\mathrm{f}} + c_q^{\ell}\ell_e$ its installation *cost* along edge e, and by r_q its *rating*. Let y_{eq} be a binary variable equal to 1 if a cable of type q is built along e, and 0 otherwise. Each built substation can be connected to at most 1 other substation. We have

$$\sum_{e=(v,v')\in E^s} \sum_{q\in Q^{\mathrm{s}}} y_{eq} \le \sum_{s\in S} x_{vs}, \quad \text{for all } v \text{ in } V^{\mathrm{s}} \tag{4}$$

We assume that cables between substations never fail.

2.2 Scenarios

We use scenarios to model weather conditions, which have an impact on the power generated by turbines, and failures. Let Ω be a finite set of scenarios. We assume that the power delivered by wind turbines is scenario-dependent, but identical for all turbines under a given scenario. Let π_ω ($\le P^{\max}$) be the power delivered by each turbine under scenario ω.

The operational cost come from the fact the network operator must pay a penalty proportional to the power it cannot bring to the shore.

No failure cost. When there is no failure, the network operates in radial mode, which means that cables between substations are not used. Let us denote by $C^{\mathrm{n}}(x, y, z, \omega)$ the curtailing under scenario ω with no failure, and $C^{\max}$ the maximum curtailing penalty that can be incurred by the network manager. It is therefore equal to

$$C^{\mathrm{n}}(x, y, z, \omega) = \sum_{e=(v_0,v)\in E^0} \overbrace{\left[\underbrace{\pi_\omega\Big(\sum_{\bar{e}=(v,t)\in E^{\mathrm{t}}} z_{\bar{e}}\Big)}_{\substack{\text{Power generated}\\\text{by turbines linked}\\\text{to } v}} - \underbrace{\min\Big(\sum_{s\in S} r_s x_{vs}, \sum_{q\in Q^0} r_q y_{eq}\Big)}_{\substack{\text{Capacity of the}\\\text{cable / substation}\\\text{linking } v}}\right]^+}^{\text{Curtailing of } v \text{ under scenario } \omega} \tag{5}$$

where $[\cdot]^+$ denotes the positive part ($[x]^+ = \max(x, 0)$).

Cost under failure. We make the simplifying hypothesis that no two equipment can fail simultaneously. We denote by $C^{\mathrm{f}}(v, x, y, z, \omega)$ the cost when the cable linking substation v to the ground or substation v

itself fails.

$$
\begin{aligned}
C^{\mathrm{f}}(v,x,y,z,\omega) = &\overbrace{\Bigg[\pi_\omega\Big(\sum_{\tilde{e}=(v,t)\in E^{\mathrm{t}}} z_{\tilde{e}}\Big) - \sum_{\tilde{e}=(v,\bar{v})\in E^{\mathrm{s}}}\sum_{q\in Q^{\mathrm{s}}} r_q y_{\tilde{e}q}\Bigg]^+}^{\text{Curtailing of } v \text{ under } \omega \text{ and failure of } v} + \\
&\sum_{\substack{e=(v_0,\bar{v})\in E^0\\ \bar{v}\neq v\\ \tilde{e}=(v,\bar{v})}} \overbrace{\Bigg[\underbrace{\pi_\omega\Big(\sum_{\tilde{e}=(\bar{v},t)\in E^{\mathrm{t}}} z_{\tilde{e}}\Big)}_{\text{Power generated by turbines linked to } \bar{v}} + \underbrace{\min\Big(\sum_{q\in Q^{\mathrm{s}}} r_q y_{\tilde{e}q}, \pi_\omega\Big(\sum_{\tilde{e}=(v,t)\in E^t} z_{\tilde{e}}\Big)\Big)}_{\text{Power sent from } v \text{ to } \bar{v}} - \underbrace{\min\Big(\sum_{s\in S} r_s x_{\bar{v}s}, \sum_{q\in Q^0} r_q y_{eq}\Big)}_{\text{Capacity of the cable / substation linking } \bar{v}}\Bigg]^+}^{\text{Curtailing of } \bar{v} \text{ under scenario } \omega \text{ and failure of } v}
\end{aligned}
\tag{6}
$$

The cost $c^{\mathrm{c}}(C)$ associated to a curtailing C is linear in the curtailing, with a strong additional penalty if the curtailing is beyond a threshold $C^{\max}$

$$
c^{\mathrm{c}}(C) = c^0 C + c^{\mathrm{p}}[C - C^{\max}]^+
$$

Finally, let us denote by $p^{\mathrm{f}}(v,x,y)$ the probability of failure of the station in v or the cable linking it to v_0.

$$
p^{\mathrm{f}}(v,x,y) = \sum_{s\in S} p_s x_{vs} + \sum_{q\in Q^0} p_q y_{eq} \quad \text{where} \quad e = (v_0, v)
$$

Let us denote by c^{c} the unit cost of curtailing. We get the following total cost.

$$
\begin{aligned}
c(x,y,z) = &\overbrace{\sum_{v\in V^{\mathrm{s}}}\sum_{s\in S} c_s x_{vs} + \sum_{e\in E^0}\sum_{q\in Q^0} c_{eq} y_{eq} + \sum_{e\in E^{\mathrm{s}}}\sum_{q\in Q^{\mathrm{s}}} c_{eq} y_{eq} + \sum_{e\in E^{\mathrm{t}}} c_e z_e}^{\text{construction cost}} \\
&+ \underbrace{\sum_{\omega\in\Omega} p_\omega \Bigg[\sum_{v\in V^{\mathrm{s}}} p^{\mathrm{f}}(v,x,y) c^{\mathrm{c}}\Big(C^{\mathrm{f}}(v,x,y,z,\omega)\Big) + \Big(1 - \sum_{v\in V^{\mathrm{s}}} p^{\mathrm{f}}(v,x,y)\Big) c^{\mathrm{c}}\Big(C^{\mathrm{n}}(x,y,z,\omega)\Big)\Bigg]}_{\text{operational cost}}
\end{aligned}
\tag{7}
$$

2.2.1 Problem statement

The objective of this hackathon is to find the grid that minimizes the total cost. It can be stated as the following problem.

$$
\begin{aligned}
\min \;& c(x,y,z) \\
\text{s.t. } \;& \text{constraints (1), (2), (3), and (4),} \\
& x \in \{0,1\}^{V^{\mathrm{s}}\times S}, \quad y \in \{0,1\}^{(E^0\times Q^0)\cup(E^{\mathrm{s}}\times Q^s)}, \quad z \in \{0,1\}^{E^{\mathrm{t}}}
\end{aligned}
\tag{8}
$$

3 Instance format and solutions

Instances. Instances are given under the `json` format, which basically contains embedded dictionaries. In these dictionaries, the keys are always strings within quotation marks. Here is an example of a `tiny.json`.

Listing 1: `tiny.json`

```
{
    "general_parameters": {
        "curtailing_cost": 0.4,
        "curtailing_penalty": 0.8,
        "fixed_cost_cable": 1000,
        "main_land_station": {
            "x": 0,
            "y": 0
        },
        "maximum_curtailing": 50000,
        "maximum_power": 1000,
        "variable_cost_cable": 10
    },
    "land_substation_cable_types": [
        {
            "fixed_cost": 2000,
            "id": 1,
            "probability_of_failure": 0.02,
            "rating": 500,
            "variable_cost": 20
        },
        {
            "fixed_cost": 3000,
            "id": 2,
            "probability_of_failure": 0.01,
            "rating": 1000,
            "variable_cost": 30
        }
    ],
    "substation_locations": [
        {
            "id": 1,
            "x": -10,
            "y": -10
        },
        {
            "id": 2,
            "x": -20,
            "y": -20
        }
    ],
    "substation_substation_cable_types": [
        {
            "fixed_cost": 1000,
            "id": 1,
            "probability_of_failure": 0.02,
            "rating": 500,
            "variable_cost": 10
        },
        {
```

Listing 2: `tiny.json` (continuation)

```
            "fixed_cost": 1500,
            "id": 2,
            "probability_of_failure": 0.01,
            "rating": 1000,
            "variable_cost": 15
        }
    ],
    "substation_types": [
        {
            "cost": 100000,
            "id": 1,
            "probability_of_failure": 0.01,
            "rating": 500
        },
        {
            "cost": 200000,
            "id": 2,
            "probability_of_failure": 0.005,
            "rating": 1000
        }
    ],
    "wind_scenarios": [
        {
            "id": 1,
            "power_generation": 100,
            "probability": 0.1
        },
        {
            "id": 2,
            "power_generation": 150,
            "probability": 0.9
        }
    ],
    "wind_turbines": [
        {
            "id": 1,
            "x": -5,
            "y": -5
        },
        {
            "id": 2,
            "x": -15,
            "y": -15
        },
        {
            "id": 3,
            "x": -25,
            "y": -25
        }
    ]
}
```

Let us now briefly describe its syntax. The json file containing the instance has the following attributes:

- `general_parameters`: This contains the general parameters of the problem, such as:

- `maximum_power`: maximum power that can be generated by the wind farm, denoted by $P^{\max}$. Note that this value is informative only, it is not needed for computing the cost or check the constraints.
- `curtailing_cost` and `curtailing_penalty`: curtailing cost c^0 and curtailing penalty c^{p}
- `maximum_curtailing`: curtailing threshold $C^{\max}$.
- `fixed_cost_cable`: fixed cost $c^{\mathrm{f}}_{\mathrm{t}}$ of installing a turbine-substation cable.
- `variable_cost_cable`: variable cost c^{ℓ}_{t} of installing a turbine-substation cable per kilometer.
- `main_land_station`: This contains the coordinates of the onshore station, denoted by v_0.

- `substation_types`: This is a list of substation types that can be built, each with its id, cost, rating, and probability of failure. These correspond to the set S and the parameters c_s, r_s, and p_s.
- `land_substation_cable_types`: This is a list of cable types that can be used to connect a substation to the onshore station, each with its id, fixed cost, variable cost, probability of failure, and rating. These correspond to the set Q^0 and the parameters c^{f}_q, c^{ℓ}_q, p_q, and r_q.
- `substation_substation_cable_types`: This is a list of cable types that can be used to connect two substations, each with its id, fixed cost, variable cost, probability of failure, and rating. These correspond to the set Q^{s} and the parameters c^{f}_q, c^{ℓ}_q, and r_q.
- `substation_locations`: This is a list of possible locations for substations, each with its id and coordinates. These correspond to the set V^{s}.
- `wind_turbines`: This is a list of wind turbines, each with its id and coordinates. These correspond to the set V^{t}.

Solutions. The json file containing the solution has the following attributes:

Listing 3: `tiny_sol.json`

```
{
    "substations": [
        {
            "id": 1,
            "land_cable_type": 1,
            "substation_type": 1
        },
        {
            "id": 2,
            "land_cable_type": 2,
            "substation_type": 2
        }
    ],
    "substation_substation_cables": [
        {
            "substation_id": 1,
            "other_substation_id": 2,
            "cable_type": 2
        }
    ],
```

Listing 4: `tiny_sol.json` (continuation)

```
    "turbines": [
        {
            "id": 1,
            "substation_id": 1
        },
        {
            "id": 2,
            "substation_id": 1
        },
        {
            "id": 3,
            "substation_id": 2
        }
    ]
}
```

- `substations`: This is a list of substations selected, each with its id, substation type, and land cable type.These correspond to the variables x_{vs} and y_{eq} ($e \in E^0, q \in Q^0$).

- `substation_substation_cables`: This is a list of edges between substations, with the two substation ids, and the cable type. These correspond to the variables y_{eq} $(e \in E^s, q \in Q^s)$.
- `turbines`: This list contains for each wind turbine its id and the id of the substation to which it is connected. These correspond to the variables z_e.

4 Some tips

You have to be very efficient to address such a problem in six hours. The team that is going to win is the one that manages to quickly produce decent solutions.

1. Divide the tasks
2. Immediately start coding the tools to parse an input file and create an output file
3. Keep it simple: it's not hard to build a feasible solution. Start by coding simple methods that give you pretty good solutions and avoid designing a very powerful algorithm that you will not be able to implement in 6 hours.