Master thesis : Wind correlation analysis between Greenland and Europe

Auteur : Jouretz, Raphaël
Promoteur(s) : Ernst, Damien
Faculté : Faculté des Sciences appliquées
Diplôme : Master en sciences informatiques, à finalité spécialisée en "computer systems and networks"
Année académique : 2017-2018
URI/URL : http://hdl.handle.net/2268.2/4562

Master Thesis

Wind correlation analysis between Greenland and Europe

Graduation Studies conducted for obtaining the Master's degree
in Computer sciences by Raphaël Jouretz

supervised by Damien Ernst

University of Liège - Faculty of Applied Sciences
Academic year 2017 - 2018

Acknowledgement

I would like to thank my departmental supervisor Prof. Damien Ernst. His constant guidance and advice played the important role in the execution of this document. He always brought me help and suggestions in order to make this document as rigorous as possible.

I would also like to express my deepest gratitude to his right-hand man Raphaël Fonteneau for his invaluable support and guidance throughout the research. His continued support led me to the right way. He always gave me his suggestions that were crucial in making this document as flawless as possible.

Finally, I would like to extend my appreciation to other members of the department for their precious advice during my research and their constant availability for any request.

Contents

1 Abstract

This article is dedicated to the development of computational methods for processing time series from wind speed signals. These computational methods are designed under the light of electricity generation from this intermittent renewable source. We propose to define the notion of critical time windows for measuring the proportion of the time during which renewable energy fields may simultaneously not produce sufficiently, i.e., with a load factor lower than a predefined threshold. Using this type of criterion opens the door to expressing optimization problems describing how to choose the localizations where to install production capacities. The proposed methodology is benchmarked using wind time series data from Europe and Greenland. In this document, two sets of time series are considered: the first one was generated using the MERRA-2 from NASA, while the second one was generated using the MAR model, a regional model taking into account Katabatic winds. First experimental results illustrate the fact that connecting Western Europe with Greenland may effectively decrease the proportion of critical time windows.

2 Introduction

Among the strategies that are currently investigated by the scientific community to carbonized the electricity mix, one can mention the so-called *global grid* approach [1, 2]. Such an approach amounts in realizing large electric interconnections in order to take advantage of natural fluctuations: for example, during the night, at a given location, it would be possible to import electricity produced through solar energy from a different location where, at that same time, the sun was shining, thus avoiding the need to build huge storage capacities. The same reasoning applies for managing seasonal fluctuations: by interconnecting both hemispheres of the world, places going through the winter may benefit from solar energy production in other places going through summer at the same time. Also, as discussed in the original papers introducing the global grid approach [1, 2], large interconnections may favour the harvesting of energy from extreme renewable energy sources.

In this document, we specifically address the problem of assessing, two which extend, a given geographical area may face time periods during which electricity generation from wind is *not sufficient*, the notion of being not enough being expressed in the form of a load factor threshold. Since one of the objective of realizing large interconnections is to increase the proportion of the time when electricity is generated from wind turbines, we specifically design computational methodologies for computing such time proportion. Our methodologies are grounded on the notion of critical time windows, which are series of Δ subsequent time steps (of one hour), where $\Delta \in \{1, \ldots, 8760\}$, during which the load factor of all considered localizations are simultaneously below a threshold $\alpha \in [0, 1]$. The notion of $(\Delta, \alpha)-$criticality opens the door to the description of several optimization problems related with the optimal selection of localizations for deploying generation capacities.

We first propose a detailed, generic formalization for wind signals followed by a few optimization problems. In particular, we investigate how the proportion of $(\Delta, \alpha)-$critical windows evolves as a function the set of geographical localizations, Δ and α. We also formalize optimization problems related with the selection of localizations to deploy generations capacities so that the resulting proportion of $(\Delta, \alpha)-$critical windows is optimized. Our methodology is challenged using data generated by two climate models, the MERRA-2 model [5] and the MAR model [3, 8] in the context of Western-Europe and Greenland interconnection. One of the advantages of Greenland is the fact that it may show wind patterns that are anti-correlated with Western Europe winds. In addition to this, the South-East part of Greenland may show Katabatic winds [3]. Katabatic wind is the technical name for a wind that carries high-density air from a higher elevation down a slope under the force of gravity. This document also investigates to which extent this two phenomena may be of interest for power generation.

The following of this document is structured as follows. We first develop in Section 3 a formalization of the notion of critical windows. Section 4 is dedicated to applying our methodology using wind data from France, Europe and Greenland. The section 6 will conclude the results of this document. Finally, the section 5 brings the formalization further by introducing new concepts.

3 formalization

This section will formalize the data-processing methodology by using set-theoretical framework.

3.1 Set of locations

Let $\mathcal{L}$ denote the set of all locations considered in this paper. Each location $l \in \mathcal{L}$ is defined by a pair $l = (\lambda^{long}, \lambda^{lat}) \in \mathbb{R}^2$, corresponding to the geographic coordinate of l. In the following, we also use the notation L to denote any subset of $L \subset \mathcal{L}$:

$$L \in \mathcal{P}(\mathcal{L})$$

where $\mathcal{P}(\mathcal{L})$ denotes the power set of $\mathcal{L}$, i.e. the set of all subsets of $\mathcal{L}$.

3.2 Time and time windows

In this document, we consider a time discretization: $t \in \{0, \dots, T-1\}$. We also assume that the time horizon $T \in \mathbb{N}$ is a multiple of the yearly number of hours, i.e.:

$$\exists k_T \in \mathbb{N} : T = k_T * T_h$$

where T_h is the number of hours in any non-leap year, i.e. $T_h = 8760$.

Let $\Delta \in \{1, \dots, T_h\}$ be a time length less or equal than the total number of hours in a non-leap year. Then, a Δ-time window starting at time t is defined as

$$w_t^\Delta = [t, t + \Delta - 1] \cap \mathbb{N}$$

In other words, a Δ-time window starting at time t is represented by the subset of $\mathbb{N}$ including all natural numbers ranging from t to $t+\Delta-1$ (both included). For a given time period starting at time T_s and finishing at time $T_f > T_s$ such that $(T_f - T_s) \geq \Delta$ and over which a signal of choice is studied, the set of all Δ-time windows that can be extracted, denoted as $\mathcal{W}_{T_s,T_f}^\Delta$, can be further defined as

$$\mathcal{W}_{T_s,T_f}^\Delta = \left\{ w_t^\Delta | t \in \{T_s, \dots, T_f - \Delta\} \right\}$$

In practice, T_s is often taken as $T_s = 0$, whereas the value of T_f can change from one study to another. For the sake of clarity, the set of Δ-time windows will be denoted as $\mathcal{W}^\Delta$ in the sequel.

3.3 Data

The approach proposed in this document relies on building a set of methodologies to compute correlation indicators from a set of wind data. This set of data is a collection of wind signals associated to locations. Formally, a wind signal $\left(s_t^l\right)_{t \in \{0,\dots,T-1\}}$ is a non-negative time series associated with a specific location $l \in \mathcal{L}$:

$$\forall l \in \mathcal{L}, \forall t \in \{0, \dots, T-1\}, s_t^l \in \mathbb{R}_{\geq 0}$$

The type σ specifies the type of resource (e.g., solar or wind), and the set Σ denote the set of all available types of resources.

3.4 Load factor time series and vectors

Let f denote a transfer function of a wind turbine. This function associates, for a given wind signal, an associated power signal, a load factor value:

$$\forall l \in \mathcal{L}, f\left(s_t^l\right) \in [0, 1].$$

In the following, we also denote by $\left(u_t^l\right)_{t\in\{0,\ldots,T-1\}}$ the time series defined as follows:

$$u_t^l = f\left(s_t^l\right).$$

This time series corresponds to the power signal of the wind signal s_t^l.

Also, we introduce a specific notation to define a vector extracted from a time series over a time window w_t^Δ:

$$\begin{array}{l}\forall\Delta\in\{1,\ldots,T_h\},\forall t\in\{0,\ldots,T-\Delta\},\\ \mathbf{u}^l\left(w_t^\Delta\right)=\left[u_t^l,u_{t+1}^l,\ldots,u_{t+\Delta-1}^l\right]\in[0,1]^\Delta.\end{array}$$

3.5 Measure mappings

For any Δ, we assume that we have access to a mapping assessing, to any load factor vector of size Δ, a measure. Let us denote by N_Δ such a mapping:

$$\begin{array}{rcl}N_\Delta:[0,1]^\Delta & \rightarrow & [0,1]\\ \mathbf{v} & \mapsto & N_\Delta(\mathbf{v}).\end{array}$$

In the following, we will consider two variants of such a mapping, in particular, the following mappings $N_{\Delta,1}$ and $N_{\Delta,\infty}$:

$$\begin{array}{rcl}\forall\mathbf{v}\in[0,1]^\Delta, N_{\Delta,1}(\mathbf{v}) & = & \dfrac{1}{\Delta}\displaystyle\sum_{i=1}^{\Delta}\mathbf{v}(i)\\ N_{\Delta,\infty}(\mathbf{v}) & = & \displaystyle\max_{i\in\{1,\ldots,\Delta\}}\mathbf{v}(i)\end{array}$$

3.6 Set of (Δ,α)-critical time windows for a given location

Let $\alpha\in[0,1]$ be a load factor threshold. For a given location $l\in\mathcal{L}$, we denote by $\mathcal{C}_{\Delta,\alpha}^l$ the set of critical time windows:

$$\forall\sigma\in\Sigma,\forall l\in\mathcal{L},\mathcal{C}_{\Delta,\alpha}^l=\left\{w\middle| w\in\mathcal{W}^\Delta, N_\Delta\left(\mathbf{u}^l(w)\right)\leq\alpha\right\}$$

The set of critical time windows gather all Δ-time windows during which the mapping N_Δ of these time windows is smaller than α.

3.7 Intersection of sets of critical time windows for a given type of source signal, and a given technology

Let $L\in\mathcal{P}(\mathcal{L})$ be a subset of locations of wind signals. We introduce the set $\Gamma_{\Delta,\alpha}^{\sigma,i}(L)$ as being the intersection of critical time windows over the subset of locations L:

$$\forall L\in\mathcal{P}(\mathcal{L}),\forall(\Delta,\alpha)\in\{1,\ldots,T_h\}\times[0,1],\Gamma_{\Delta,\alpha}(L)=\bigcap_{l\in L}\mathcal{C}_{\Delta,\alpha}^l$$

Intuitively, such a set $\Gamma_{\Delta,\alpha}(L)$ is a set of time windows during which, wherever the location $l\in L$, the mapping N_Δ of the time window is lower than α.

3.8 Assessing the criticality of a set of locations

Let $L \in \mathcal{P}(\mathcal{L})$ by a set of locations. We define the criticality of relying on a the set of locations L for harvesting energy as the cardinality of the intersection of sets of critical windows as a mapping $\gamma_{\Delta,\alpha} : \mathcal{P}(\mathcal{L}) \to [0,1]$:

$$\forall L \in \mathcal{P}(\mathcal{L}), \gamma_{\Delta,\alpha}(L) = \frac{Card\left(\Gamma_{\Delta,\alpha}(L)\right)}{Card(\mathcal{W}^{\Delta})}$$

where $Card$ is a mapping that associates, to any finite set S, the cardinality of S. Concretely, $\gamma_{\Delta,\alpha}(L)$ is the proportion of critical Δ-time windows within the set of localizations L and considering the measure mapping N_{Δ}. The mapping $\gamma_{\Delta,\alpha}$ is used in the following of this paper as a fundamental ingredient in the definition of other indicators and optimization problems.

3.9 Computing the average proportion of critical windows found in a given zone $L \in \mathcal{P}(\mathcal{L})$

Let us define the first indicator $I^{(1)}_{\Delta,\alpha} : \mathcal{P}(\mathcal{L}) \to [0,1]$,

$$\forall L \in \mathcal{P}(\mathcal{L}), I^{(1)}_{\Delta,\alpha}(L) = \frac{1}{Card(L)} \sum_{l \in L} \gamma_{\Delta,\alpha}(\{l\})$$

3.10 Computing the average proportion of critical windows common to all locations in a given zone $L \in \mathcal{P}(\mathcal{L})$

Let us define the second indicator $I^{(2)}_{\Delta,\alpha} : \mathcal{P}(\mathcal{L}) \to \mathbb{R}$

$$\forall L \in \mathcal{P}(\mathcal{L}), I^{(2)}_{\Delta,\alpha}(L) = \gamma_{\Delta,\alpha}(L)$$

3.11 Find in a given zone $L \in \mathcal{P}(\mathcal{L})$, n points such that the proportion of critical windows, common to all signals in the zone is maximal

Let us define the third mapping $I^{(3)}_{\Delta,\alpha} : \mathcal{P}(\mathcal{L}) \times \mathbb{N} \to \mathbb{R}$:

$$\forall L \in \mathcal{P}(\mathcal{L}), \forall n \in \mathbb{N}, I^{(3)}_{\Delta,\alpha}(L,n) = \max_{\substack{L' \subset L \\ \text{s.t. } Card(L') = n}} I^{(2)}_{\Delta,\alpha}(L')$$

We also introduce a mapping $\mathcal{P}^{(3)}_{\Delta,\alpha} : \mathcal{P}(\mathcal{L}) \times \mathbb{N} \to \mathcal{P}(\mathcal{L})$ that outputs a set of sets of locations $\mathcal{P}^{(3)}_{\Delta,\alpha}(L,n)$ such that:

$$\forall L \in \mathcal{P}(\mathcal{L}), \forall n \in \mathbb{N}, \mathcal{P}^{(3)}_{\Delta,\alpha}(L,n) = \arg\max_{\substack{L' \subset L \\ \text{s.t. } Card(L') = n}} I^{(2)}_{\Delta,\alpha}(L')$$

By definition, one has:

$$\forall L' \in \mathcal{P}^{(3)}_{\Delta,\alpha}(L,n), I^{(2)}_{\Delta,\alpha}(L') = I^{(3)}_{\Delta,\alpha}(L,n)$$

3.12 Find in a zone $L \in \mathcal{P}(\mathcal{L})$, n points such that the proportion of critical windows, common to all signals in the zone is minimal

Let us define a fourth mapping $I^{(4)}_{\Delta,\alpha} : \mathcal{P}(\mathcal{L}) \times \mathbb{N} \to \mathbb{R}$:

$$\forall L \in \mathcal{P}(\mathcal{L}), \forall n \in \mathbb{N}, I^{(4)}_{\Delta,\alpha}(L,n) = \min_{\substack{L' \subset L \\ \text{s.t. } Card(L') = n}} I^{(2)}_{\Delta,\alpha}(L')$$

We also introduce a mapping $\mathcal{P}^{(4)}_{\Delta,\alpha} : \mathcal{P}(\mathcal{L}) \times \mathbb{N} \to \mathcal{P}(\mathcal{L})$ that outputs a set of sets of locations $\mathcal{P}^{(4)}_{\Delta,\alpha}(L,n)$ such that:

$$\forall L \in \mathcal{P}(\mathcal{L}), \forall n \in \mathbb{N}, \mathcal{P}^{(4)}_{\Delta,\alpha}(n) = \arg\min_{\substack{L' \subset L \\ \text{s.t. } Card(L') = n}} I^{(2)}_{\Delta,\alpha}(L')$$

By definition, one has:

$$\forall L' \in \mathcal{P}^{(4)}_{\Delta,\alpha}(L,n), I^{(2)}_{\Delta,\alpha}(L') = I^{(4)}_{\Delta,\alpha}(L,n)$$

4 Application

The purpose of this section is to answer the case study about finding the correlation between wind signals in Greenland and Europe.

Previously in section 3, we have defined a correlation metric able to compare wind signals from specific regions. This tool will be used to extract correlation values from winds in Europe and Greenland, and therefore be a way of evaluating if a connection between Greenland and Europe is interesting.

In order to run this tools, it requires wind data from the past years, especially in Europe and Greenland. Hence, the following sections will explain how to fetch wind data and how to use them.

4.1 Meteorological model and reanalysis

A meteorological model is a mathematical model used to predict weather based on current weather conditions. To generate the pieces of data, these models use meteorological reanalyses. It consists of meteorological methods used to forecast and assimilate historical observations from weather stations. It presents a clear picture of past weather, independent of the instruments used to take measurements.

Around the world, there exists lots of weather stations that measure atmospheric conditions. These measures provide information allowing weather forecasts and weather and climate studies at specific locations on the earth.

The models that we are interested in use Numerical weather prediction models (NWP models) and data assimilation systems to reanalyze the observations, creating global data sets. Reanalysis also holds the NWP model constant, keeping the historical record uninfluenced by

artificial factors.

It is important to keep in mind that data exported from reanalyses are `estimations` of weather data. Reanalysis signals are less sensitive to extremes and tend to underestimate them.

This next sections explain what is a reanalysis of observations, how to extract contents from reanalysis, the possibilities of data exportation from this database, and finally what is exported for our study.

4.2 Wind turbine VESTAS V80-2MW

In the formalization at section 3.4, we introduced a transfer function denoted f. This function is used to associate a power signal u_t^l to a given wind signal s_t^l:

$$u_t^l = f(s_t^l)$$

For our applications, we decided to choose a transfer function of an existing wind turbine. We selected the `VESTAS V80-2MW` power curve. Indeed, this wind turbine fetches the wind speeds at 100 meters high, has a rate power of $2MW$, and has a cut-out speed of $25m/s$ (see appendix 7.1 for more information about the wind turbine configuration).

The following table 1 provides relations between the wind speeds W_tab and the produced powers P_tab for specific wind values. These are given by the wind turbine manufacturer.

W_tab (m/s)	3	4	5	6	7	8	9	10	11	12	13	14
P_tab (kW)	0	58	149	277	461	698	996	1331	1645	1854	1955	1988

15	16	17	18	19	20	21	22	23	24	25	26	27
1996	1998	2000	2000	2000	2000	2000	2000	2000	2000	2000	0	0

Table 1: Relation between wind values from W_tab and power values from P_tab for the VESTAS V80-2MW.

Let $W\_tab[i]$ refers to the i^{th} element of the first row in table 1, e.g. $W\_tab[3] = 5$ m/s. Similarly, let $P\_tab[i]$ refers to the i^{th} element of the second row in table 1, e.g. $P\_tab[3] = 149$ kW.

Let us illustrate a model of the power curve of the wind turbine VESTAS V80-2MW by drawing a continuous curve on a graph, joining all the points from table 1. The figure 1 shows the relation between the wind speeds and the power production, the coordinates from the previous table 1 being interpolated.

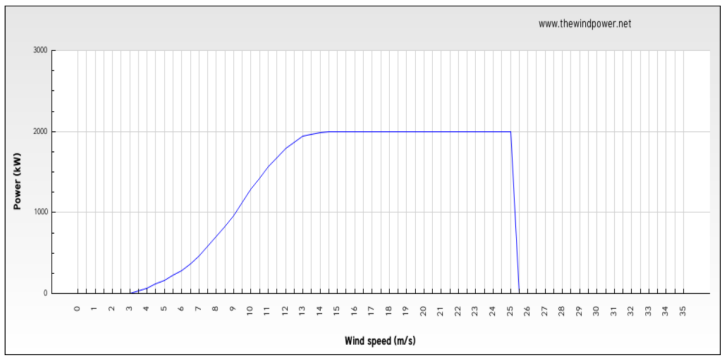


Figure 1: Power curve of VESTAS V80-2MW (from www.thewindpower.net).

The analytic expression of the model of the `VESTAS V80-2MW` transfer function is defined as follows:

Let s_t^l a wind signal at a given location l and $\left(s_t^l\right)_t$ the wind speed of the wind signal at time t. We obtain the associated power value $\left(u_t^l\right)_t$ by applying the Vestas transfer function on the wind speed. We do this as follows:

Let i the index $\in [1;24] \cap \mathbb{N}$ of W_tab such that

$$W\_tab[i] \leq \left(s_t^l\right)_t < W\_tab[i+1]$$

The model of the VESTAS V80-2MW power curve f_{vestas} is defined as:

$$f_{vestas} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$$

$$x \mapsto \begin{cases} P\_tab[i] + (\frac{x - W\_tab[i]}{W\_tab[i+1] - W\_tab[i]}) * (P\_tab[i+1] - P\_tab[i]) & \text{if } x \in [3;27[\cap\mathbb{N} \\ 0 & \text{else} \end{cases} \quad \forall t$$

Hence, the model of the `VESTAS V80-2MW` transfer function will be used to convert a wind signal into a power signal in the following applications. To do this, we simply apply this transfer function to a wind signal s_t^l:

$$u_t^l = f_{vestas}\left(s_t^l\right)$$

4.3 Weather data from NASA

In this section, we will describe how the NASA stores weather data, how to export data relevant for our study, and finally how to manipulate exported data to get wind signals from Europe and Greenland during the period from 1980 to 2016.

4.3.1 Possibilities of data exportation from the NASA database

The NASA uses reanalysis to compute a lot of information about geographic data on the Earth (see subsection 4.1). They gather information about atmospheric composition (e.g. Carbon Monoxyde, Ozone), Water and Energy Cycles (e.g. radiation, optical depth), and Climate

Variability (e.g. temperature, wind speed, clouds). All these pieces of data are stored in a database and are freely accessible on their official `website`.

On the `website` of the NASA database, a user is able to export various data from this database. The use can specify his request by specifying some available exportation options.

1. Data category. The user selects the category of the data he wants to export. A menu invites him to choose the category of data to export (e.g. temperature, wind speed, optical depth, radiation). In addition, some options can be chosen. Depending on the type of data, one can export data sampled 1-hourly, 3-hourly, instantaneous, or averaged.

2. Geographical zone. Once the type of data and the options chosen, the user has to select a geographical zone in order to extract data from the selected area. He must encode the minimum longitude, minimum latitude, maximum longitude, and maximum latitude of the desired geographical zone.

3. Temporal interval. After selecting the type and the geographical zone of data, the user can select a temporal interval to extract data only from this interval of time. All data are available starting from the 1^{st} of January 1980.

4. Parameters. Depending on the category previously selected, data type will be available according to the category. These data types are presented as a list of parameters that the user can check. All the checked parameters will be exported at the end of the process.

5. Extra options. Other options are available, depending on the category of wanted data to extract (e.g. time subset, regridding, output file format).

6. Exportation. At the end of the selection, the user can download weather data he chose from the selected geographical zone and the temporal interval. By clicking on the download button, he has access to the links that will download the weather files.

The downloaded files are in a netCDF4 format. Each file contains multidimensional array filled with values representing exported data over the different dimensions. The size of these arrays depends on the selected geographical area, on the interval time, but also on the number of samples that the NASA extracts from this area. In fact, the NASA database contains information for specific geographical coordinates, not for all coordinates on the Earth. Indeed, the NASA gathers data for discretized coordinates. On map surface, data is sampled at a coordinate resolution of $\frac{2}{3}$-degree (precisely 0,625) longitude by $\frac{1}{2}$-degree latitude (see `File Specification on NASA database website, page 7`).

4.3.2 Exported data from the NASA database

For our study, we export data related to wind speeds for Europe and Greenland, from 1980 to 2016. Hence, here are the entries that were selected from the NASA website:

- `Data Product`: MERRA-2 inst1_2d_asm_Nx (first on the list)
- `Spatial Search`:

  - `for Greenland`: West=-56.777, South=57.535, North=88.473, East=-12.129 (Figure 2)
  - `for Europe`: West=-10, South=16, North=71, East=32 (Figure 3)
- `Temporal Order Option`: from 1980, jan 01 to 2016, dec 31 (37 years)
- `Parameters`[1]: the following parameters were exported:
  - `U10M`: Eastward wind at 10 m above displacement height.
  - `U50M`: Eastward wind at 50 m above surface.
  - `V10M`: Northward wind at 10 m above the displacement height.
  - `V50M`: Northward wind at 50 m above surface.
- `Additional Options`: Unchanged (–)
- `Regridding Options`: Unchanged (NONE)
- `Output File Format`: Unchanged (netCDF4)

Europe and Greenland maps.

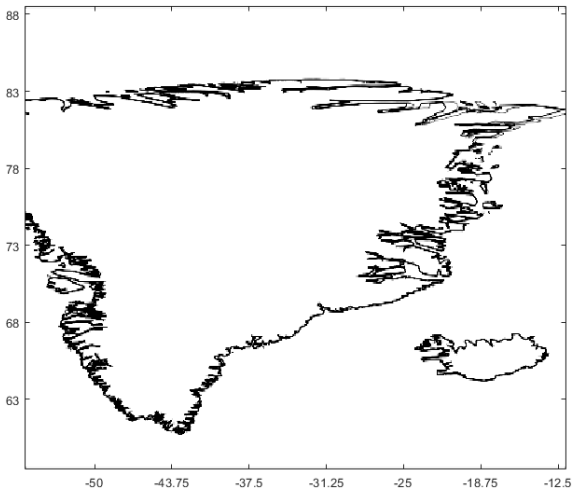


Figure 2: Map of Greenland.

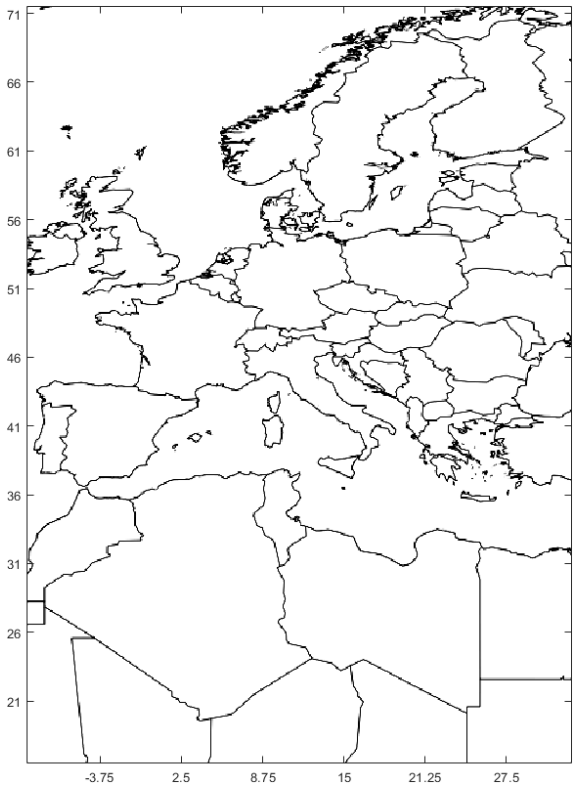


Figure 3: Map of Europe.

Once all the options well configured, a `Start Search` button redirects the user to a page explaining how to download the pieces of data.

[1]As explained in section 4.3.1, parameters are data types depending on the category previously selected

Once redirected, lots of netCDF4 files (.nc4, one file per day) containing all the wind speeds in m/s will be available and downloadable (see File Specification on NASA database website, page 48).

4.3.3 NetCDF4 structure

NetCDF4 is a data format that supports the creation, access, and sharing of array-oriented scientific data that is self-documented. This format requires specific methods in order to extract data from the files. This kind of functions are easily usable via Python (python.org/netCDF4) or via Matlab functions from the `netCDF4 library`.

Each downloaded netCDF4 file represents extracted weather data of a whole day. As explained previously, all these netCDF4 files contain grids of wind signals, one signal for each sampled coordinates from the geographical zone. Since downloaded wind speeds are sampled hourly by the NASA, these signals are composed of 24 values.

Each netCDF4 file contains exactly four grids of wind signals[2], one for each parameters (U10M, U50M, V10M, and V50M). These four grids are accessible via Python or Matlab as 3D matrices filled with double values representing the wind speed signals over 24 hours. The first dimension of these matrices is the `longitude index` of the geographical zone, the second dimension is the `latitude index` of the geographical zone, and the last dimension is the `time index`. Here are the sizes of the matrices:

- `Greenland`: 71 x 61 x 24 matrix, as shown in figure 3,
- `Europe`: 68 x 111 x 24 matrix, as shown in figure 2.

4.3.4 NASA data reshaping

As explained in the previous section, a single netCDF4 file represents 24 hours of data for an entire zone (Europe or Greenland). For the 37 years of data, from 1980 to 2016, there are 10 leap years and 27 that are not. In total, there are

$$24 \times 366 \times 10 + 24 \times 365 \times 27 = 324360 \text{ hours}$$

that are split into

$$324360/24 = 13515 \text{ netCDF4 files}$$

for the Europe and the Greenland zone.

Previous section 4.3.3 explained that files from NASA contain wind values split into 4 different 3D grids: one for eastward winds at 10 meters high, one northward winds at 10 meters high, one for eastward winds at 50 meters high, and one for northward winds at 50 meters high. For our study, the direction of the wind is not important. Only the wind resultant force that matters.

To obtain the resultant force of a wind signal s_t^l at 50 meters high and at a given location $l = (\lambda^{long}, \lambda^{lat})$, we apply the Pythagorean theorem on the eastward wind and on the northward wind for each time t:

[2] A wind signal consists of a sequence of successive wind values in time. When talking about a wind signal from a specific location in Europe (resp. Greenland), it corresponds to the sequence of successive wind values that occur at this location, from 1^{st} January 1980 to $31th$ December 2016

$$\forall t \in 0, ..., T-1, s_t^l = \sqrt{\left(\texttt{U50M}_t^l\right)^2 + \left(\texttt{V50M}_t^l\right)^2}$$

with

- T: the total number of wind values into the wind signal s_t^l (324360 in our case),
- $\texttt{U50M}_t^l$: the eastward wind value at 50 meters high, at location l, and at time t,
- $\texttt{V50M}_t^l$: the northward wind value at 50 meters high, at location l, and at time t.

We can also precise that $l \in \mathcal{L}_E$ in case of Europe location ($\mathcal{L}_\mathcal{E}$ is the set of all Western Europe locations) and $l \in \mathcal{L}_G$ in case of Greenland location ($\mathcal{L}_\mathcal{G}$ is the set of all Greenland locations). More precisely,

$$\forall l = (\lambda^{long}, \lambda^{lat}) \in \mathcal{L}_E : \lambda^{long} \in \{-10, -9.375, ..., 31.875\}, \lambda^{lat} \in \{16, 16.5, ..., 71\}$$

$$\forall l = (\lambda^{long}, \lambda^{lat}) \in \mathcal{L}_G : \lambda^{long} \in \{-56.25, -55.625, ..., -12.5\}, \lambda^{lat} \in \{58, 58.5, ..., 88\}$$

To read the $\texttt{U50M}_t^l$ and $\texttt{V50M}_t^l$ values for each time $t \in 0, ..., T-1$, we have to load sequentially all the 13515 files from a zone (Greenland or Europe). This is not efficient enough. Sometimes, we have to load all the wind signals from a given zone (Europe or Greenland) sequentially in order to analyze them. For example, Europe has $68 \times 111 = 7548$ wind signals, and it is not efficient to load all the files for each wind signal.

Therefore data has been reshaped in order to access directly a wind signal at a specific location. To do this, a python script was written to load sequentially all netCDF4 files of a zone (Europe or Greenland) and to store one by one the wind signals into separated `numpy` vectors.

Hence, each `numpy` vector contains 324360 values, one for each wind speed value at the location l represented by the vector. The n^{th} values of the vector corresponds to the wind speed at time $t = n - 1$, at the location l. Vectors are then saved into the hard drive disk using the following representative name:

$$z\texttt{\_1980\_2016\_lon}\lambda^{long}\texttt{\_lat}\lambda^{lat}\texttt{\_high}h\texttt{.npy}$$

with

- z: the zone where the location represented by the vector saved in the file is ($z \in \{$`Europe, Greenland`$\}$),
- λ^{long} and λ^{lat}: the longitude and the latitude of the location represented by the vector,
- h: the high of the wind signal represented by the vector ($h \in \{10, 50\}$).

To go further and to improve a little more the computation efficiency, we also decided to save the power signal corresponding to each wind signal into other separated files. Hence, for a location $l = (\lambda^{long}, \lambda^{lat})$, we have a file containing the wind speeds of the wind signal s_t^l (as described previously), but we also created another file containing the power signal $u_t^l = f_{vestas}(s_t^l)$ obtained using the model `Vestas` power curve f_{vestas}. We named these files as follows:

$$z\texttt{\_1980\_2016\_lon}\lambda^{long}\texttt{\_lat}\lambda^{lat}\texttt{\_}h\texttt{high\_power.npy}$$

with

- z: the zone where the location represented by the vector saved in the file is ($z \in \{$`Europe, Greenland`$\}$),
- λ^{long} and λ^{lat}: the longitude and the latitude of the location represented by the vector,
- h: the high of the wind signal represented by the vector ($h \in \{10, 50\}$).

4.3.5 NASA data representation

As explained in the previous section, we have exported wind signals from the NASA database and stored them into numpy vectors. Each vector contains 324360 ordered values, one for each hour of the 37 years of data (from January, 1^{st} 1980 to December, 31^{th} 2016.

When wind data was downloaded from the NASA database, we did a rectangle selection around the wind signals areas of interest. Hence, for the Greenland zone, we have exported

$$71 \times 61 = 4331 \text{ wind signals}$$

and for Europe, we have exported

$$68 \times 111 = 7548 \text{ wind signals}$$

The figures 2 and 3 show the locations of the wind signals that are considered and exported for our study. Each wind signals is represented by a colored dot on the figures.

Europe and Greenland maps.

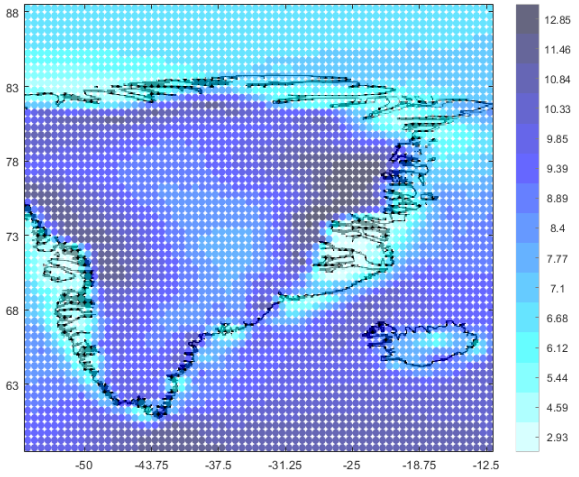


Figure 4: Kmean on means of wind speeds at 50 meters high in Greenland.

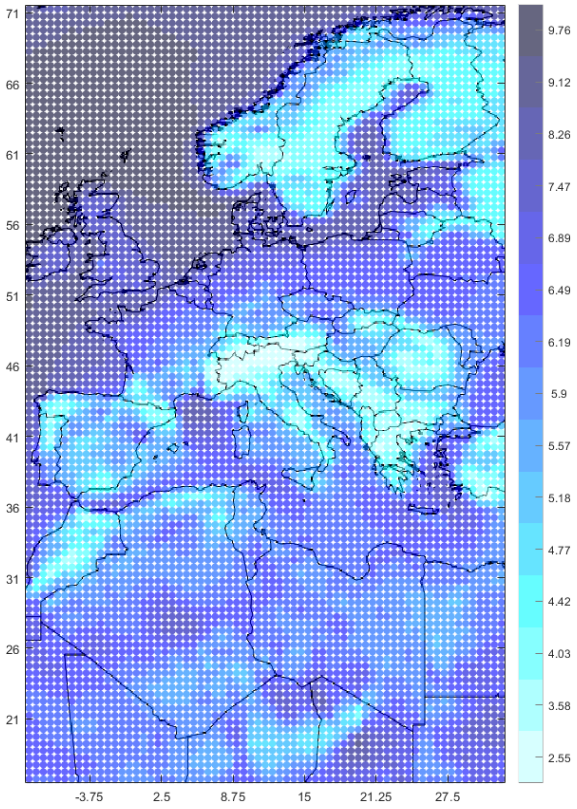


Figure 5: Kmean on means of wind speeds at 50 meters high in Europe.

The color assigned to each dot corresponds to the mean wind value over the 37 years of the wind signal at 50 meters high it corresponds to. To ease the readability of the dot colors, we applied a kmean clustering algorithm on these colored dots (see the appendix 7.0.1 for more details about kmean). Hence, we displays these dots clustered into 15 different clusters. The color assigned to each cluster depends on its centroid. The more the centroid is windy, the darker the assigned color is.

4.4 Weather data from MAR model

The MAR model is a Regional Climate Models (RCM) specifically developed for simulating polar climate specificities, furthermore abundantly evaluated over the Greenland ice sheet [4].

MAR requires prescription of atmospheric fields (temperature, relative humidity, wind speed, pressure) at its lateral boundaries as well as sea surface conditions from a global forcing dataset such as reanalysis like ERA-Interim or MERRA-2 [3].

MAR is well known for its fine representation of physical processes in polar regions. It is one of the few models able to simulate realiztic surface mass balance, air-snow interactions, and atmospheric circulation over ice sheets, including katabatic winds.

4.4.1 Katabatic wind in the MAR model

As explained previously, the MAR model is one of the few models able to simulate further phenomena that other models do not. For example, the MAR model simulates better atmospheric circulation over ice sheets that MERRA-2 do not consider in its model, in particular katabatic winds.

Katabatic winds are winds that carry high-density air from higher elevation down a slope under the force of gravity.
The figure 6 illustrates katabatic winds occur on the Greenland Ice Sheet.

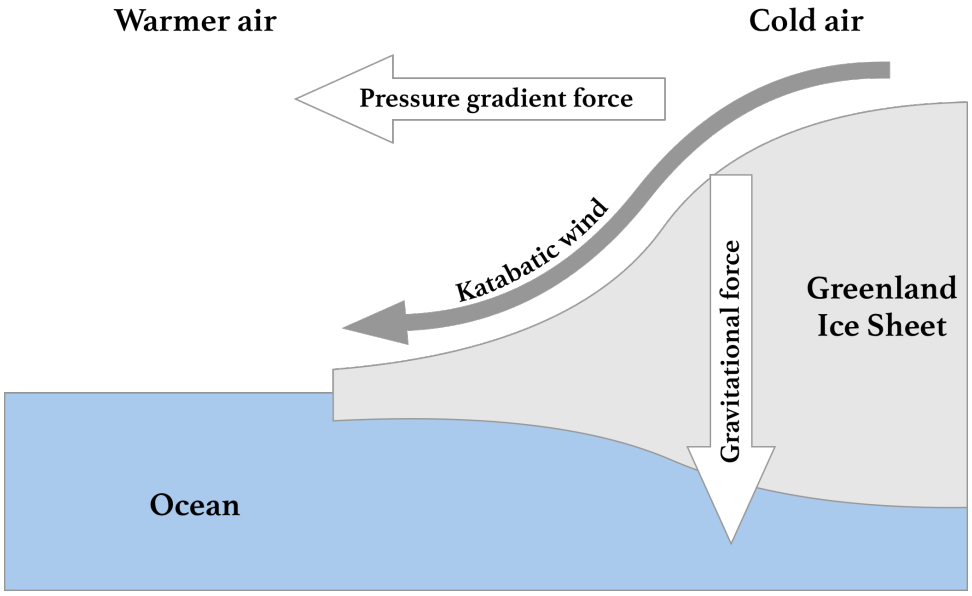


Figure 6: An illustration of katabatic winds in Greenland, carrying high-density air from a higher elevation down a slope under the force of gravity.

Katabatic winds are created at the ice sheet level (from 0 to 10 meters high) and blow to the coast. When moving to the coasts, the katabatic winds are influence by the relief of the land and may reach several meters high once on the coasts. Hence, at a certain height, these winds can be interesting for our study. They may be convenient for wind turbine to produce continuous high power.

Since MERRA-2 do not consider katabatic winds, we will perform some of the following computations twice. Once with the MERRA-2 data, and once with MAR data. Then, comparisons between both datasets will follow.

4.4.2 Exported data from the MAR model

For the second database, we were in contact with Dr. Xavier Fettweis, a Climatologist researcher who studies in particular the MAR model.

In order to fill our needs, Xavier Fettweis kindly run the MAR model on the Greenland map. He used MAR version 3.9 with lateral boundaries forced using MERRA-2 reanalysis.

Hence, he was able to provide wind data from 2000 to 2017, at the same scale than the MERRA-2 data scale (0.5×0.625). He generated results for a precise region of Greenland that seemed to be more windy and interesting for our study. This region of interest is located at the south coast of Greenland. More precisely, a location $l = (\lambda^{long}, \lambda^{lat})$ in Greenland belongs to the region of interest if

$$-50.625 \leq \lambda^{long} \leq -40 \quad \text{and } 59.5 \leq \lambda^{lat} \leq 63.5$$

The figure 7 shows the selected region of interest (in ligh color).

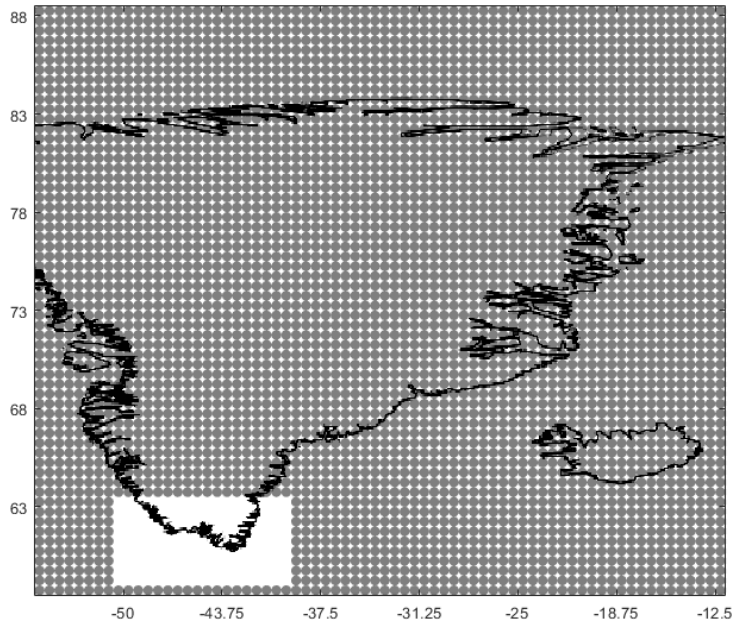


Figure 7: Representation of the region of interest in Greenland from Mar model (white area)

Indeed, in the next sections, we will show that this region appears to have interesting wind speeds.

From the MAR model, we exported all the wind signals of the region illustrated at figure 7. We will call this set of location:

$$\mathcal{L}_{MAR}$$

and contains the 162 following locations:

$$\forall l = (\lambda^{long}, \lambda^{lat}) \in \mathcal{L}_{MAR} : \lambda^{long} \in \{-50.625, -50, ..., -40\}, \lambda^{lat} \in \{59.5, 60, ..., 63.5\}$$

Exported data are stored into 36 netCDF4 files (see section 4.3.3 for further information about netCDF4 files). 18 netCDF4 files are used to store information about hourly wind speeds at 10 meters high for locations $l \in \mathcal{L}_{MAR}$ from January, 1^{st} 2000 to December, 31^{th} 2017. The 18 others files are used to store information about same hourly wind data as previous ones, except for the wind speeds at 50 meters high. Since there is 18 netCDF4 files per wind speed high, each netCDF4 file contains data of a specific year between 2000 and 2017 for the given wind speed high (10 or 50 meter).

In the netCDF4 files, we found similarities with the netCDF4 files from NASA. Indeed, here are the matrices stored into the netCDF4 files generated by MAR model that contain wind data:

- UUZ: matrix that contains the Eastward wind at 10 meters or 50 meters high above surface (similar to U10M and U50M from MERRA-2 netCDF4 files),
- VVZ: matrix that contains the Northward wind at 10 meters or 50 meters high above surface (similar to V10M and V50M from MERRA-2 netCDF4 files).

4.4.3 MAR data reshaping

In comparison to section 4.3.4, we reshaped the MAR databases in order to be more efficient when loading data of a single location $l \in \mathcal{L}_{MAR}$. Indeed, for the 18 years of data, from 2000 to 2017 with 5 leap years and 13 that are not, there are

$$366 \times 24 + 365 \times 24 \times 13 = 157800 \text{ hours}$$

that are split into 18 netCDF4 files for wind speeds at 10 meters high and into 18 other netCDF4 files for wind speeds at 50 meters high.

As explained in previous section 4.4.2, netCDF4 files contains wind values splits into 2 different grid: one for eastward winds and the other for northward wind. Since the direction is not important for our study, we only kept the wind resulting force.

To compute the resulting force of a wind signal s_t^l at a location $l \in \mathcal{L}_{MAR}$, we apply the Pythagorean theorem on the eastward wind and on the northward wind for each time t:

$$\forall t \in 0, ..., T-1, s_t^l = \sqrt{(\text{UUZ}_t^l)^2 + (\text{VVZ}_t^l)^2}$$

with

- T: the total number of wind values into the wind signal s_t^l (157800 in our case),
- UUZ_t^l: the eastward wind value at 10 meters or 50 meters high, at location $l \in \mathcal{L}_{MAR}$, and at time t,
- VVZ_t^l: the northward wind value at 10 meters or 50 meters high, at location $l \in \mathcal{L}_{MAR}$, and at time t.

Using this technique, we are able to generate all wind signals s_t^l for location $l \in \mathcal{L}_{MAR}$ as `numpy` vectors and to store them into separed files.

Hence, each `numpy` vector contains 157800 values, one for each wind speed at the location l represented by the vector. The n^{th} value of the vector corresponds to the wind speed at time $t = n - 1$, at the location l. Vectors are then saved into the hard drive disk using the following representative name:

$$\texttt{greenland\_2000\_2017\_lon}\lambda^{long}\texttt{\_lat}\lambda^{lat}\texttt{\_high}h\texttt{.npy}$$

with

- λ^{long} and λ^{lat}: the longitude and the latitude of the location represented by the vector,
- h: the high of the wind signal represented by the vector ($h \in \{10, 50\}$).

4.4.4 MAR data representation

As explained in section 4.4.2, we have 162 locations as region of interest. At each of these locations $l \in \mathcal{L}_{MAR}$ corresponds a wind signal s_t^l containing wind values from January, 1^{st} 2000 to December, 31^{th} 2017. In order to illustrate these 162 wind signals, we computed the mean value of each wind signal at 50 meters high. A kmean algorithm has then been applied to these 162 means using 15 clusters. The figure 8 shows the results of the kmean algorithm:

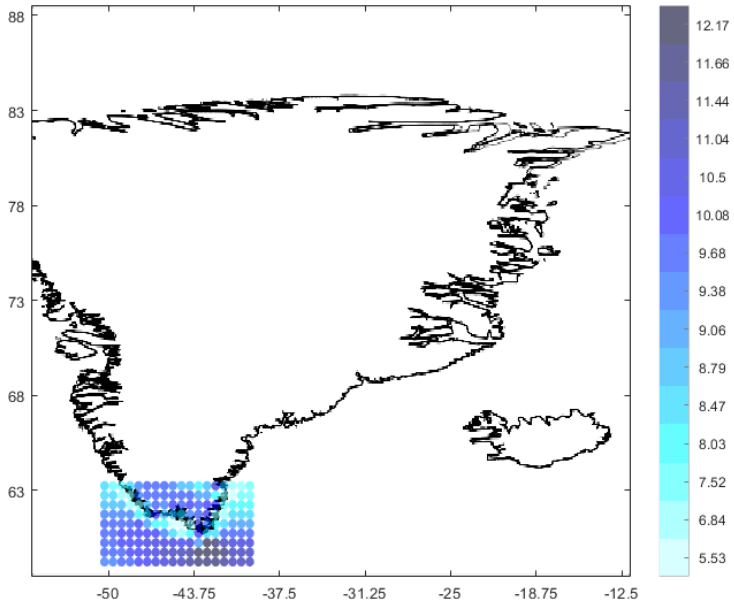


Figure 8: kmean using 15 clusters applied to the mean of the wind signals of region of interest for MAR data

4.5 Comparisons between both weather models

Even though MAR model is forced at its lateral boundaries with data from MERRA-2 model, both models do not output same data. Indeed, this subsection will show that MAR model outputs more accurate climate data in Greenland due to the extra computation of katabatic winds.

In order to compare wind data from MERRA-2 and MAR models, we select a common region of interest and a common time interval that occurs in both databases. For the common region of interest that will be used, we selected the region exported with MAR model as it is fully included in the region exported with MERRA-2 model. For the common time interval, since MERRA-2 database includes data from 1980 to 2016 and MAR database includes data from 2000 to 2017, we used the period beginning on the 1^{st} January 2000 and ending on the 31^{th} December 2016 to compare both databases.

Hence, the figures 9 and 10 represent the means of the wind signal at 50 meters high during the common time interval at each location from the common region of interest. For better readability, mean values are clustered into 15 different clusters using kmean algorithm.

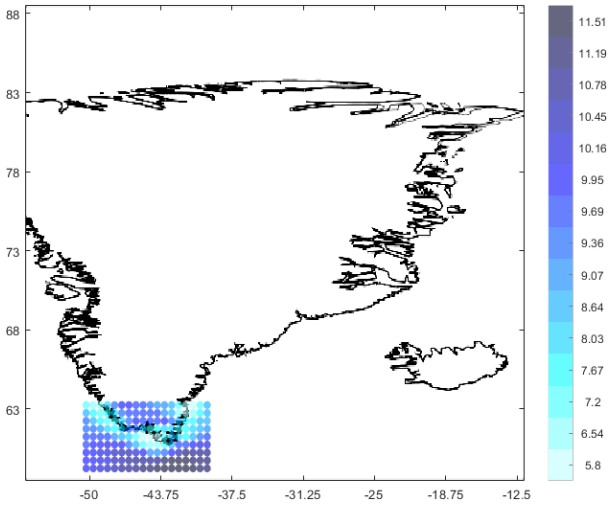


Figure 9: kmeans on Greenland data generated with MERRA-2 model

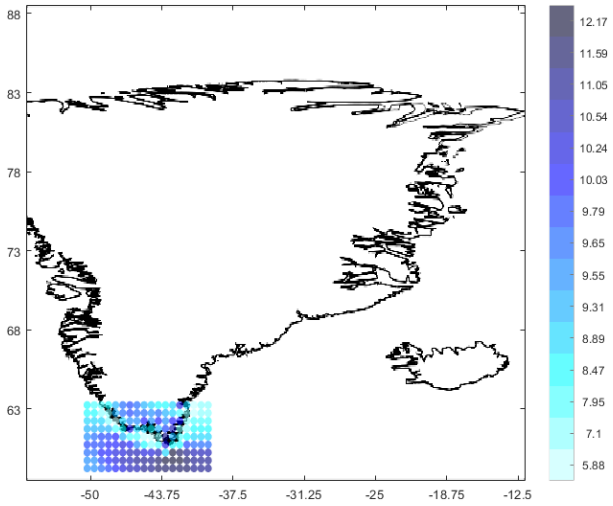
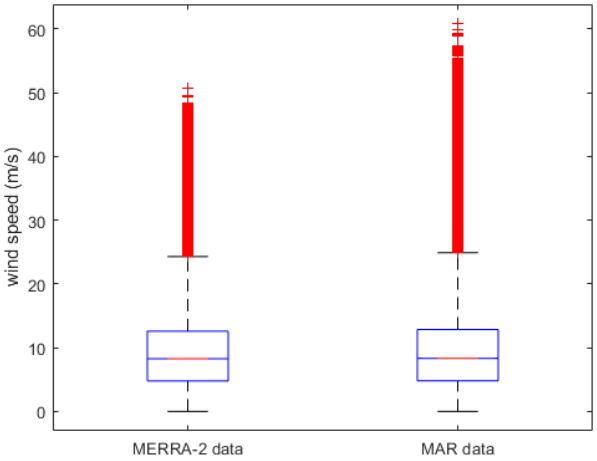


Figure 10: kmeans on Greenland data generated with MAR model

We can see the zone with strong wind signals (dark blue) is more expanded with MERRA-2 data than with MAR data. But, within this strong wind zone, wind signals are stronger with MAR model than with MERRA-2 data.

On figure 11, boxplots have been generated by considering all the values from wind signals in the common region of interest of Greenland during the common time interval for MERRA-2 data and for MAR data.



	MERRA-2 model	**MAR model**
maximum	50.82	60.86
minimum	0.00	0.00
median	8.28	8.34
first quartile	4.79	4.82
third quartile	12.59	12.86

Figure 11: At left-hand side, boxplots comparing wind speed data from MERRA-2 model and from MAR model. At right-hand side, results from boxplots are summarized.

By looking both boxplots, values from MERRA-2 data and from MAR model seem to share the same distribution. Indeed, median, quartiles and whiskers are seemly equal on both boxplots. The difference remains in the outliers. MAR data has bigger extremes than MERRA-2 data.

In order to have a better view of the difference between both datasets, we selected a critical location on the Greenland coast at coordinates (-43.75,60.5) (see figure 12) and we generated

some boxplots according to this location (see figure 13).

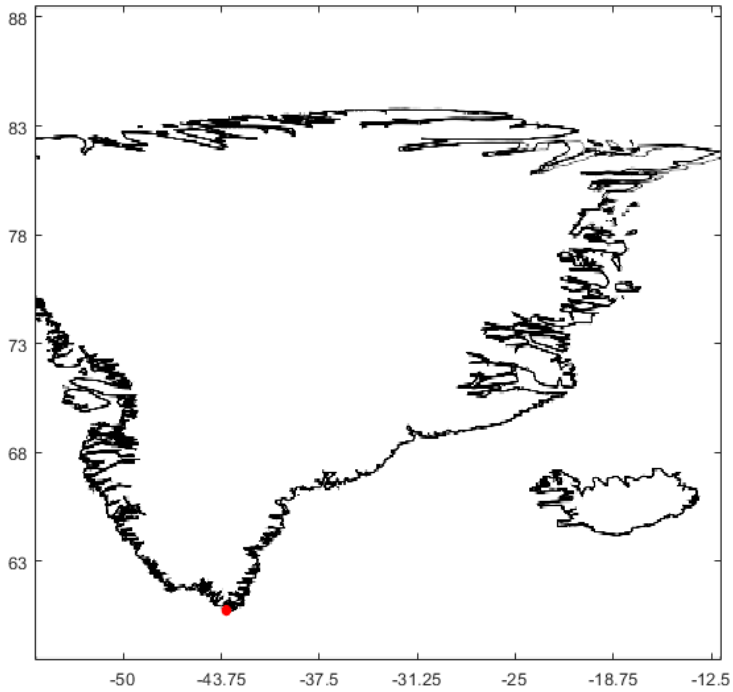
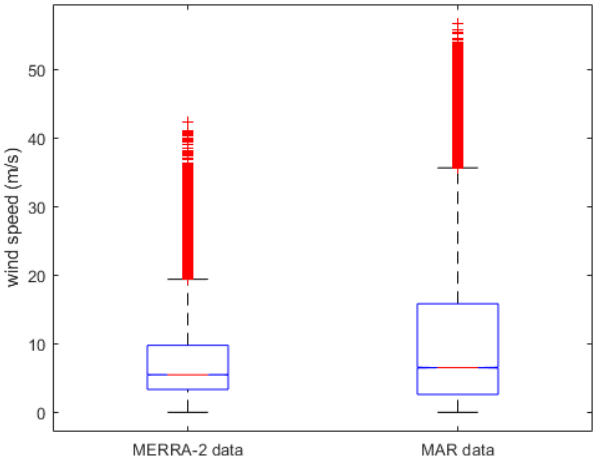


Figure 12: location selected in Greenland coast at coordinates (-43.75,60.5) represented by a red dot.



	MERRA-2 model	**MAR model**
maximum	42.27	56.65
minimum	0.02	0.01
median	5.49	6.52
first quartile	3.34	2.63
third quartile	9.77	15.84

Figure 13: At left-hand side, boxplots comparing wind speed data of location at coordinates (-43.75,60.5) from MERRA-2 model and from MAR model. At right-hand side, results from boxplots are summarized.

At this location, there is a significant difference between wind values from both models. Minimum value and first quartile are bigger for MERRA-2 data, but median, first quartile, third quartile and maximum value are bigger for MAR data.

These figures introduce the idea of bigger wind signal for MAR model than for MERRA-2 model in general.

A last interesting analysis would to evaluate how models are accurate. Hence, the figure 14 tries to illustrate the accuracy of both models by comparing their wind speed data with real climate observations at QAS_L over 2015 and 2016 (from www.promice.dk).

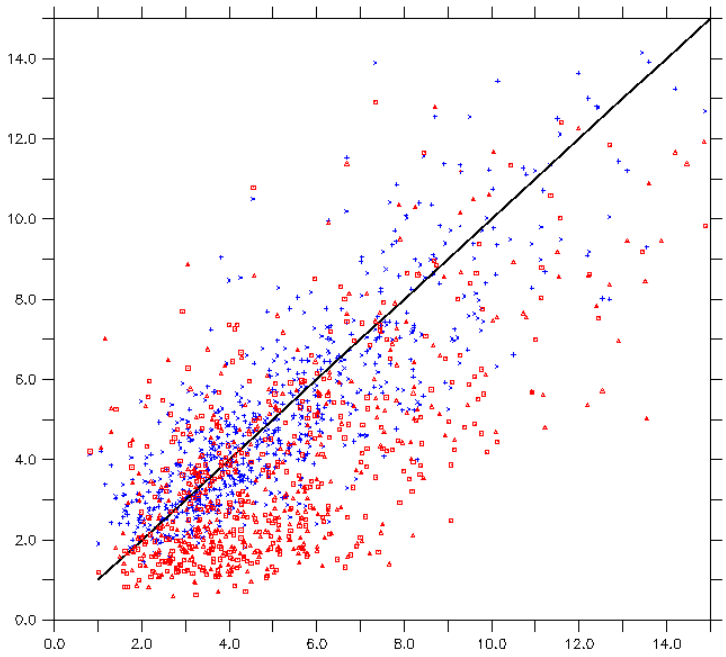


Figure 14: scatter plot of MAR(blue) or MERRA-2 (red) vs observations.

As results, we obtain the following statistics:

	R	RMSE
MERRA-2 model	0.65	2.6 m/s
MAR model	0.83	1.6 m/s

Due to the previous observation, MAR model tends to be more accurate than MERRA-2 on Greenland coast.

4.6 Question 1: average proportion of critical windows found in a given zone

4.6.1 Introduction

A lot of question are interesting to answer, regarding the databases we have and the tool we developed. The formalization described in section 3 will be used in this section and the following ones to generate results about critical windows. As reminder, a critical window is an time Δ of a power signal during which the average power production of this signal is below a certain threshold α.

Based on the formalization, we will define new concepts and answer some questions in order to provide information about the number of critical windows in a specific zone. First, we will

answer questions by considering different regions as Europe, France, Greenland, Greenland and Europe, and Greenland and France. The two last regions, when considering Greenland with Europe of France, correspond to the situations when these regions are connected together (using a physical submarine cable for example). Finally, we will compare the different results and observe if the link between Greenland and Europe modifies these results.

The first question that will be answered is the following one:

Question 1

In a specific zone, what is the average proportion of critical windows over the 17 years for a wind signal found in this zone ?

4.6.2 First methodology

In order to answer the `question 1`, we will use the related indicator $I^{(1)}_{\Delta,\alpha}$ defined at subsection 3.9. Using this, we are able to answer directly the question.

The indicator $I^{(1)}_{\Delta,\alpha}$ depends on two variables: the length Δ and the threshold α of the windows that will be considered on the power signals. Hence, we will answer `question 1` by using several values of the length Δ and the threshold α.

The length Δ of the windows will take the successive values `24, 48, 72, 96, 120, 144`, and finally `168` hours. For the threshold, we will consider five different values: `10%, 15%, 20%, 25%`, and finally `30%`.

In order to generate the first results, we will consider the following variant mapping $N_{\Delta,\infty}$:

$$\forall \mathbf{v} \in [0,1]^{\Delta}, N_{\Delta,\infty}(\mathbf{v}) \quad = \quad \max_{i \in \{1,\dots,\Delta\}} \mathbf{v}(i)$$

Here below are the answers of `question 1` for the zone corresponding to Europe (table 2), France (table 3), Greenland (table 4), France and Greenland (table 8), and Europe and Greenland (table 6) using the mapping $N_{\Delta,\infty}$.

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**
1	0.44	0.58	0.67	0.73	0.78	0.81	0.84
5	0.37	0.51	0.61	0.68	0.73	0.77	0.80
9	0.31	0.46	0.56	0.64	0.69	0.74	0.77
13	0.26	0.42	0.52	0.60	0.66	0.71	0.75
17	0.23	0.38	0.49	0.57	0.63	0.68	0.72
21	0.20	0.35	0.46	0.54	0.61	0.66	0.70
24	0.18	0.33	0.44	0.52	0.59	0.64	0.69
48	0.10	0.23	0.33	0.41	0.49	0.54	0.60
72	0.07	0.17	0.26	0.34	0.41	0.48	0.53
96	0.05	0.13	0.21	0.29	0.36	0.42	0.48
120	0.03	0.11	0.18	0.25	0.32	0.38	0.44
144	0.02	0.09	0.16	0.22	0.29	0.35	0.41
168	0.02	0.08	0.14	0.20	0.26	0.32	0.38

Table 2: Proportions of critical windows (using mapping $N_{\Delta,\infty}$) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of a wind signal in Europe over the 17 years using MERRA-2 database.

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**
1	0.44	0.59	0.69	0.76	0.81	0.85	0.88
5	0.36	0.51	0.62	0.70	0.76	0.81	0.84
9	0.30	0.45	0.57	0.66	0.72	0.77	0.81
13	0.25	0.40	0.52	0.62	0.69	0.74	0.79
17	0.21	0.36	0.49	0.58	0.66	0.71	0.76
21	0.18	0.33	0.45	0.55	0.63	0.69	0.74
24	0.16	0.31	0.43	0.53	0.61	0.67	0.73
48	0.09	0.19	0.30	0.40	0.49	0.56	0.63
72	0.06	0.13	0.22	0.32	0.41	0.48	0.56
96	0.04	0.10	0.17	0.26	0.34	0.42	0.50
120	0.03	0.08	0.14	0.21	0.30	0.37	0.45
144	0.03	0.06	0.11	0.18	0.26	0.33	0.41
168	0.02	0.05	0.10	0.16	0.23	0.30	0.37

Table 3: Proportions of critical windows (using mapping $N_{\Delta,\infty}$) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of a wind signal in France over the 17 years using MERRA-2 database.

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**
1	0.30	0.40	0.46	0.52	0.56	0.60	0.64
5	0.23	0.32	0.39	0.44	0.49	0.53	0.57
9	0.18	0.26	0.33	0.38	0.43	0.47	0.51
13	0.14	0.22	0.28	0.34	0.38	0.43	0.47
17	0.12	0.19	0.25	0.30	0.34	0.39	0.43
21	0.10	0.16	0.22	0.27	0.31	0.35	0.39
24	0.08	0.15	0.20	0.24	0.29	0.33	0.37
48	0.03	0.07	0.11	0.14	0.17	0.21	0.24
72	0.02	0.04	0.06	0.09	0.12	0.14	0.17
96	0.01	0.02	0.04	0.06	0.08	0.10	0.13
120	0.01	0.02	0.03	0.04	0.06	0.08	0.10
144	0.00	0.01	0.02	0.03	0.05	0.06	0.08
168	0.00	0.01	0.02	0.02	0.03	0.05	0.06

Table 4: Proportions of critical windows (using mapping $N_{\Delta,\infty}$) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of a wind signal in Greenland over the 17 years using MERRA-2 database.

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**
1	0.31	0.40	0.46	0.52	0.56	0.60	0.63
5	0.22	0.31	0.38	0.44	0.48	0.52	0.56
9	0.17	0.25	0.32	0.37	0.42	0.47	0.51
13	0.13	0.21	0.27	0.32	0.37	0.42	0.46
17	0.10	0.17	0.23	0.28	0.33	0.37	0.42
21	0.08	0.15	0.20	0.25	0.30	0.34	0.38
24	0.07	0.13	0.18	0.23	0.27	0.31	0.36
48	0.02	0.06	0.09	0.12	0.16	0.19	0.23
72	0.01	0.03	0.05	0.07	0.10	0.13	0.15
96	0.01	0.02	0.03	0.05	0.07	0.09	0.11
120	0.00	0.01	0.02	0.03	0.05	0.06	0.08
144	0.00	0.01	0.01	0.02	0.04	0.05	0.06
168	0.00	0.00	0.01	0.02	0.03	0.04	0.05

Table 5: Proportions of critical windows (using mapping $N_{\Delta,\infty}$) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of a wind signal in Greenland over the 17 years using MAR database.

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**
1	0.44	0.58	0.66	0.72	0.77	0.80	0.83
5	0.36	0.51	0.60	0.67	0.72	0.76	0.80
9	0.30	0.46	0.55	0.63	0.68	0.73	0.77
13	0.26	0.41	0.51	0.59	0.65	0.70	0.74
17	0.23	0.38	0.48	0.56	0.62	0.67	0.72
21	0.20	0.35	0.45	0.53	0.60	0.65	0.69
24	0.18	0.33	0.43	0.51	0.58	0.63	0.68
48	0.10	0.22	0.32	0.40	0.48	0.53	0.59
72	0.06	0.17	0.25	0.33	0.40	0.46	0.52
96	0.04	0.13	0.21	0.28	0.35	0.41	0.47
120	0.03	0.11	0.18	0.25	0.31	0.37	0.43
144	0.02	0.09	0.15	0.22	0.28	0.34	0.40
168	0.02	0.07	0.14	0.20	0.26	0.31	0.37

Table 6: Proportions of critical windows (using mapping $N_{\Delta,\infty}$) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of a wind signal in Europe and Greenland over the 17 years using MERRA-2 database.

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**
1	0.44	0.58	0.66	0.72	0.77	0.80	0.83
5	0.36	0.51	0.60	0.67	0.72	0.76	0.80
9	0.30	0.46	0.55	0.63	0.68	0.73	0.77
13	0.26	0.41	0.51	0.59	0.65	0.70	0.74
17	0.23	0.38	0.48	0.56	0.62	0.67	0.72
21	0.20	0.35	0.45	0.53	0.60	0.65	0.69
24	0.18	0.33	0.43	0.51	0.58	0.63	0.68
48	0.10	0.22	0.32	0.40	0.47	0.53	0.59
72	0.06	0.17	0.25	0.33	0.40	0.46	0.52
96	0.04	0.13	0.21	0.28	0.35	0.41	0.47
120	0.03	0.11	0.18	0.25	0.31	0.37	0.43
144	0.02	0.09	0.15	0.22	0.28	0.34	0.40
168	0.02	0.07	0.13	0.20	0.26	0.31	0.37

Table 7: Proportions of critical windows (using mapping $N_{\Delta,\infty}$) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of a wind signal in Europe and Greenland over the 17 years using MERRA-2 database for Europe and MAR database for Greenland.

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**
1	0.40	0.54	0.63	0.69	0.74	0.78	0.81
5	0.32	0.46	0.56	0.63	0.69	0.73	0.77
9	0.26	0.40	0.50	0.58	0.64	0.69	0.73
13	0.22	0.35	0.46	0.54	0.60	0.65	0.70
17	0.18	0.31	0.42	0.50	0.57	0.62	0.67
21	0.16	0.28	0.39	0.47	0.54	0.59	0.64
24	0.14	0.26	0.36	0.45	0.52	0.58	0.63
48	0.07	0.16	0.25	0.33	0.40	0.46	0.52
72	0.05	0.11	0.18	0.25	0.33	0.39	0.45
96	0.03	0.08	0.14	0.20	0.27	0.33	0.39
120	0.02	0.06	0.11	0.17	0.23	0.29	0.35
144	0.02	0.05	0.09	0.14	0.20	0.26	0.31
168	0.02	0.04	0.07	0.12	0.17	0.23	0.28

Table 8: Proportions of critical windows (using mapping $N_{\Delta,\infty}$) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of a wind signal in France and Greenland over the 17 years using MERRA-2 database.

Δ \ α	10%	20%	30%	40%	50%	60%	70%
1	0.41	0.54	0.63	0.69	0.74	0.78	0.81
5	0.32	0.46	0.56	0.63	0.68	0.73	0.76
9	0.26	0.40	0.50	0.58	0.64	0.69	0.73
13	0.21	0.35	0.45	0.54	0.60	0.65	0.69
17	0.18	0.31	0.41	0.50	0.57	0.62	0.67
21	0.15	0.28	0.38	0.47	0.54	0.59	0.64
24	0.14	0.26	0.36	0.45	0.52	0.57	0.62
48	0.07	0.15	0.24	0.32	0.40	0.46	0.52
72	0.04	0.10	0.17	0.25	0.32	0.38	0.44
96	0.03	0.08	0.13	0.20	0.27	0.33	0.39
120	0.02	0.06	0.11	0.16	0.23	0.29	0.35
144	0.02	0.05	0.09	0.14	0.20	0.25	0.31
168	0.02	0.04	0.07	0.12	0.17	0.23	0.28

Table 9: Proportions of critical windows (using mapping $N_{\Delta,\infty}$) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of a wind signal in France and Greenland over the 17 years using MERRA-2 database for France and MAR database for Greenland.

4.6.3 Observations

On the tables from table 2 to table 9, we can observe a same pattern. The proportion of critical windows increases (resp. decreases) when the value of the threshold α increases (resp. decreases). This is expected since there are more power values below the threshold if this threshold increases. The figure 15 illustrates this situation.

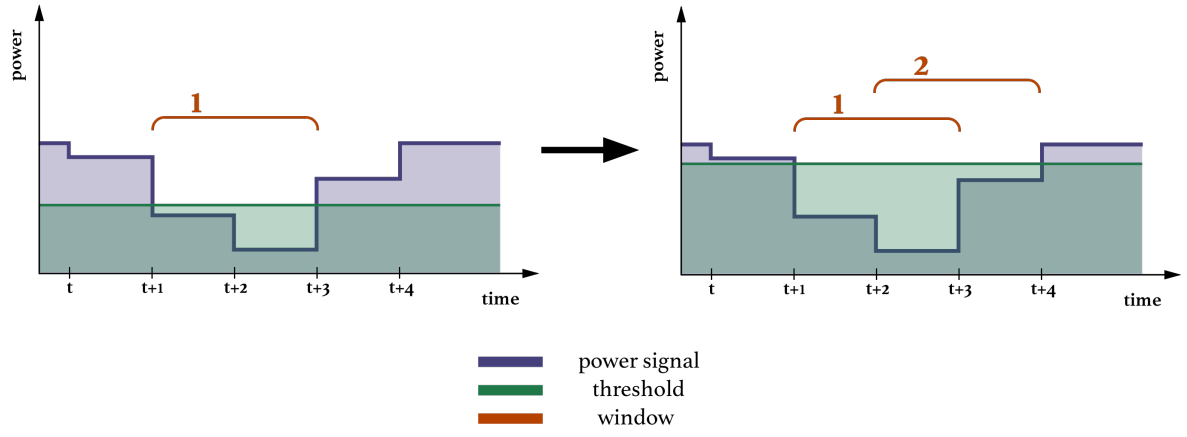


Figure 15: Situation when more critical windows appear when the threshold increases.

However, the number of critical windows decreases (resp.increases) when the length of critical windows increases. Indeed, the bigger the length of critical windows are, the more consecutive power values from the power signal should be below the threshold. And since the signal fluctuates around the threshold, small critical windows are more frequent than big critical windows. The Figure 16 illustrates the situation.

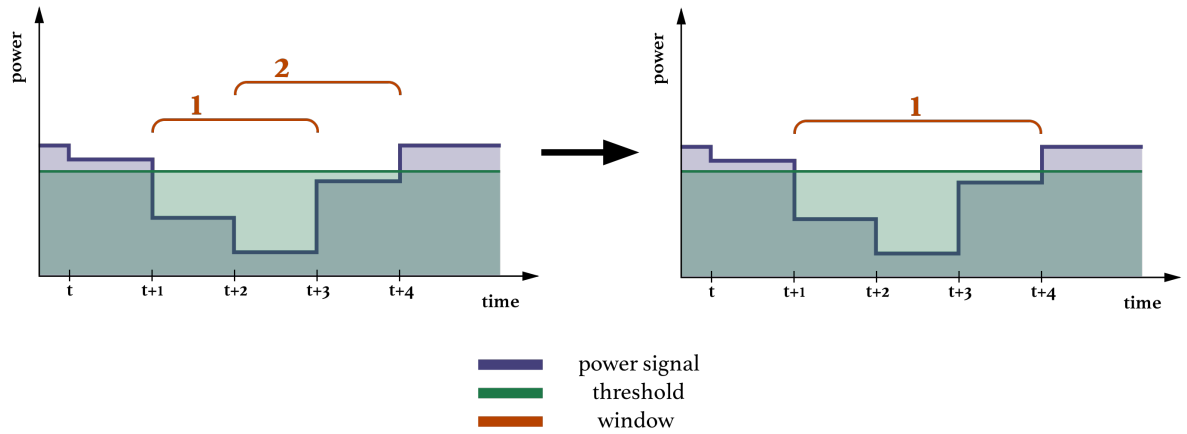


Figure 16: Situation when there are more small critical windows than big critical windows for a given treshold.

Between MERRA-2 and MAR, the tables showing simulation results with MERRA-2 data and MAR are equal or with a small difference that does not exceed 0.02.

Between Europe and France, proportions are almost the same in tables 2 and 3. From tables of France, proportions are very similar to proportions in tables of Europe. Therefore, analyzing proportions of critical windows on wind signal in France is representative of analyzing proportions in Europe.

Between Europe (resp. France) and Greenland, we observe that, for a given window length Δ and a given threshold α, there are a lot of more critical windows for wind signals in Europe (resp. France) than in Greenland 4.

When considering Europe and Greenland together, we see that the values in the tables 6 and 7 are the same as the tables for Europe alone. This is due to the Europe zone that is bigger than the Greenland zone.

When considering Europe and France together, proportions in the tables 8 and 9 are smaller than proportions in the tables of France alone. Indeed, proportions in Greenland tables are lower than proportions in France tables. Therefore, when considering both zone together, the proportions will decrease.

4.6.4 Drawback

The method used to compute critical windows with the mapping $N_{\Delta,\infty}$ gives a first idea of which zones has more critical windows than the others. However, this techniques is limited since it classify a window by only considering one value of the power signal. Indeed, for a wind

signal s_t^l, a window w_t^Δ is classified according to the biggest value of $\mathbf{u}^l\left(w_t^\Delta\right)$ (see section 3.4). The other values of $\mathbf{u}^l\left(w_t^\Delta\right)$ are weightless in the classification.

This mapping may be misleading in certain situations. For example, let us consider a power signal and a window. The power values during this window are all below the threshold, excepted one value. Therefore, with the mapping $N_{\Delta,\infty}$, this window will not be classified as critical. So, using this mapping, the proportion showed in the previous tables give an underestimated idea of the number of power values of the power signals that are below the threshold.

To solve this, a second mapping $N_{\Delta,1}$ has been introduced in section 3.5 that takes into account more values of the power signal to classify a window. In the second part of this section, same computations are done with the second mapping.

4.6.5 Second methodology

The second methodology is very similar to the first one. It will answer the `question 1` by using the indicator $I^{(1)}_{\Delta,\alpha}$ defined in section 3.9. Same values for Δ and α will be considered. For the next computations, the following second mapping $N_{\Delta,1}$ will be used:

$$\forall \mathbf{v} \in [0,1]^\Delta, N_{\Delta,1}(\mathbf{v}) \quad = \quad \frac{1}{\Delta}\sum_{i=1}^{\Delta}\mathbf{v}(i)$$

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**
1	0.44	0.58	0.67	0.73	0.78	0.81	0.84
5	0.44	0.58	0.67	0.73	0.78	0.81	0.85
9	0.43	0.58	0.67	0.73	0.78	0.82	0.85
13	0.42	0.57	0.66	0.73	0.78	0.82	0.86
17	0.41	0.56	0.66	0.73	0.78	0.82	0.86
21	0.40	0.56	0.66	0.73	0.78	0.83	0.87
24	0.39	0.55	0.66	0.73	0.78	0.83	0.87
48	0.35	0.53	0.64	0.73	0.79	0.85	0.90
72	0.32	0.51	0.64	0.73	0.80	0.87	0.92
96	0.30	0.50	0.63	0.73	0.81	0.88	0.93
120	0.29	0.49	0.63	0.74	0.82	0.89	0.94
144	0.28	0.48	0.63	0.74	0.83	0.89	0.95
168	0.27	0.48	0.63	0.74	0.83	0.90	0.96

Table 10: Average proportions of critical windows (using mapping $N_{\Delta,1}$) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of a wind signal in Europe over 17 years using MERRA-2 database.

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**
1	0.44	0.60	0.69	0.76	0.81	0.85	0.88
5	0.44	0.59	0.69	0.76	0.81	0.85	0.88
9	0.43	0.59	0.69	0.76	0.82	0.86	0.89
13	0.42	0.58	0.69	0.77	0.82	0.86	0.90
17	0.40	0.58	0.69	0.77	0.82	0.87	0.90
21	0.39	0.57	0.69	0.77	0.83	0.87	0.91
24	0.39	0.57	0.69	0.77	0.83	0.87	0.91
48	0.34	0.55	0.68	0.78	0.84	0.90	0.94
72	0.30	0.53	0.68	0.78	0.86	0.91	0.95
96	0.28	0.52	0.68	0.79	0.87	0.93	0.96
120	0.26	0.51	0.68	0.80	0.88	0.94	0.97
144	0.24	0.50	0.68	0.81	0.89	0.95	0.98
168	0.23	0.49	0.68	0.81	0.90	0.95	0.99

Table 11: Average proportions of critical windows (using mapping $N_{\Delta,1}$) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of a wind signal in France over 17 years using MERRA-2 database.

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**
1	0.31	0.40	0.46	0.52	0.56	0.60	0.64
5	0.29	0.39	0.46	0.51	0.56	0.60	0.64
9	0.27	0.37	0.44	0.50	0.56	0.61	0.66
13	0.25	0.35	0.43	0.50	0.56	0.61	0.67
17	0.23	0.33	0.42	0.49	0.55	0.62	0.68
21	0.21	0.32	0.40	0.48	0.55	0.62	0.69
24	0.20	0.31	0.39	0.47	0.55	0.63	0.70
48	0.14	0.24	0.34	0.44	0.55	0.65	0.75
72	0.11	0.20	0.31	0.42	0.55	0.67	0.79
96	0.09	0.18	0.29	0.41	0.55	0.69	0.81
120	0.07	0.16	0.27	0.40	0.55	0.70	0.83
144	0.06	0.15	0.26	0.40	0.55	0.71	0.85
168	0.05	0.14	0.25	0.39	0.55	0.72	0.86

Table 12: Average proportions of critical windows (using mapping $N_{\Delta,1}$) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of a wind signal in Greenland over 17 years using MERRA-2 database.

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**
1	0.31	0.40	0.46	0.52	0.56	0.60	0.63
5	0.29	0.39	0.46	0.51	0.56	0.60	0.64
9	0.26	0.37	0.44	0.50	0.56	0.61	0.66
13	0.24	0.35	0.43	0.50	0.56	0.62	0.67
17	0.22	0.33	0.41	0.49	0.56	0.62	0.68
21	0.20	0.31	0.40	0.48	0.56	0.63	0.70
24	0.19	0.30	0.39	0.48	0.55	0.63	0.71
48	0.13	0.23	0.34	0.44	0.55	0.66	0.76
72	0.10	0.19	0.30	0.42	0.55	0.68	0.80
96	0.08	0.17	0.28	0.40	0.55	0.70	0.83
120	0.06	0.15	0.26	0.39	0.55	0.71	0.85
144	0.05	0.14	0.25	0.39	0.55	0.72	0.86
168	0.05	0.13	0.24	0.38	0.55	0.73	0.88

Table 13: Average proportions of critical windows (using mapping $N_{\Delta,1}$) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of a wind signal in Greenland over 17 years using MAR database.

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**
1	0.44	0.58	0.66	0.72	0.77	0.80	0.83
5	0.43	0.57	0.66	0.72	0.77	0.81	0.84
9	0.42	0.57	0.66	0.72	0.77	0.81	0.84
13	0.41	0.56	0.66	0.72	0.77	0.81	0.85
17	0.40	0.56	0.65	0.72	0.77	0.82	0.86
21	0.39	0.55	0.65	0.72	0.78	0.82	0.86
24	0.39	0.55	0.65	0.72	0.78	0.83	0.87
48	0.34	0.52	0.63	0.72	0.79	0.84	0.89
72	0.31	0.50	0.63	0.72	0.80	0.86	0.91
96	0.30	0.49	0.62	0.72	0.80	0.87	0.93
120	0.28	0.48	0.62	0.73	0.81	0.88	0.94
144	0.27	0.47	0.62	0.73	0.82	0.89	0.95
168	0.26	0.47	0.62	0.73	0.82	0.90	0.95

Table 14: Average proportions of critical windows (using mapping $N_{\Delta,1}$) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of a wind signal in Europe and Greenland over 17 years using MERRA-2 database.

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**
1	0.44	0.58	0.66	0.72	0.77	0.80	0.83
5	0.43	0.57	0.66	0.72	0.77	0.81	0.84
9	0.42	0.57	0.66	0.72	0.77	0.81	0.84
13	0.41	0.56	0.66	0.72	0.77	0.81	0.85
17	0.40	0.56	0.65	0.72	0.77	0.82	0.86
21	0.39	0.55	0.65	0.72	0.78	0.82	0.86
24	0.39	0.55	0.65	0.72	0.78	0.83	0.87
48	0.34	0.52	0.63	0.72	0.79	0.84	0.89
72	0.31	0.50	0.63	0.72	0.80	0.86	0.91
96	0.29	0.49	0.62	0.72	0.80	0.87	0.93
120	0.28	0.48	0.62	0.73	0.81	0.88	0.94
144	0.27	0.47	0.62	0.73	0.82	0.89	0.95
168	0.26	0.47	0.62	0.73	0.82	0.90	0.95

Table 15: Average proportions of critical windows (using mapping $N_{\Delta,1}$) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of a wind signal in Europe and Greenland over 17 years using MERRA-2 database for Europe and MAR database for Greenland.

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**
1	0.41	0.54	0.63	0.69	0.74	0.78	0.81
5	0.40	0.53	0.63	0.69	0.74	0.78	0.82
9	0.38	0.53	0.62	0.69	0.74	0.79	0.82
13	0.37	0.52	0.62	0.69	0.75	0.79	0.83
17	0.35	0.51	0.61	0.69	0.75	0.80	0.84
21	0.34	0.50	0.61	0.69	0.75	0.80	0.85
24	0.33	0.50	0.61	0.69	0.75	0.81	0.85
48	0.28	0.46	0.59	0.68	0.76	0.83	0.88
72	0.25	0.44	0.58	0.68	0.77	0.84	0.91
96	0.22	0.42	0.57	0.69	0.78	0.86	0.92
120	0.21	0.41	0.57	0.69	0.79	0.87	0.93
144	0.19	0.40	0.56	0.69	0.80	0.88	0.95
168	0.18	0.39	0.56	0.69	0.80	0.89	0.95

Table 16: Average proportions of critical windows (using mapping $N_{\Delta,1}$) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of a wind signal in France and Greenland over 17 years using MERRA-2 database.

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**
1	0.41	0.54	0.63	0.69	0.74	0.78	0.81
5	0.40	0.53	0.63	0.69	0.74	0.78	0.82
9	0.38	0.53	0.62	0.69	0.74	0.79	0.82
13	0.37	0.52	0.62	0.69	0.75	0.79	0.83
17	0.35	0.51	0.61	0.69	0.75	0.80	0.84
21	0.34	0.50	0.61	0.69	0.75	0.80	0.85
24	0.33	0.50	0.60	0.69	0.75	0.81	0.85
48	0.28	0.46	0.59	0.68	0.76	0.83	0.89
72	0.24	0.44	0.57	0.68	0.77	0.85	0.91
96	0.22	0.42	0.57	0.68	0.78	0.86	0.93
120	0.20	0.41	0.56	0.69	0.79	0.87	0.94
144	0.19	0.40	0.56	0.69	0.80	0.88	0.95
168	0.18	0.39	0.56	0.69	0.80	0.89	0.96

Table 17: Average proportions of critical windows (using mapping $N_{\Delta,1}$) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of a wind signal in France and Greenland over 17 years using MERRA-2 database for France and MAR database for Greenland.

4.6.6 Observations

First, we observe the same pattern as the tables in previous methodology at subsection 4.6.2. The proportion of critical windows increases (resp. decreases) when the threshold increases (resp. decreases) and the proportion decreases when the length of the windows increases. This is due to the same reasons as previous explanations illustrated by the figures 15 and 16.

Secondly, when comparing the tables illustrating mapping $N_{\Delta,1}$ and the tables obtained with mapping $N_{\Delta,\infty}$, we observe that proportions are the same for window length $\Delta = 1$. This is an expected situation.
Indeed, for a window length $\Delta = 1$, a power signal u_t^l, and a window w_t^Δ, the associated vector $\mathbf{u}^l\left(w_t^\Delta\right)$ contains only one value and therefore

$$N_{\Delta,1}\left(\mathbf{u}^l\left(w_t^\Delta\right)\right) = N_{\Delta,\infty}\left(\mathbf{u}^l\left(w_t^\Delta\right)\right)$$

However, for the other proportions, we observe that values from the tables illustrating mapping $N_{\Delta,1}$ are strictly bigger than values obtained with mapping $N_{\Delta,\infty}$. Indeed, for a power signal u_t^l, a window w_t^Δ, and a threshold α, if the vector $\mathbf{u}^l\left(w_t^\Delta\right)$ contains a power value x such that $x > \alpha$, the mapping $N_{\Delta,\infty}$ leads to classify the window w_t^l as non critical. However, for mapping $N_{\Delta,1}$, the classification of this window also depends on the other power values from $\mathbf{u}^l\left(w_t^\Delta\right)$. If $mean\left(\mathbf{u}^l\left(w_t^\Delta\right)\right) > \alpha$, the window is classified as non critical. But,

if $mean\left(\mathbf{u}^l\left(w_t^\Delta\right)\right) > \alpha$ this window is classified as critical. This is why, when using the second mapping $N_{\Delta,1}$, there is more windows classified as critical. The figure 17 illustrates this situation.

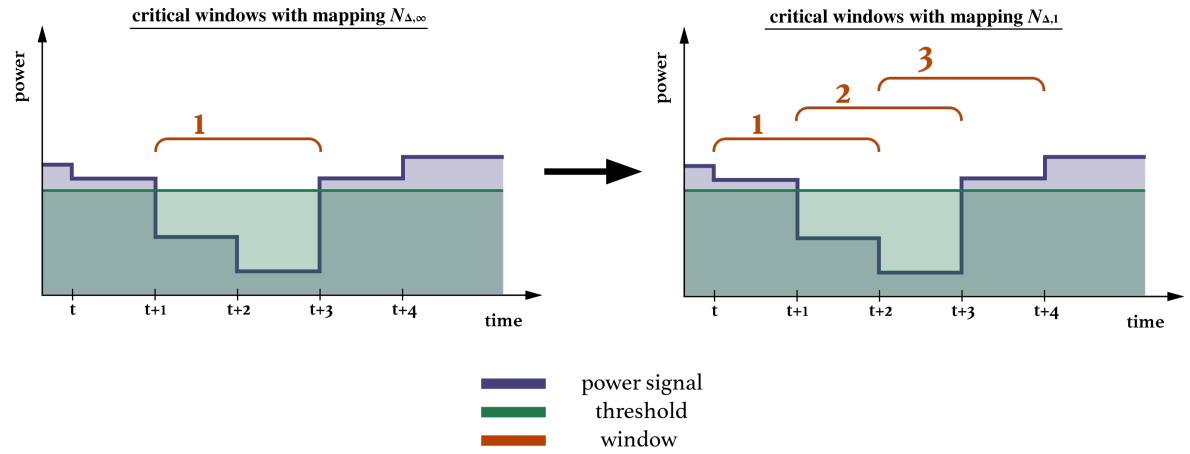


Figure 17: Situation when there are more critical windows with mapping $N_{\Delta,1}$ than mapping $N_{\Delta,\infty}$.

Also, all the observations for the tables generated with the first mapping can also be applied to the tables generated with the second mapping.

4.7 Question 2: average proportion of critical windows common to all signals in a zone

4.7.1 Introduction

This section will try to answer a second interesting question, which is the following one:

Question 2

In a specific zone, what is the average proportion of critical windows over the 17 years common to all wind signals found in this zone ?

4.7.2 Methodology

To answer the `question 2`, we will use the second indicator $I^{(2)}_{\Delta,\alpha}$ introduced in subsection 3.10. This indicator allow us to answer the `question 2` directly.

Like the `question 1`, the indicator $I^{(2)}_{\Delta,\alpha}$ depends on the length Δ and the threshold α of the considered windows. Hence, to observe the influence of the two variables, we will run the same computations by changing both variables for each computation. The length Δ will take the successive values: $1, 4, 9, 13, 17$, then 21 hours while the threshold will take the successive values: $30\%, 40\%, 50\%, 60\%$, and finally 70%.

Here below are the tables showing the proportions of critical windows with length Δ and threshold α, considering wind signal in Europe, France, Greenland, then Europe and Greenland, and finally France and Greenland. Computations are done with both mapping $N_{\Delta,\infty}$ and $N_{\Delta,1}$.

Table 18.a. Proportions of common critical windows in Europe (mapping $N_{\Delta,\infty}$ with MERRA-2).

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**	**100%**
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00

Table 18.b. Proportions of common critical windows in Europe (mapping $N_{\Delta,1}$ with MERRA-2).

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**	**100%**
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	1.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	1.00
13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	1.00
17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.03	1.00
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.04	1.00

Table 18: Average proportions of common critical windows (using mapping $N_{\Delta,\infty}$ on table 18.a and mapping $N_{\Delta,1}$ on table 18.b) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of all wind signals in Europe over 17 years using MERRA-2 database.

Table 19.a. Proportions of common critical windows in France (mapping $N_{\Delta,\infty}$ with MERRA-2).

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**	**100%**
1	0.00	0.00	0.01	0.03	0.07	0.11	0.16	0.21	0.30	1.00
5	0.00	0.00	0.01	0.02	0.04	0.07	0.11	0.16	0.24	1.00
9	0.00	0.00	0.00	0.01	0.03	0.05	0.08	0.13	0.20	1.00
13	0.00	0.00	0.00	0.01	0.02	0.04	0.07	0.11	0.17	1.00
17	0.00	0.00	0.00	0.00	0.01	0.03	0.05	0.09	0.15	1.00
21	0.00	0.00	0.00	0.00	0.01	0.02	0.04	0.07	0.13	1.00

Table 19.b. Proportions of common critical windows in France (mapping $N_{\Delta,1}$ with MERRA-2).

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**	**100%**
1	0.00	0.00	0.01	0.03	0.07	0.11	0.16	0.21	0.30	1.00
5	0.00	0.00	0.02	0.04	0.08	0.12	0.18	0.24	0.33	1.00
9	0.00	0.00	0.02	0.05	0.09	0.14	0.20	0.27	0.37	1.00
13	0.00	0.00	0.02	0.05	0.10	0.15	0.22	0.30	0.41	1.00
17	0.00	0.01	0.02	0.06	0.11	0.17	0.24	0.33	0.44	1.00
21	0.00	0.01	0.03	0.06	0.12	0.18	0.26	0.35	0.48	1.00

Table 19: Average proportions of common critical windows (using mapping $N_{\Delta,\infty}$ on table 19.a and mapping $N_{\Delta,1}$ on table 19.b) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of all wind signals in France over 17 years using MERRA-2 database.

Table 20.a. Proportions of common critical windows in Greenland (mapping $N_{\Delta,\infty}$ with MERRA-2).

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**	**100%**
1	0.00	0.00	0.01	0.02	0.03	0.05	0.07	0.09	0.13	1.00
5	0.00	0.00	0.01	0.01	0.02	0.03	0.05	0.06	0.10	1.00
9	0.00	0.00	0.00	0.01	0.01	0.02	0.03	0.05	0.07	1.00
13	0.00	0.00	0.00	0.01	0.01	0.02	0.02	0.04	0.06	1.00
17	0.00	0.00	0.00	0.00	0.01	0.01	0.02	0.03	0.05	1.00
21	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.02	0.04	1.00

Table 20.b. Proportions of common critical windows in Greenland (mapping $N_{\Delta,1}$ with MERRA-2).

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**	**100%**
1	0.00	0.00	0.01	0.02	0.03	0.05	0.07	0.09	0.13	1.00
5	0.00	0.00	0.01	0.02	0.03	0.05	0.07	0.10	0.14	1.00
9	0.00	0.00	0.01	0.02	0.04	0.06	0.08	0.11	0.17	1.00
13	0.00	0.00	0.01	0.02	0.04	0.06	0.09	0.13	0.20	1.00
17	0.00	0.00	0.01	0.02	0.04	0.07	0.10	0.15	0.23	1.00
21	0.00	0.00	0.01	0.02	0.04	0.07	0.11	0.17	0.26	1.00

Table 20: Average proportions of common critical windows (using mapping $N_{\Delta,\infty}$ on table 20.a and mapping $N_{\Delta,1}$ on table 20.b) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of all wind signals in Greenland over 17 years using MERRA-2 database.

Table 21.a. Proportions of common critical windows in Greenland (mapping $N_{\Delta,\infty}$ with MAR).

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**	**100%**
1	0.00	0.00	0.00	0.01	0.02	0.03	0.04	0.06	0.10	1.00
5	0.00	0.00	0.00	0.00	0.01	0.02	0.03	0.04	0.07	1.00
9	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.03	0.05	1.00
13	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.02	0.04	1.00
17	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.03	1.00
21	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.02	1.00

Table 21.b. Proportions of common critical windows in Greenland (mapping $N_{\Delta,1}$ with MAR).

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**	**100%**
1	0.00	0.00	0.00	0.01	0.02	0.03	0.04	0.06	0.10	1.00
5	0.00	0.00	0.00	0.01	0.02	0.03	0.05	0.07	0.11	1.00
9	0.00	0.00	0.00	0.01	0.02	0.04	0.06	0.09	0.13	1.00
13	0.00	0.00	0.01	0.01	0.02	0.04	0.07	0.10	0.16	1.00
17	0.00	0.00	0.01	0.01	0.03	0.05	0.08	0.12	0.19	1.00
21	0.00	0.00	0.01	0.01	0.03	0.05	0.09	0.14	0.22	1.00

Table 21: Average proportions of common critical windows (using mapping $N_{\Delta,\infty}$ on table 21.a and mapping $N_{\Delta,1}$ on table 21.b) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of all wind signals in Greenland over 17 years using MAR database.

Table 22.a. Proportions of common critical windows in Europe and Greenland (mapping $N_{\Delta,\infty}$ with MERRA-2).

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**	**100%**
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00

Table 22.b. Proportions of common critical windows in Europe and Greenland (mapping $N_{\Delta,1}$ with MERRA-2).

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**	**100%**
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	1.00
17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	1.00
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	1.00

Table 22: Average proportions of common critical windows (using mapping $N_{\Delta,\infty}$ on table 22.a and mapping $N_{\Delta,1}$ on table 22.b) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of all wind signals in Europe and Greenland over 17 years using MERRA-2 database.

Table 23.a. Proportions of common critical windows in Europe and Greenland (mapping $N_{\Delta,\infty}$ with MERRA-2 and MAR).

Δ \ α	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00

Table 23.b. Proportions of common critical windows in Europe and Greenland (mapping $N_{\Delta,1}$ with MERRA-2 and MAR).

Δ \ α	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	1.00
17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	1.00
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	1.00

Table 23: Average proportions of common critical windows (using mapping $N_{\Delta,\infty}$ on table 23.a and mapping $N_{\Delta,1}$ on table 23.b) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of all wind signals in Europe and Greenland over 17 years using MERRA-2 database for Europe and MAR database for Greenland.

Table 24.a. Proportions of common critical windows in France and Greenland (mapping $N_{\Delta,\infty}$ with MERRA-2).

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**	**100%**
1	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.02	0.05	1.00
5	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.03	1.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02	1.00
13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	1.00
17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	1.00
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	1.00

Table 24.b. Proportions of common critical windows in France and Greenland (mapping $N_{\Delta,1}$ with MERRA-2).

Δ \ α	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**	**100%**
1	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.02	0.05	1.00
5	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.03	0.06	1.00
9	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.04	0.08	1.00
13	0.00	0.00	0.00	0.00	0.01	0.01	0.03	0.05	0.10	1.00
17	0.00	0.00	0.00	0.00	0.01	0.02	0.03	0.06	0.12	1.00
21	0.00	0.00	0.00	0.00	0.01	0.02	0.04	0.07	0.14	1.00

Table 24: Average proportions of common critical windows (using mapping $N_{\Delta,\infty}$ on table 24.a and mapping $N_{\Delta,1}$ on table 24.b) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of all wind signals in France and Greenland over 17 years using MERRA-2 database.

Table 25.a. Proportions of common critical windows in France and Greenland (mapping $N_{\Delta,\infty}$ with MERRA-2 and MAR).

Δ \ α	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
1	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.04	1.00
5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02	1.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	1.00
13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	1.00
17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00

Table 25.b. Proportions of common critical windows in France and Greenland (mapping $N_{\Delta,1}$ with MERRA-2 and MAR).

Δ \ α	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
1	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.04	1.00
5	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.05	1.00
9	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.03	0.06	1.00
13	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.04	0.08	1.00
17	0.00	0.00	0.00	0.00	0.00	0.01	0.03	0.05	0.10	1.00
21	0.00	0.00	0.00	0.00	0.01	0.01	0.03	0.06	0.13	1.00

Table 25: Average proportions of common critical windows (using mapping $N_{\Delta,\infty}$ on table 25.a and mapping $N_{\Delta,1}$ on table 25.b) with length Δ (hours) and threshold α (percentage of 2MW, the installed power) of all wind signals in France and Greenland over 17 years using MERRA-2 database for France and MAR database for Greenland.

4.7.3 Observations

As explained for the observation of the second methodology of `question 1`, there is no difference between both mappings when considering $\Delta = 1$. However, when increasing Δ, proportions are becoming smaller with mapping $N_{\Delta,\infty}$ while they are becoming bigger with the other mapping.

Also, the proportions in column $\alpha = 100\%$ are always equal to 1. Which is expected since all the signals appear below the threshold of 100% the maximum power production allowed.

For Europe, all the proportions in the table generated with the first mapping are null (figure 18.a). Which means that, at any time, there always are some locations that can produce a certain power and therefore do not have critical windows at that time. This is why the proportions of

common critical windows (to all locations in the zone) is null. This is due to the zone that is very big (see 3) and includes a lot of different regions with different wind regimes.

For the Europe table generated with the second mapping (figure 18.b), same observations are made. However, some values appear for $\alpha = 80\%$ and $\alpha = 90\%$ but these proportions remain very small (smaller than 0.04).

Between Europe and France, since the France zone is small, wind regimes from this zone tend to be less various than from the Europe zone (figures 19). We can observe that some proportions appear from $\alpha = 30\%$.

Between Europe and Greenland, same observations can be done than previous point. However, proportions are smaller in Greenland tables than in France tables. So, there are less critical windows in Greenland than in France (figures 20 and 21). Section 4.6 also showed these results.

Considering Europe and Greenland together, since there are null proportions in Europe tables, there are also null proportions when considering Europe zone with another zone (figures 22 and 23).

Considering France and Greenland together, there are still some proportions that appear above $\alpha = 50\%$ for both mappings. However, those do not exceed 0.14, which is small (figures 24 and 25).

In general, we need to set a treshold above 30% to see some small proportions of common critical windows appear. Proportions are even smaller when considering big zone due to the various wind regimes that may occur in this zone.

4.8 Genetic programming

4.8.1 Introduction

In order to answer the next section, we need to develop an efficient methodology. Indeed, the next question requires to analyze all the wind signals from a zone (Europe or Greenland) and to find the n wind signals that maximize or minimize a given criteria. A naive way to answer such problem is to look for all the possible subsets of n wind signals within the set of all wind signals of the zone and to find the subset which is optimal for the given criterion.

Hence, for a zone containing in total N wind signals, the maximization and minimization problem will consider

$$\binom{N}{n} = \frac{N!}{n!(N-n)!}$$

possible subsets. However, if there are many wind signals in the zone (big N in the previous equation) or if we have to consider big subsets (big n in the previous equation), computations may be not efficient. Let us illustrate this inefficiency using the Europe zone.

In the Europe zone, we consider 4828 locations, which correspond to the following ones:

$$\forall l = (\lambda^{long}, \lambda^{lat}) \in \mathcal{L}_E : \lambda^{long} \in \{-10, -9.375, ..., 31.875\}, \lambda^{lat} \in \{36, 36.5, ..., 71\}$$

For example, in order to find among the 4828 wind signals a subset of size 10 that minimizes or maximizes a given criterion in an exhaustive way, the problem will review

$$\binom{4828}{10} = \frac{4828!}{10!(4828-10)!} = 1.8787 \times 10^{+30}$$

possible subsets, which corresponds to an unreasonable amount of computation time.

In order to avoid this huge amount of subset possibilities, we decided to approach this problem using heuristics, in particular Genetics Algorithm [6].

4.8.2 Principle of genetics algorithm

A genetics algorithm [6] (denoted as GA) is a heuristic algorithm that belongs to the larger class of evolutionary algorithm selection (denoted as EA). Genetics algorithms are based on the process of natural selection by relying on bio-inspired operators such as mutation, crossover and selection [7].

A genetics algorithm starts from a population of random candidates (also called individuals) as starting solution. Each candidate corresponds to a solution of the problem and is often represented as binary vectors. The purpose of genetics algorithms is to make the population evolves toward better solutions by merging, mutating or altering the current population.

By merging, mutating, and keeping best candidates to create a new population, this algorithm tends to lead to the best candidates possible after several population evolutions.

4.8.3 Algorithm

Genetics algorithms run following several key steps. The diagram 18 illustrates these key steps and their interactions.

Here below is explained the principles of the different Genetics algorithm steps.

1. Initialization. As input, genetics algorithm require an initial population (also called solution). This first population contains several hundreds of randomly generated initial individuals. An individual is characterized by a set of parameters known as Genes. Genes are joined into a string to form a Chromosome (the individual). Typically, a chromosome is encoded into a binary string (string of 1s and 0s).

2. Fitness assignment. In order to evaluate the individuals of a given population, a score is assigned to each individuals using a fitness function. The fitness function determines how fit is an individual. The probability that an individual will be selected for reproduction (as solution of the problem) is based on its fitness score.

3. Selection. The evolution of the population is based on several phases. The first one is based on the creation of a new population by selecting the fittest individuals from the previous population, which are called `parents`.

4. Crossover. The second phase of the evolution consists in creating new candidates called `children`. So, using the selected parents from the previous step, several parent genes are crossed over in order to create the children. The crossover technique may be randomly or wisely defined.

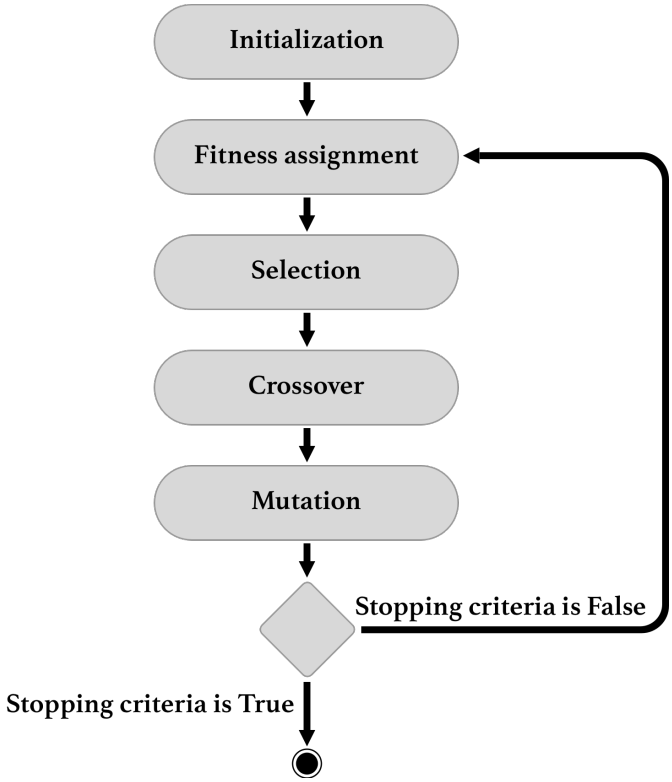


Figure 18: Diagram representing the key steps and their interactions of genetics algorithms.

5. Mutation. Since genetics algorithms tend to imitate natural evolution of a population, the algorithm procedure embeds a mutation phase. Indeed, when parents are kept from one population to the other, or when new candidates are created from parents, these individuals mutate with a small probability. This corresponds to slightly modify some of the new individuals at random time.

6. Stopping criteria. The algorithm terminates once the population has converged or when the number of iterations exceeds a given threshold. Then, the individuals contained into the final population constitute a set of solutions for the problem.

Other steps may be added to reinforce the algorithm such as heuristics or elitist selections.

4.8.4 Implementation

In order to use the genetics algorithm to solve the problem, the different steps described in subsection 4.8.3 need to be adjusted to our problem case. Hence, below is explained how the steps were adapted.

1. Initialization. We decided to start our implementation of the genetics algorithm with a starting population of 100 random individuals.

For our problem, an individual corresponds to a subset L of 10 locations from the original zone $\mathcal{L}$ (Europe, Greenland, France, etc). We need to convert the representation of an individual into a binary string in order to use it with the genetics algorithm. Thus, an individual is represented by a vector containing $|\mathcal{L}|$ binary values. Each index i of the vector is associated

to a precise location $l_i \in \mathcal{L}$. If the value from the vector at index i is `True` (or equals to 1), it means that location l_i belongs to the subset represented by the individual ($l_i \in L$). In total, the vector contains 10 $True$ values. The figure 19 illustrates an example of individual representing a subset containing locations $\{l_0, l_3, l_5, ...\}$.

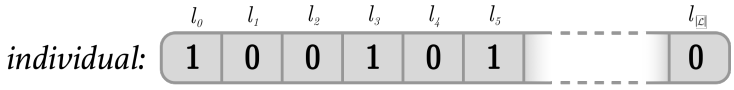


Figure 19: Example of individual representing a subset of locations $\{l_0, l_3, l_5, ...\}$ for the genetics algorithm.

Hence, in order to create 100 random individuals for the starting population, each individual contains 10 $True$ values at random indices.

2. Fitness assignment. To keep the best candidates from one population to another, we will define a `fitness function` that associates a score to each individual.

As explained in the initialization step above, each individual represents a subset L of 10 locations from a set $\mathcal{L}$. The score of an individual will be based on the number of critical windows of length Δ below a threshold α that are common to all locations contained in the subset represented by the individual. More formally, let L be a subset of locations from a set $\mathcal{L}$ of locations and Ind an individual that represents the subset L, the fitness function is defined as follows:

$$fitness(Ind) = \gamma_{\Delta,\alpha}(L)$$

with

- Δ: the length of the critical windows we want to compute (6 hours),
- α: the threshold of the critical windows we want to compute (50%),
- $\gamma_{\Delta,\alpha}()$: the function that computes the cardinality of the critical windows intersection (see section 3.8)

3. Selection. In the case of computing the minimum (resp. maximum) intersection, the 30 individuals with the smallest (resp. biggest) fitness score will be called parents as next population and will be kept for the next population.

For the other individuals that are not selected as parent, they may also be part of the next population with a probability of 0.05.

4. Crossover. Once individuals from the current population are selected as parents for the new population (because of their best fitness score), new individuals will be created using these parents and will also compose the new population. These children are created by crossing over several parents. For our implementation, we decided to considering four different parents for each child. Hence, a new individual is created by selected a total of 10 $True$ values from their four parents. The figure 20 illustrate the creation of a child by crossing over two parents.

These new children complete the new population of individuals.

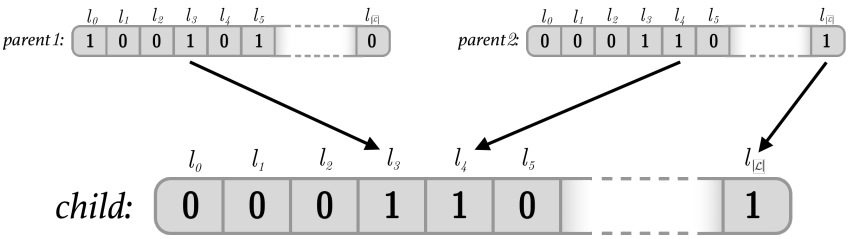


Figure 20: Example of individual created by cross over two parents.

5. Mutation. individuals can also mutate from one population to another with a probability of 0.01. The mutation of an individual consists in randomly permuting a $False$ value and a $True$ values from the individual.

The figure 21 shows the mutation of an individual. The set L of locations represented by the mutated individual now includes the location l_1 and does not include the location l_3 anymore.

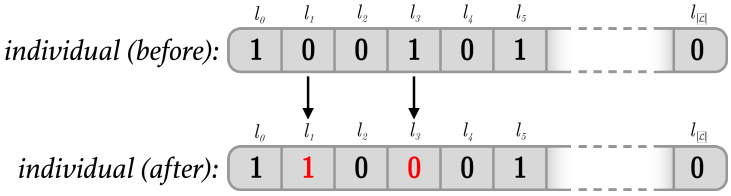


Figure 21: Example of mutation of an individual that permutes a $False$ value and a $True$ values randomly.

6. Stopping criteria. The genetics algorithm stops after 100 iterations of the population evolution. The 10 individuals with the best fitness score are then extracted from the final population.

Using these approaches, we are able to create a genetics algorithm that fits our needs. We call our custom genetics algorithm as follows:

$$\mathcal{A}_{ga}^{min,6,50}$$

when minimizing the common critical window intersection with $\delta = 6$ and $\alpha = 50\%$. Or

$$\mathcal{A}_{ga}^{max,6,50}$$

when maximizing the common critical window intersection with $\delta = 6$ and $\alpha = 50\%$.

4.9 Question 3: 10 points from a given zone such that the proportion of common critical windows is maximal

4.9.1 Introduction

This section will answer the following question:

Question 3

In a specific zone, what are the 10 points such that the average number of critical windows, common to all signals in the zone, is maximal?

The purpose of answering this question is to find the 10 locations for which there are a lot of common critical windows considering

$$\Delta = 6 \text{ and } \alpha = 50\%.$$

Finding these 10 locations will highlight a worst case scenario if wind turbine farms are installed at these 10 locations.

Indeed since these locations share critical windows, all wind turbine farms installed at these 10 locations will produce low power during these common critical windows.

4.9.2 First methodology

The first idea was to run an instance of the genetics algorithm maximizing the common critical window intersection (see $\mathcal{A}_{ga}^{max,6,50}$ algorithm at section 4.8.3) with an initial population of 100 random individuals. At the end of the 100 iterations of the algorithm, we will only select the 10 best candidates representing each a subset of 10 locations that should maximize the number of common critical windows.

While considering them together, these 10 best candidates will give 100 locations. To answer the `question 3` and to only select 10 locations from these 100 ones, we will plot these 100 locations on a graph and choose the 10 ones that are selected several times.

In order to print the results of the 10 best candidates and their 10 locations, we plotted all of the $10 \times 10 = 100$ locations as colored dots on the same map. the color chosen to plot the locations on the map is semi-transparent red. Locations that are selected by several best candidates will lead to dots that are painted several times at the same place on the map. Since dots are painted with a semi-transparent color, locations that are chosen several times will be represented as darker dots than locations that are chosen only once because of color overlay.

Here below are plotted the locations selected by the 10 best candidates computed with the $\mathcal{A}_{ga}^{max,6,50}$ algorithm on the Europe zone and on the Greenland zone (using MERRA-2 data and then MAR data) with both mappings.

4.9.2.1 Europe ($\mathcal{L} = \mathcal{L}_E$)

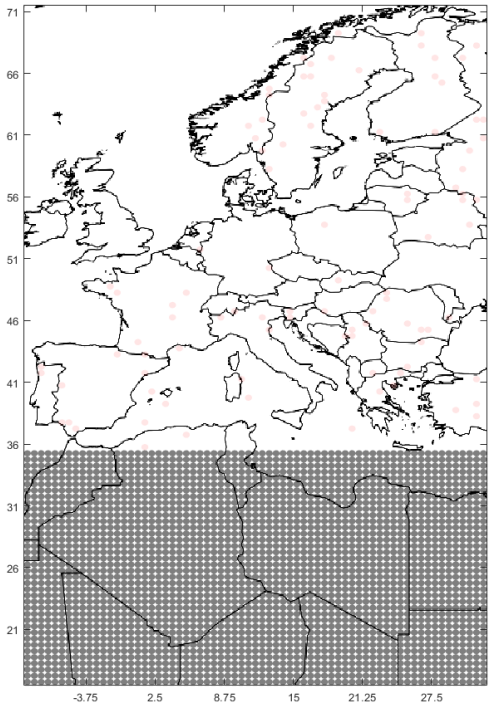


Figure 22.a: $N_{\Delta,\infty}$ on MERRA-2 data

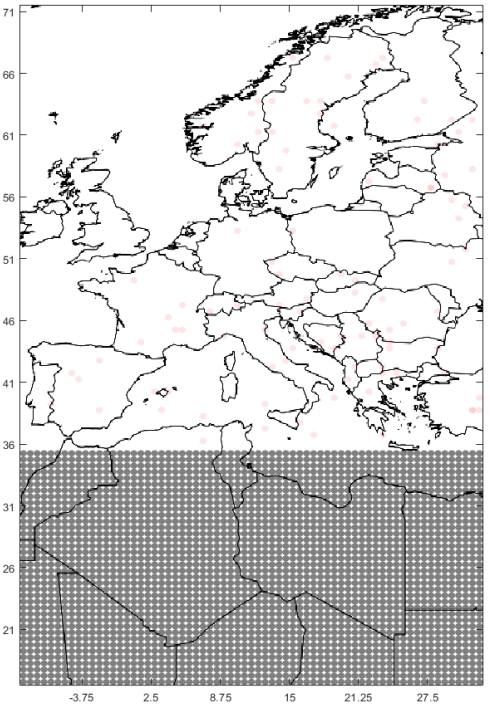


Figure 22.b: $N_{\Delta,1}$ on MERRA-2 data

Figure 22: The figures 22.a and 22.b show the results of the $\mathcal{A}_{ga}^{max,6,50}$ algorithm on Europe using the MERRA-2 database with respectively mappings $N_{\Delta,\infty}$ and $N_{\Delta,1}$.

4.9.2.2 Greenland ($\mathcal{L} = \mathcal{L}_G$)

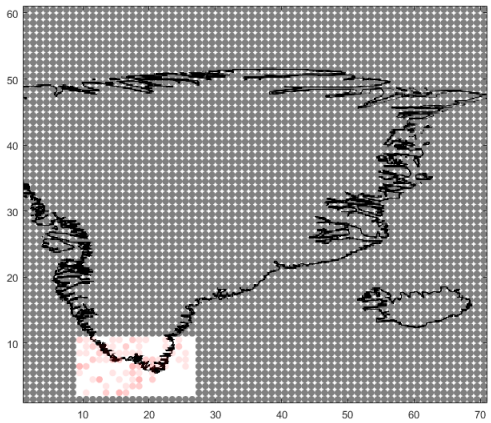


Figure 23.a: $N_{\Delta,\infty}$ on MERRA-2 data

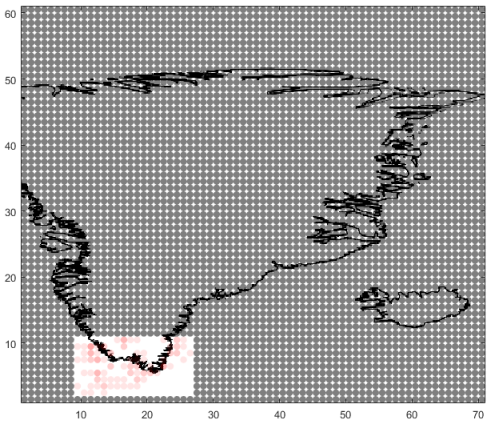


Figure 23.b: $N_{\Delta,1}$ on MERRA-2 data

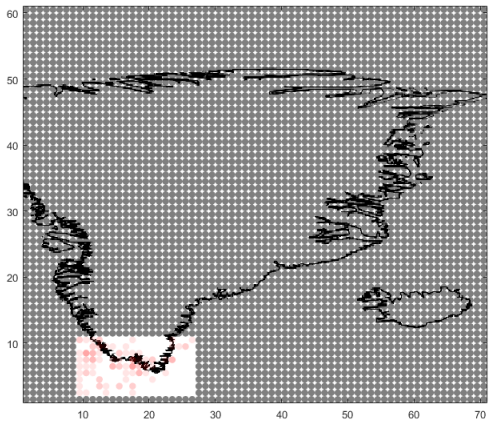


Figure 23.c: $N_{\Delta,\infty}$ on MAR data

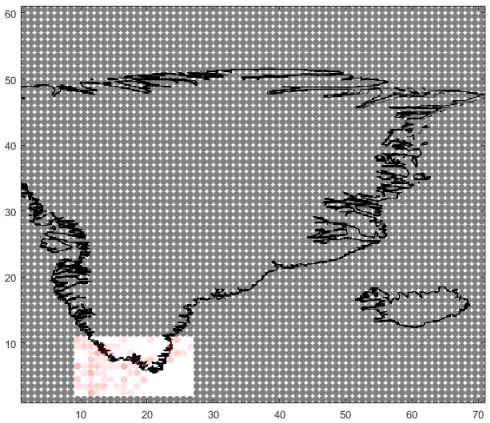


Figure 23.d: $N_{\Delta,1}$ on MAR data

Figure 23: The figures 23.a and 23.b show the results of the $\mathcal{A}_{ga}^{max,6,50}$ algorithm on Greenland region of interest using the MERRA-2 database with respectively mappings $N_{\Delta,\infty}$ and $N_{\Delta,1}$. While Figure 23.c and 23.d show the results of the $\mathcal{A}_{ga}^{max,6,50}$ algorithm using the MAR database with respectively mappings $N_{\Delta,\infty}$ and $N_{\Delta,1}$.

4.9.3 Observation

For Europe, locations on figures 22.a and 22.b are mainly selected on the land and several ones are located in the Mediterranean sea. Also, there are no overlay. Even though both mappings give different results, the selected locations stay in the same neighborhood.

For Greenland, we have the same observations than for Europe. The four figures 23.a, 23.b, 23.c, and 23.d show different selected locations, but chosen in the same neighborhood. Also there are some overlay. For example, figure 23 has 4 dot layers at location $l = (-50.625; 61.5)$. Finally, selected locations with MAR data are gathered compared to selected locations with MERRA-2 data.

Given the 100 selected locations from the 10 best candidates generated by the $\mathcal{A}_{ga}^{max,6,50}$ algorithm, we cannot select only 10 locations to answer the `question 3` properly. Indeed, there

are not 10 locations that seem to be chosen really more times than the others from the 100 selected locations.

Therefore, another approach needs to be considered in order to find 10 locations to answer the `question 3`.

4.9.4 Second methodology

Since the first methodology did not give interesting results, we will try to find another approach to answer the `question 3`.

As second methodology, we will introduce another algorithm defined as

$$\mathcal{A}_{se}^{max,6,50}$$

that runs in an semi-exhaustive way.

The first step of the algorithm consists in selecting 10 starting locations that are properly distributed on the zone to analyze. For each starting location, we will look at all the other locations from the zone to analyze and select the 9 other locations that maximize the number of common critical windows. More formally, the algorithm 1 shows the pseudocode of the semi-exhaustive approach.

Considering all the starting locations and their selected locations, this give $10 \times 10 = 100$ locations in total. In order to answer the `question 3`, we will plot these 100 locations on a map and select the 10 locations that are chosen the most.

To plot these 100 locations, we will proceed as previous methodology (see semi-transparent colored dots procedure).

Algorithm 1: Semi-exhaustive algorithm

Data:

$\mathcal{L}$ the set of all locations contained in the zone
Δ the length of the windows to detect
α the threshold of the windows to detect

return:

the collection $\mathcal{C}$ containing the 10 starting locations and their 9 selected locations (accept occurrences)

Function *semiExhautive($\mathcal{L}$, $isMax$, Δ, α)* **:**

```
//set local variables
nbPoints = 10 ;
nbSimulations = 10 ;
//init a collection allowing occurrences
C = newCollection() ;
for simul = 1 to nbSimulations do
    l = getStartLocation(simul) ;
    L = {l} ;
    for round = 1 to nbPoints-1 do
        L_copy = L ;
        for l ∈ L \ L do
            if isMax and γ_{Δ,α}(L ∪ {l}) ≥ γ_{Δ,α}(L_copy) then
                L_copy = L ∪ {l} ;
            end
            if not(isMax) and γ_{Δ,α}(L ∪ {l}) ≤ γ_{Δ,α}(L_copy) then
                L_copy = L ∪ {l} ;
            end
        end
        L = L_copy ;
    end
    C = C ∪ L ;
end
return C ;
end
```

4.9.4.1 Europe ($\mathcal{L} = \mathcal{L}_E$)

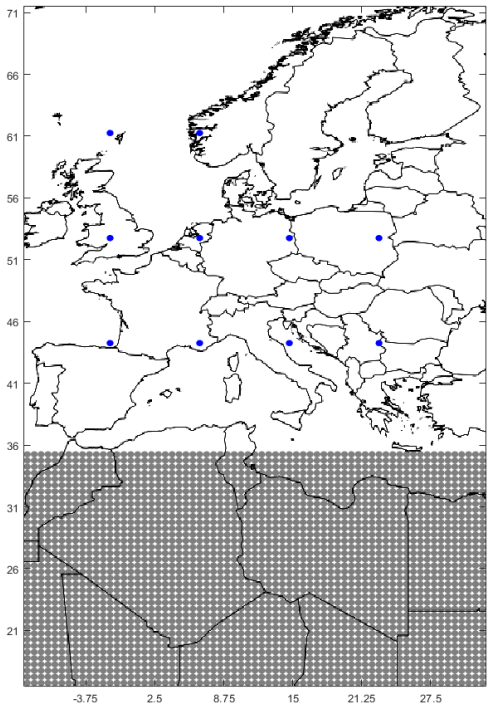


Figure 24.a: The 10 starting points

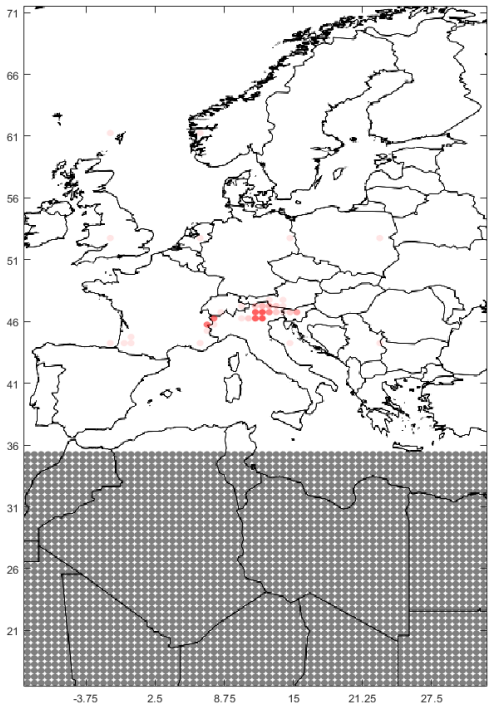


Figure 24.b: $N_{\Delta,\infty}$ on MERRA-2 data

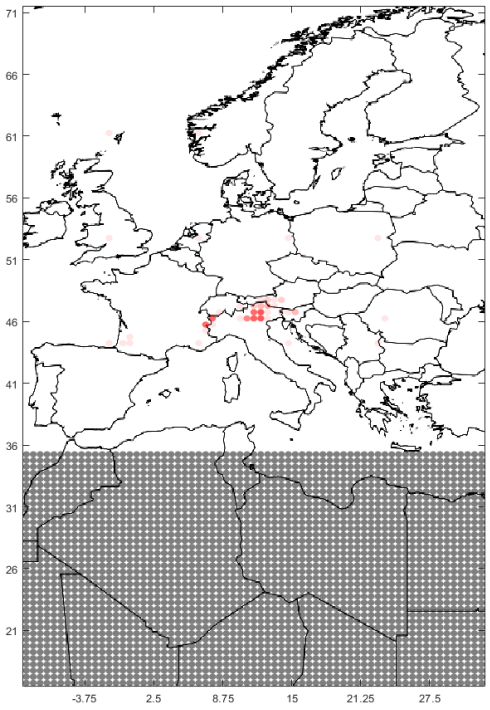


Figure 24.c: $N_{\Delta,1}$ on MERRA-2 data

Figure 24: Using 10 starting points distributed in Europe (Figure 24.a), the figures 24.b and 24.c show the results of the $\mathcal{A}_{se}^{max,6,50}$ algorithm using the MERRA-2 database with respectively mappings $N_{\Delta,\infty}$ and $N_{\Delta,1}$.

4.9.4.2 Greenland ($\mathcal{L} = \mathcal{L}_G$)

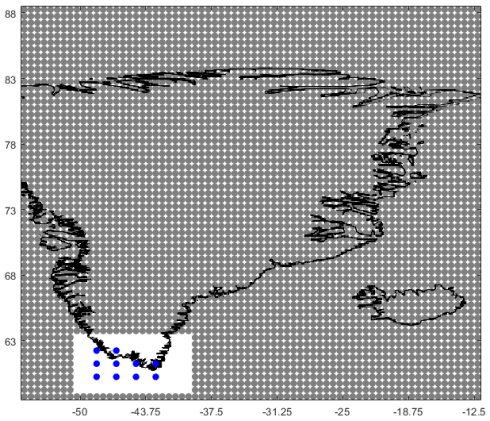


Figure 25.a: The 10 starting points

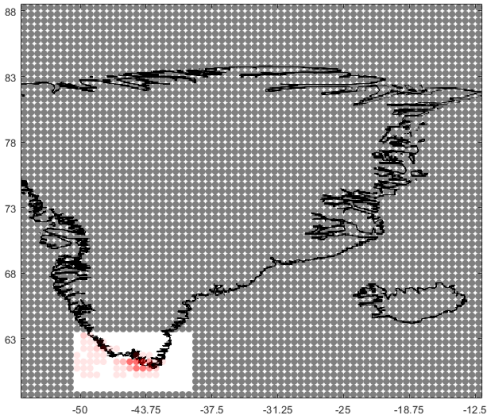


Figure 25.b: $N_{\Delta,\infty}$ on MERRA-2 data

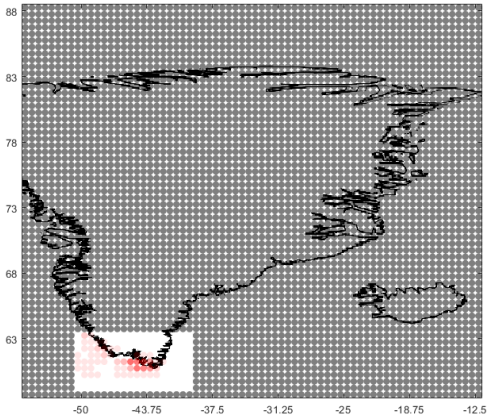


Figure 25.c: $N_{\Delta,1}$ on MERRA-2 data

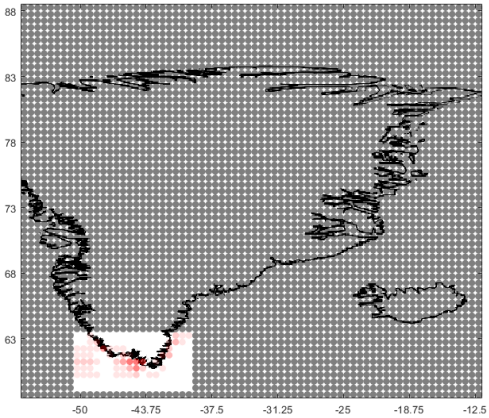


Figure 25.d: $N_{\Delta,\infty}$ on MAR data

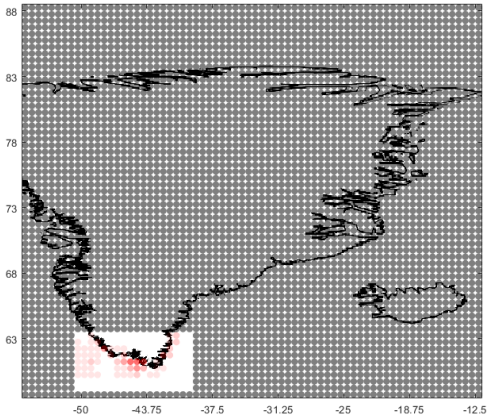


Figure 25.e: $N_{\Delta,1}$ on MAR data

Figure 25: Using 10 starting points distributed in the region of interest in Greenland (Figure 25.a), the figures 25.b and 25.c show the results of the $\mathcal{A}_{se}^{max,6,50}$ algorithm using the MERRA-2 database with respectively mappings $N_{\Delta,\infty}$ and $N_{\Delta,1}$. While Figure 25.d and 25.e show the results of the $\mathcal{A}_{se}^{max,6,50}$ algorithm using the MAR database with respectively mappings $N_{\Delta,\infty}$ and $N_{\Delta,1}$.

4.9.4.3 France ($\mathcal{L} = \mathcal{L}_F$)

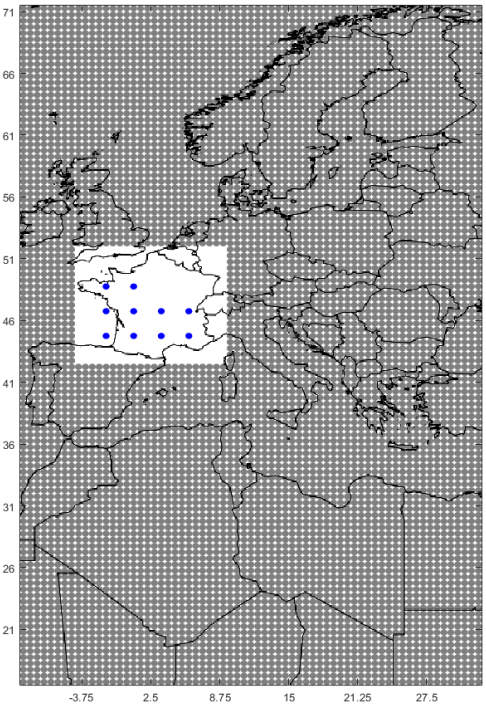


Figure 26.a: The 10 starting points

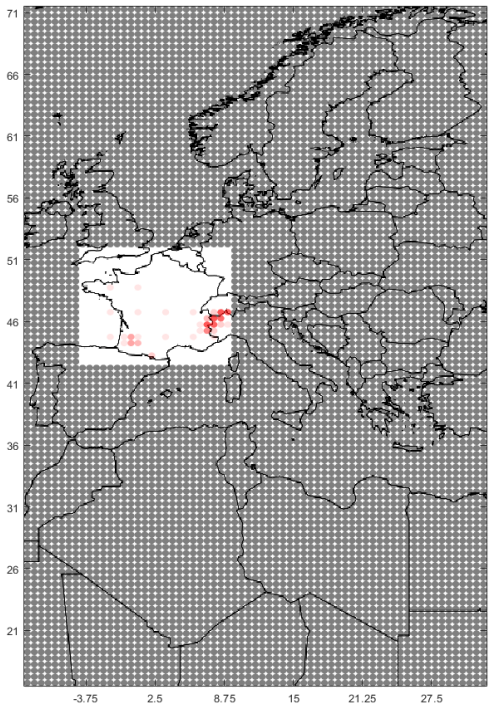


Figure 26.b: $N_{\Delta,\infty}$ on MERRA-2 data

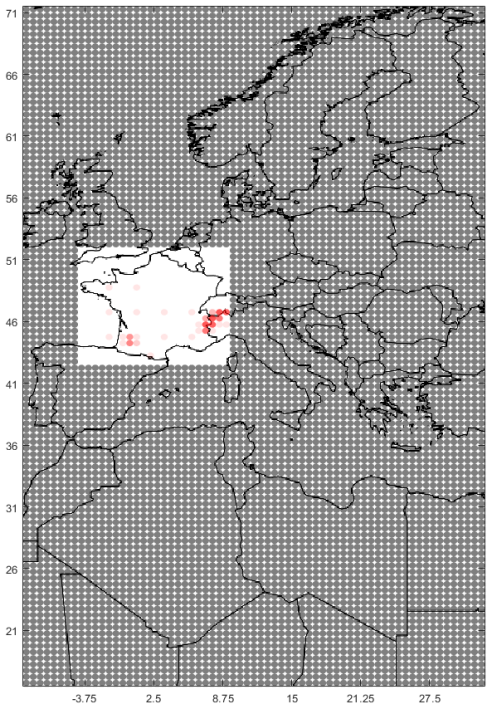


Figure 26.c: $N_{\Delta,1}$ on MERRA-2 data

Figure 26: Using 10 starting points distributed in France (Figure 26.a), the figures 26.b and 26.c show the results of the $\mathcal{A}_{se}^{max,6,50}$ algorithm using the MERRA-2 database with respectively mappings $N_{\Delta,\infty}$ and $N_{\Delta,1}$.

4.9.4.4 Greenland and Europe ($\mathcal{L} = \mathcal{L}_G \cup \mathcal{L}_E$)

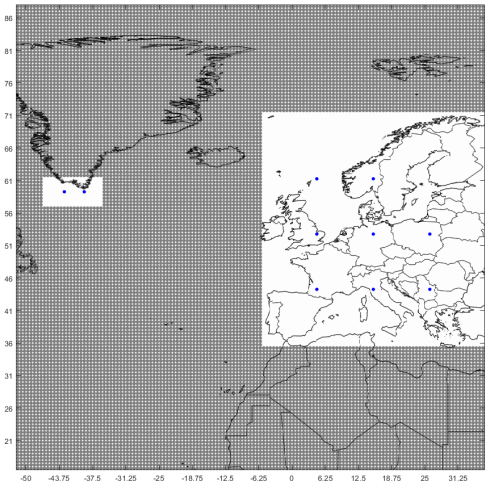


Figure 27.a: The 10 starting points

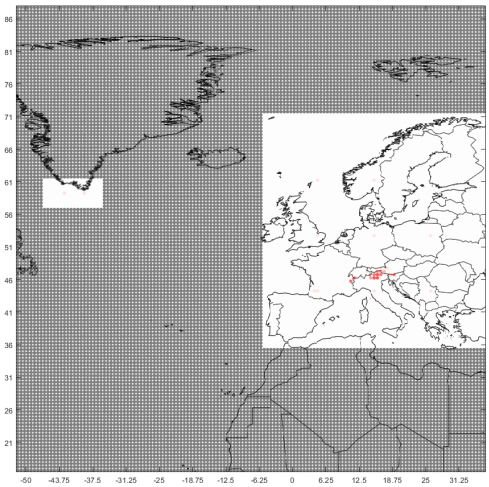


Figure 27.b: $N_{\Delta,\infty}$ on MERRA-2 data

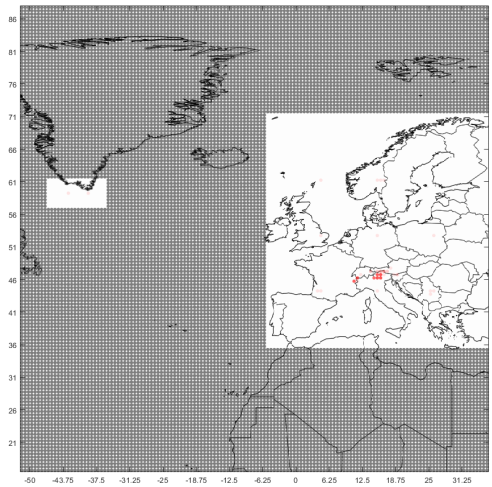


Figure 27.c: $N_{\Delta,1}$ on MERRA-2 data

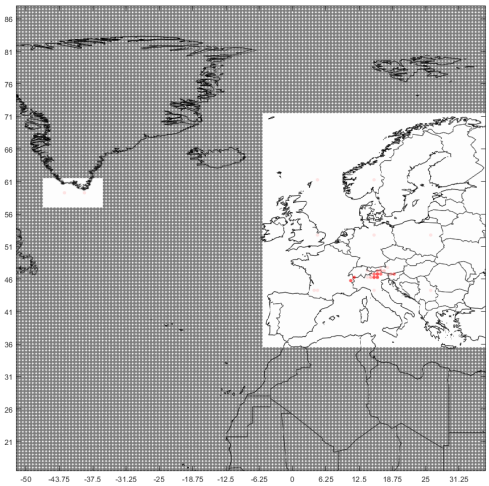


Figure 27.d: $N_{\Delta,\infty}$ on MAR data

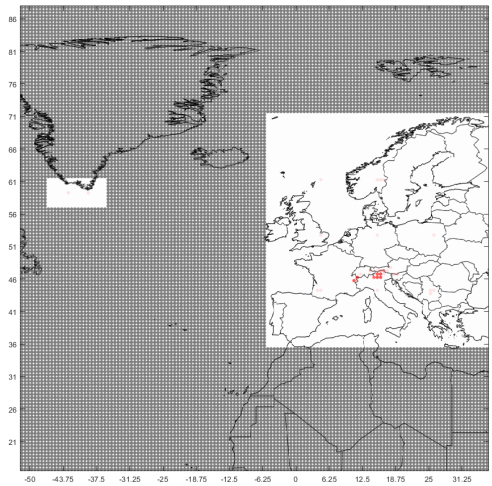


Figure 27.e: $N_{\Delta,1}$ on MAR data

Figure 27: Using 10 starting points distributed in Europe and in the region of interest in Greenland (Figure 27.a), the figures 27.b and 27.c show the results of the $\mathcal{A}_{se}^{max,6,50}$ algorithm using the MERRA-2 database with respectively mappings $N_{\Delta,\infty}$ and $N_{\Delta,1}$. While Figure 27.d and 27.e show the results of the $\mathcal{A}_{se}^{max,6,50}$ algorithm using the MAR database with respectively mappings $N_{\Delta,\infty}$ and $N_{\Delta,1}$.

4.9.4.5 Greenland and France ($\mathcal{L} = \mathcal{L}_G \cup \mathcal{L}_F$)

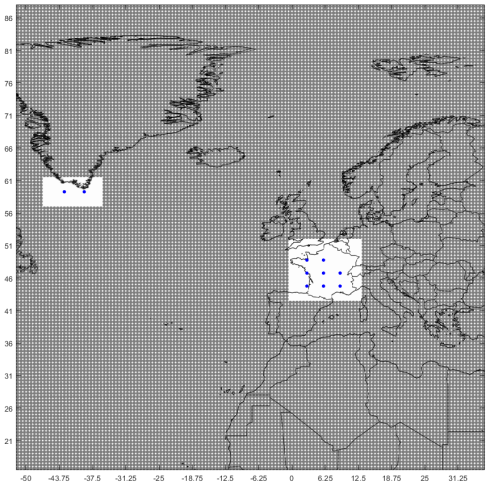


Figure 35.a: The 10 starting points

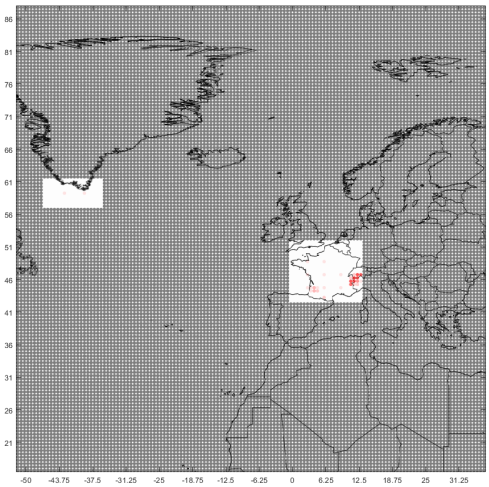


Figure 35.b: $N_{\Delta,\infty}$ on MERRA-2 data

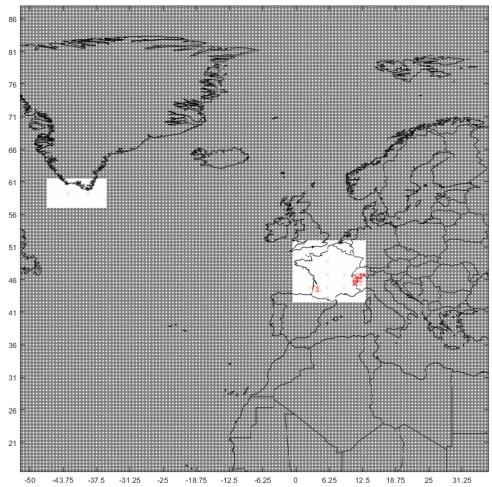


Figure 35.c: $N_{\Delta,1}$ on MERRA-2 data

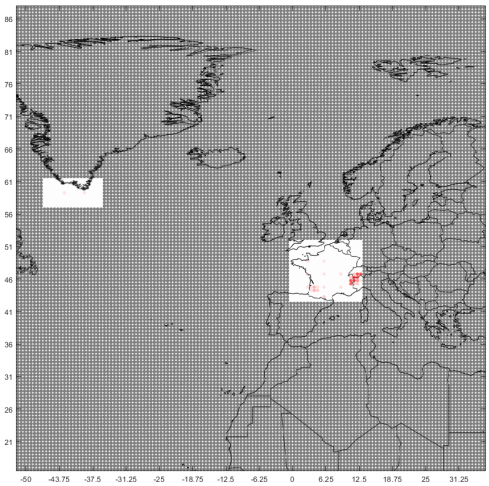


Figure 35.d: $N_{\Delta,\infty}$ on MAR data

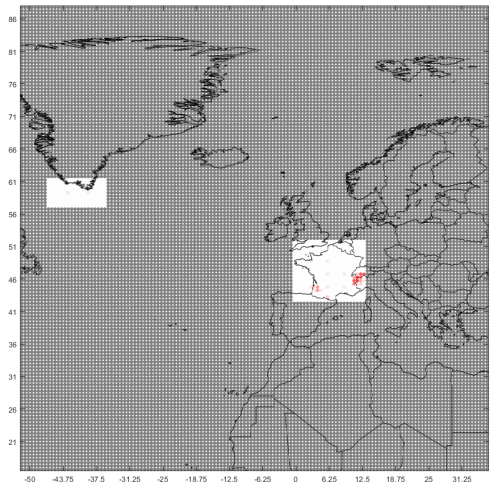


Figure 35.e: $N_{\Delta,1}$ on MAR data

Figure 28: Using 10 starting points distributed in France and in the region of interest in Greenland (Figure 35.a), the figures 35.b and 35.c show the results of the $\mathcal{A}_{se}^{max,6,50}$ algorithm using the MERRA-2 database with respectively mappings $N_{\Delta,\infty}$ and $N_{\Delta,1}$. While Figure 35.d and 35.e show the results of the $\mathcal{A}_{se}^{max,6,50}$ algorithm using the MAR database with respectively mappings $N_{\Delta,\infty}$ and $N_{\Delta,1}$.

4.9.5 Observation

For Europe, both mappings give the almost the same results (figures 24). All the locations that maximize the number of critical windows are located in the Alps. Indeed, wind speeds are very low in this region due in particular to mountains (see figure 5). We also see 7 locations that are selected several times (up to 10 layers at $l = (11.25; 46)$ for figure 24.b).

For Greenland, considering each dataset separately, MERRA-2 data and MAR data show results that are almost the same with both mapping. We observe distinct locations that are selected several time (up to 8 layers at $l = (-46.25; -61)$ for figure 25.b). Even though these locations are not the same depending on the dataset, there are all located at the south coast of Greenland.
-46.25 -61

For France, since the zone also includes the west part of the Alps, locations that are selected the most are located in this section of the Alps (up to 10 layers at $l = (6.25; 4.5)$ for figure 26.b). We can also observe such locations in the South of Nouvelle-Aquitaine, which corresponds to a zone with low wind speeds (see figure 5).

For Greenland and Europe, both mappings and both datasets give the same results. The locations that maximize the number of common critical windows are all located in the Alps, as expected (up to 10 layers at $l = (11.25; 46)$ for figure 27.b). There is no locations selected in Greenland.

For Greenland and France, both mappings and both datasets give also the same results. As for the zone with only France, locations are situated at the west side of the Alps and at the south of the Nouvelle-Aquitaine (up to 10 layers at $l = (6.25; 4.5)$ for figure 35.b). There is no locations selected in Greenland.

As expected, the second methodology highlights that locations that maximize the number of common critical windows are situated in the less windy regions. For Greenland, it corresponds to the South coast (see light blue dots in figure 4). For Europe, it mainly corresponds to the Alps (see light blue dots in figure 5).

4.10 Question 4: 10 points from a given zone such that the proportion of common critical windows is minimal

4.10.1 Introduction

This section will answer the following question:

Question 4

In a specific zone, what are the 10 points such that the average number of critical windows, common to all signals in the zone, is minimal?

The purpose of answering this question is the find the 10 locations from a zone that minimize the number of common critical windows with

$$\Delta = 6 \text{ and } \alpha = 50\%.$$

Finding these 10 locations will illustrate an example of optimal scenario to install wind turbine farms at these locations.

Indeed, since we are looking for the locations that have very few common critical windows, we will tend to a situation where, at any time, there is at least one location with power production.

4.10.2 First methodology

As for answering previous `question 3`, we will try to use our implementation of the genetics algorithm to find the 10 locations that minimize the number of common critical windows. We will run the algorithm with an initial population of 100 random individuals. At the end of the 100 iterations, we will choose the 10 best candidates.

We will then plot the $10 \times 10 = 100$ locations as colored dots on a map. The 10 locations that minimize the common critical windows are the ones that are selected more often.

Hence, here below are plotted the locations selected by the 10 best candidates computed with the $\mathcal{A}_{ga}^{min,6,50}$ algorithm (see section 4.8.3) on the Europe zone and on the Greenland zone (using MERRA-2 data and then MAR data) with both mappings.

4.10.2.1 Europe ($\mathcal{L} = \mathcal{L}_E$)

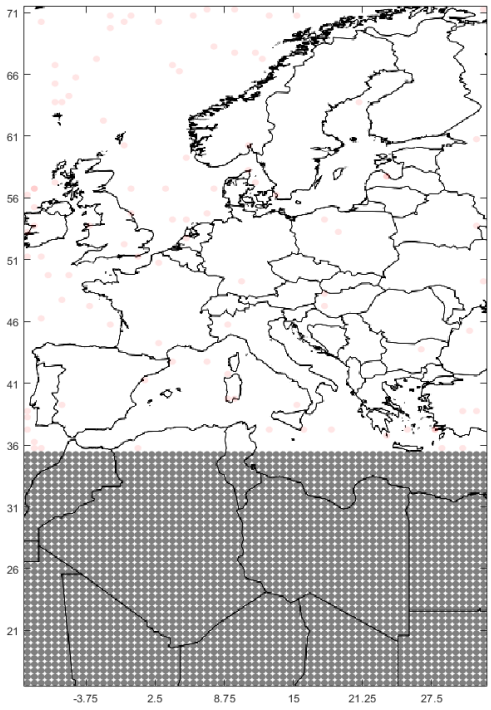


Figure 29.a: $N_{\Delta,\infty}$ on MAR data

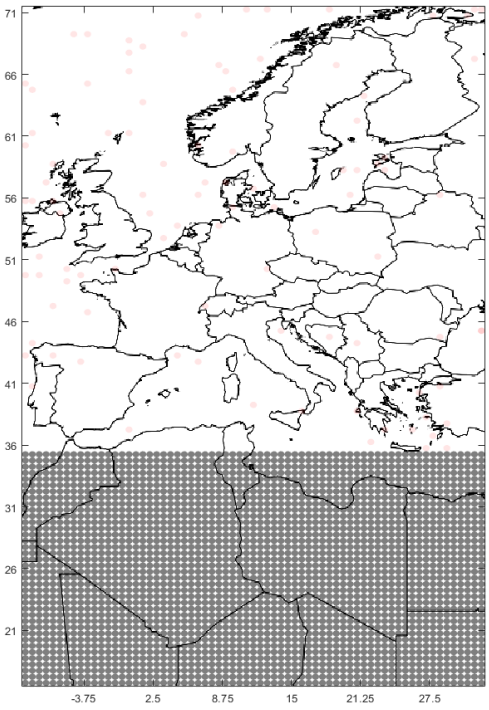


Figure 29.b: $N_{\Delta,1}$ on MAR data

Figure 29: The figures 29.a and 29.b show the results of the $\mathcal{A}_{ga}^{min,6,50}$ algorithm on Europe using the MERRA-2 database with respectively mappings $N_{\Delta,\infty}$ and $N_{\Delta,1}$.

4.10.2.2 Greenland ($\mathcal{L} = \mathcal{L}_G$)

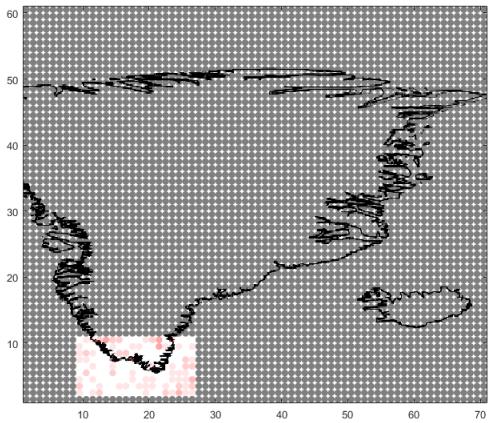


Figure 30.a: $N_{\Delta,\infty}$ on MERRA-2 data

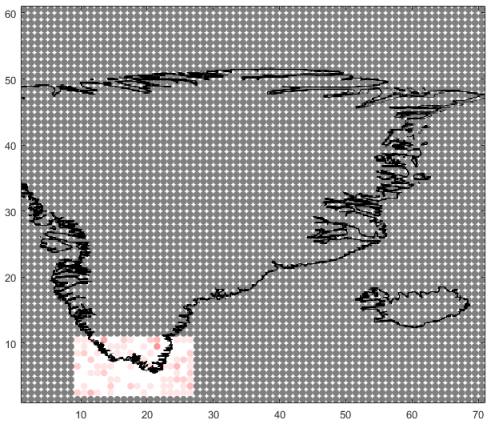


Figure 30.b: $N_{\Delta,1}$ on MERRA-2 data

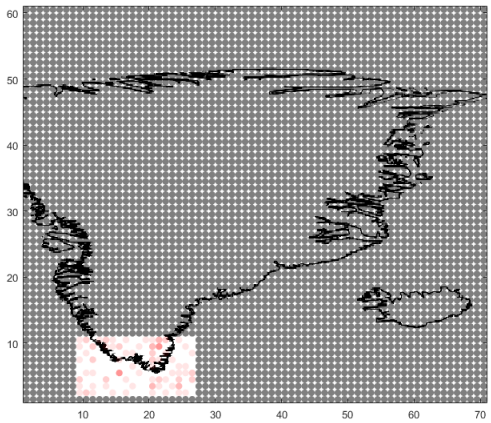


Figure 30.c: $N_{\Delta,\infty}$ on MERRA-2 data

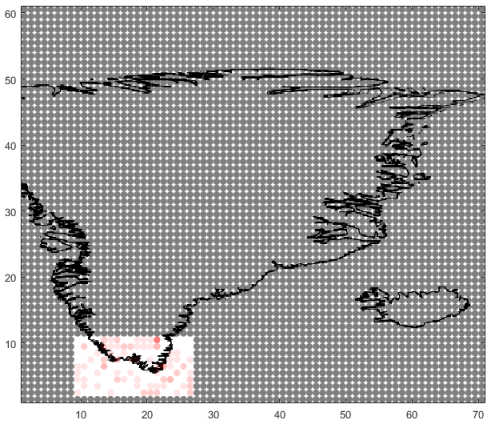


Figure 30.d: $N_{\Delta,1}$ on MERRA-2 data

Figure 30: The figures 30.a and 30.b show the results of the $\mathcal{A}_{se}^{min,6,50}$ algorithm on Greenland region of interest using the MERRA-2 database with respectively mappings $N_{\Delta,\infty}$ and $N_{\Delta,1}$. While Figure 30.c and 30.d show the results of the $\mathcal{A}_{se}^{min,6,50}$ algorithm using the MAR database with respectively mappings $N_{\Delta,\infty}$ and $N_{\Delta,1}$.

4.10.3 Observation

For Europe, as for results on `question 3` for the genetics algorithm methodology, the 100 selected locations are very disparate. There is no location selected several times. Also, we do not see any correspondence between both mappings. However, in contrary to figures 22.a and 22.b, locations are mainly selected in the seas, where wind speeds are very high (see figure 5 with dark blue locations).

For Greenland, observations are the same as for Europe. But since the area is smaller than the Europe area, there are a few locations that are selected several times.

As for `question 3`, the results generated by the $\mathcal{A}_{ga}^{min,6,50}$ algorithm does not give good results to answer the `question 4`. Indeed, we cannot select 10 locations that seem to be chosen more times than the others from the 100 selected locations.

Therefore, we will also consider the semi-exhaustive approach in order to find 10 locations to answer the `question 4`.

Second methodology As second methodology, we will introduce another algorithm defined as

$$\mathcal{A}_{se}^{min,6,50}$$

that runs in an semi-exhaustive way.

As, the $\mathcal{A}_{se}^{max,6,50}$ algorithm, the first step of this algorithm consists in selecting 10 starting locations that are properly distributed on the zone to analyze.

For each starting location, we will look at all the other locations from the zone to analyze and select the 9 other locations that minimize the number of common critical windows.

Considering all the starting locations and their selected locations, this give $10 \times 10 = 100$ locations in total. In order to answer the `question 4`, we will plot these 100 locations on a map and select the 10 locations that are chosen the most.

To plot these 100 locations, we will proceed as previous methodology (see semi-transparent colored dots procedure).

4.10.3.1 Europe ($\mathcal{L} = \mathcal{L}_E$)

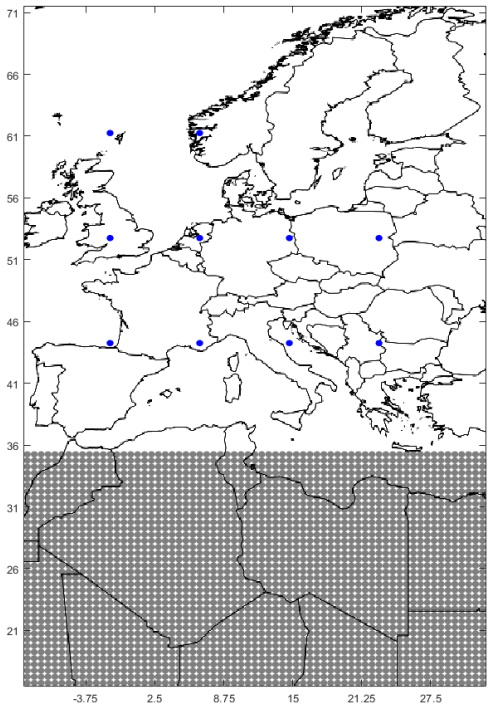


Figure 31.a: The 10 starting points

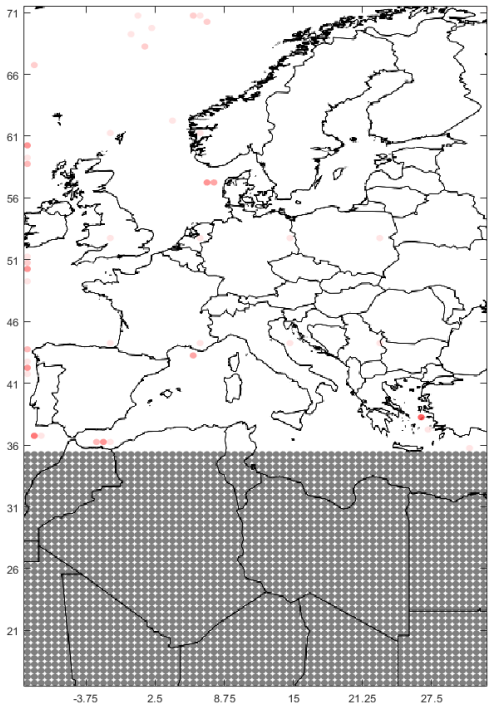


Figure 31.b: $N_{\Delta,\infty}$ on MERRA-2 data

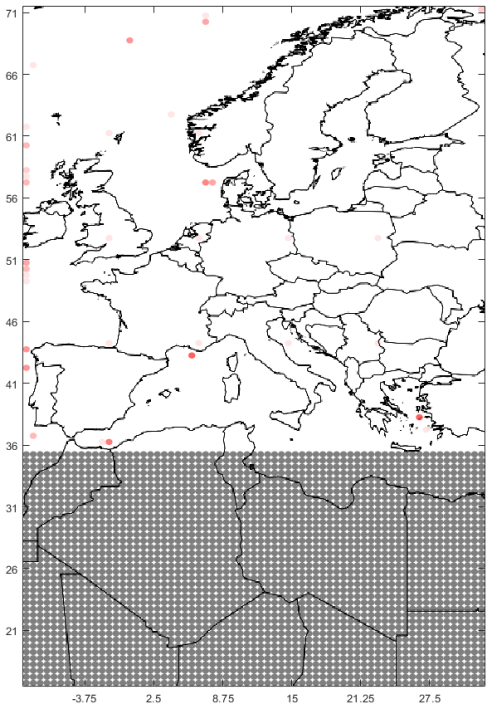


Figure 31.c: $N_{\Delta,1}$ on MERRA-2 data

Figure 31: Using 10 starting points distributed in Europe (Figure 31.a), the figures 31.b and 31.c show the results of the $\mathcal{A}_{se}^{min,6,50}$ algorithm using the MERRA-2 database with respectively mappings $N_{\Delta,\infty}$ and $N_{\Delta,1}$.

4.10.3.2 Greenland ($\mathcal{L} = \mathcal{L}_G$)

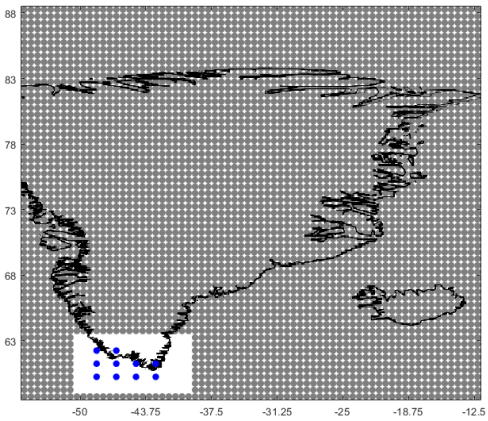


Figure 32.a: The 10 starting points

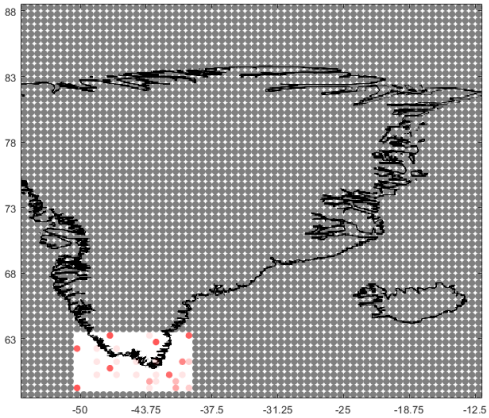


Figure 32.b: $N_{\Delta,\infty}$ on MERRA-2 data

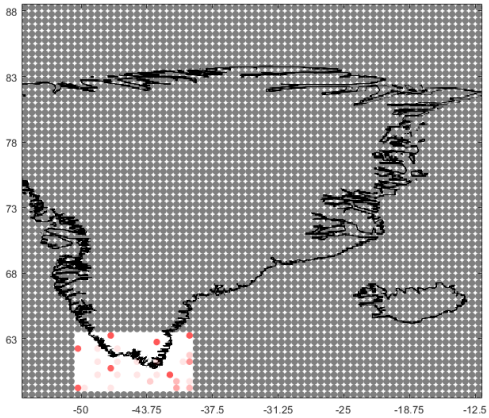


Figure 32.c: $N_{\Delta,1}$ on MERRA-2 data

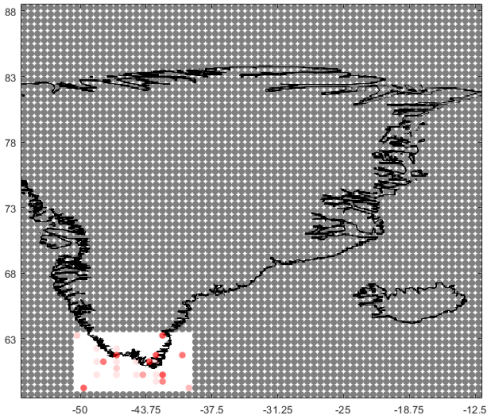


Figure 32.d: $N_{\Delta,\infty}$ on MAR data

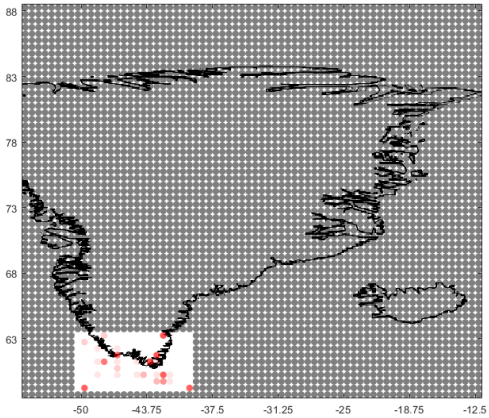


Figure 32.e: $N_{\Delta,1}$ on MAR data

Figure 32: Using 10 starting points distributed in the region of interest in Greenland (Figure 32.a), the figures 32.b and 32.c show the results of the $\mathcal{A}_{se}^{min,6,50}$ algorithm using the MERRA-2 database with respectively mappings $N_{\Delta,\infty}$ and $N_{\Delta,1}$. While Figure 32.d and 32.e show the results of the $\mathcal{A}_{se}^{min,6,50}$ algorithm using the MAR database with respectively mappings $N_{\Delta,\infty}$ and $N_{\Delta,1}$.

4.10.3.3 France ($\mathcal{L} = \mathcal{L}_F$)

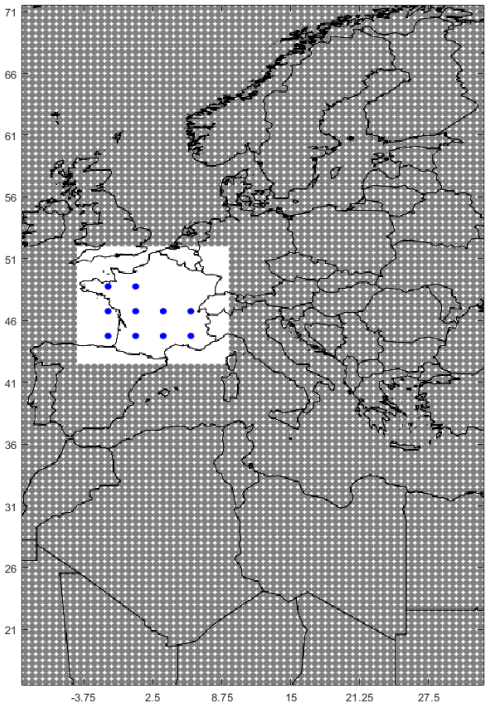


Figure 33.a: The 10 starting points

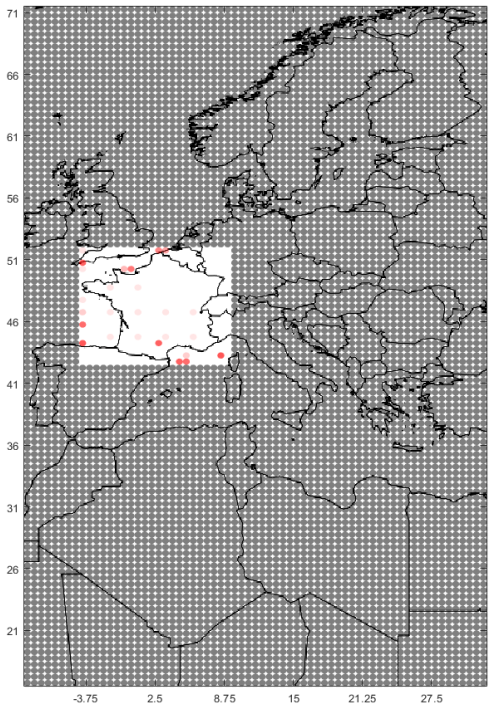


Figure 33.b: $N_{\Delta,\infty}$ on MERRA-2 data

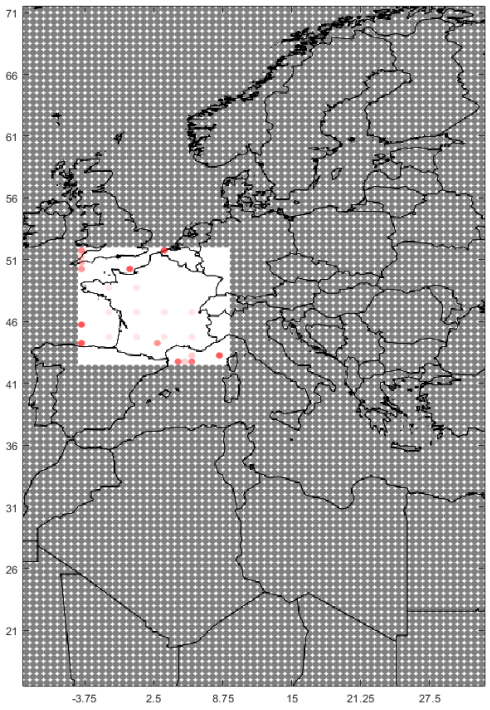


Figure 33.c: $N_{\Delta,1}$ on MERRA-2 data

Figure 33: Using 10 starting points distributed in France (Figure 33.a), the figures 33.b and 33.c show the results of the $\mathcal{A}_{se}^{min,6,50}$ algorithm using the MERRA-2 database with respectively mappings $N_{\Delta,\infty}$ and $N_{\Delta,1}$.

4.10.3.4 Greenland and Europe ($\mathcal{L} = \mathcal{L}_G \cup \mathcal{L}_E$)

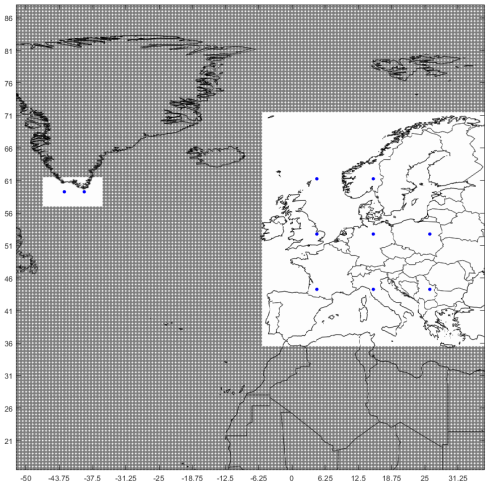


Figure 34.a: The 10 starting points

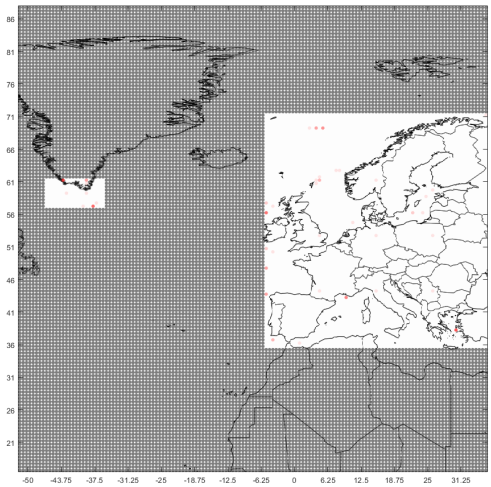


Figure 34.b: $N_{\Delta,\infty}$ on MERRA-2 data

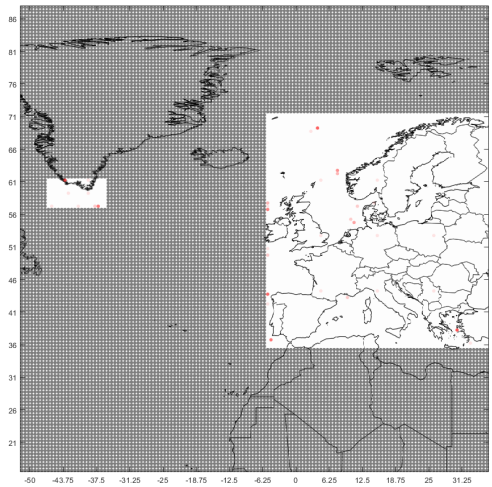


Figure 34.c: $N_{\Delta,1}$ on MERRA-2 data

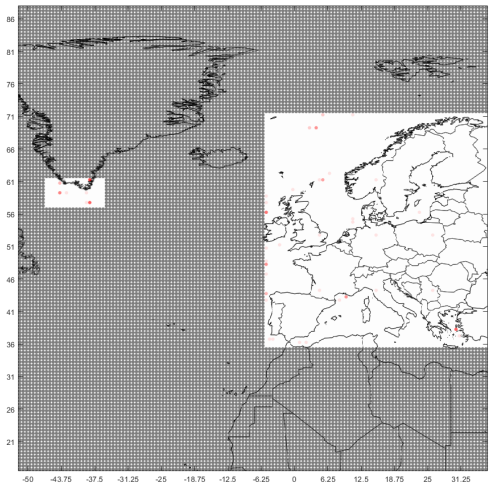


Figure 34.d: $N_{\Delta,\infty}$ on MAR data

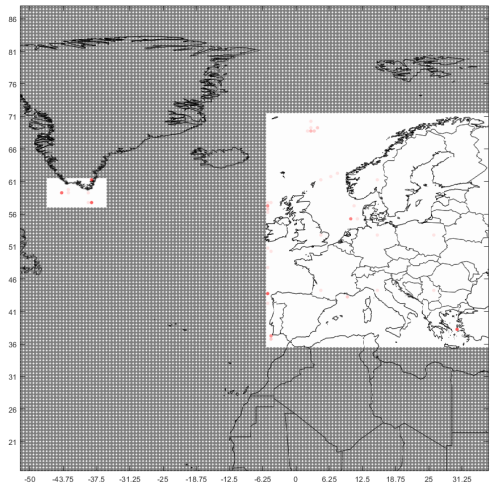


Figure 34.e: $N_{\Delta,1}$ on MAR data

Figure 34: Using 10 starting points distributed in Europe and in the region of interest in Greenland (Figure 34.a), the figures 34.b and 34.c show the results of the $\mathcal{A}_{se}^{min,6,50}$ algorithm using the MERRA-2 database with respectively mappings $N_{\Delta,\infty}$ and $N_{\Delta,1}$. While Figure 34.d and 34.e show the results of the $\mathcal{A}_{se}^{min,6,50}$ algorithm using the MAR database with respectively mappings $N_{\Delta,\infty}$ and $N_{\Delta,1}$.

4.10.3.5 Greenland and France ($\mathcal{L} = \mathcal{L}_G \cup \mathcal{L}_F$)

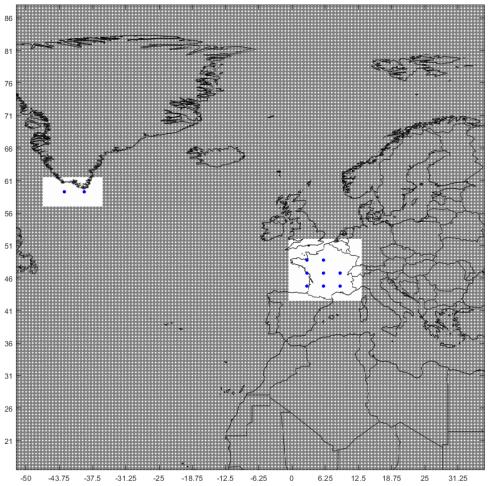


Figure 35.a: The 10 starting points

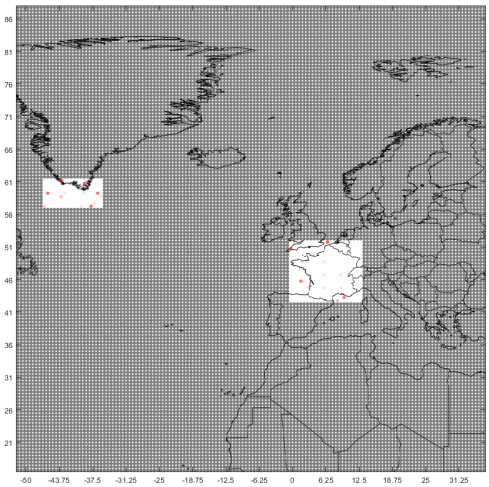


Figure 35.b: $N_{\Delta,\infty}$ on MERRA-2 data

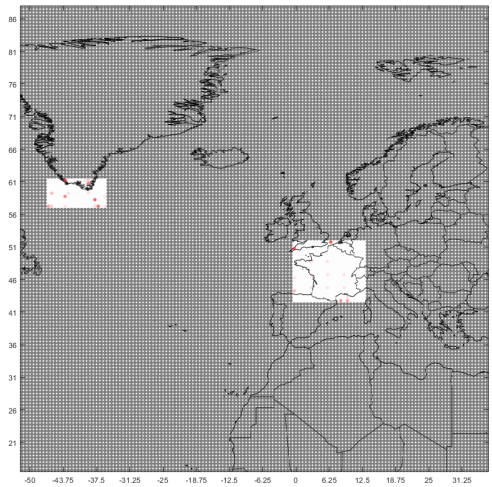


Figure 35.c: $N_{\Delta,1}$ on MERRA-2 data

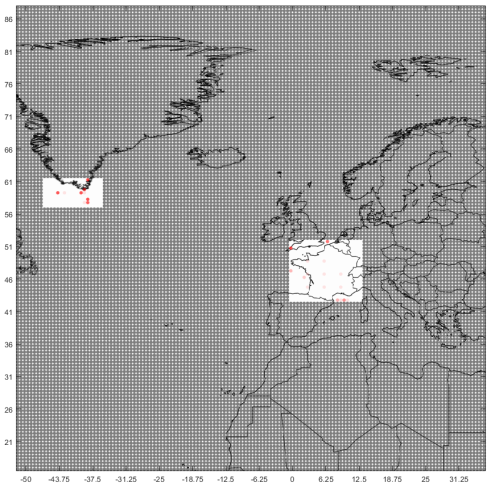


Figure 35.d: $N_{\Delta,\infty}$ on MAR data

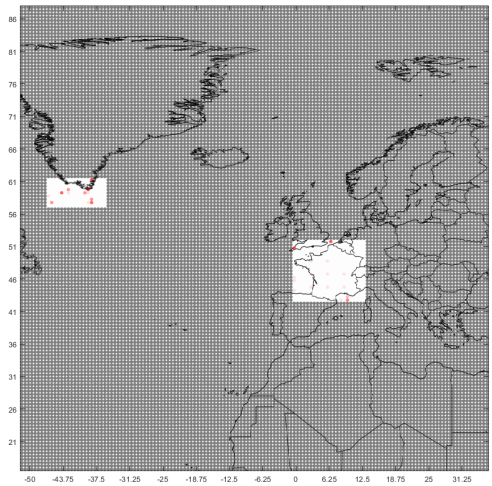


Figure 35.e: $N_{\Delta,1}$ on MAR data

Figure 35: Using 10 starting points distributed in France and in the region of interest in Greenland (Figure 35.a), the figures 35.b and 35.c show the results of the $\mathcal{A}_{se}^{min,6,50}$ algorithm using the MERRA-2 database with respectively mappings $N_{\Delta,\infty}$ and $N_{\Delta,1}$. While Figure 35.d and 35.e show the results of the $\mathcal{A}_{se}^{min,6,50}$ algorithm using the MAR database with respectively mappings $N_{\Delta,\infty}$ and $N_{\Delta,1}$.

4.10.4 Observation

For Europe, all the selected locations are mainly situated in the seas. Both mappings select locations that often match to the other mapping. Also, we can observe that a lot of locations are situated at the boundaries of the zone in order to find wind regimes as different as possible.

For Greenland, we firstly observe location that are selected more times than the others (see dark red dots). This is also due to the area of interest that is smaller than the Europe area. Also, we see the chosen mapping does not modify the results, while the chosen dataset does. Indeed, MERRA-2 dataset maps show dark red dots that are disparate. However, MAR dataset maps show dark red dots that all are located on the sea and on the coast.

For France, we observe same behaviours as for Europe map and Greenland map. First, there is no main difference between both mappings. Then, we clearly see locations that are selected more than the others. It is also due to the France map that is small. Finally, there are a lot of locations that are situated at the boundaries of the France zone.

For Greenland and Europe, there are only small changes between both mappings. Also, even if MERRA-2 and MAR datasets give different results, the selected locations tend to be in the same region for both datasets. We can see dark red dots on the Greenland zone, which mean the algorithm has found locations in Greenland that really minimize the number of common critical windows with locations in Europe.

For Greenland and France, we observe the same behaviours as for Greenland and Europe. However, for these figures, dark red dots in Greenland are really accentuated compared to the previous observations. There are more locations that minimize the number of common critical windows with locations in France than with locations in Europe.

As expected, the second methodology highlights the locations that minimize the number of common critical windows. There are mainly located on the sea and on the boundaries of the map due to big wind speeds at these regions. When considering Greenland with Europe or France, selected locations are situated on both maps.

5 Future works

Also, even if this document presents an interesting introduction to the tool we developed to analyze the critical windows, there are lots of new approaches that may also be introduced.

5.1 New optimization problems

As further analyzes, new optimization problems could be defined. For example, let us define formally new problems in order to avoid $(\Delta, \alpha)-$criticality in a given geographical zone.

For a given pair $(\Delta, \alpha) \in \{1, \dots, T\} \times [0, 1]$, $(\Delta, \alpha)-$criticality is avoided in a given set of locations $L \in \mathcal{P}(\mathcal{L})$ as soon as $I^{(2)}_{\Delta,\alpha}(L) = 0$. This corresponds to a situation for which the empirical average (over years) of the number of critical time-windows simultaneously observed within L is equal to zero, or, in other words, that it has never happened that all localizations for L simultaneously show (over a Δ-time period) a load factor small than α.

Assumption *In the following, we consider two values Δ and α such that: $I^{(2)}_{\Delta,\alpha}(\mathcal{L}) = 0$ (non-criticality is obtained at least when considering all localizations).*

Minimizing the cardinality of a set of localizations while avoiding $(\Delta, \alpha)-$criticality One first optimization problem is to identify one subset of locations of minimal cardinality that still leads to a non-critical situation.

$$\begin{aligned} &\min_{L \in \mathcal{P}(\mathcal{L})} && Card(L) \\ &\text{s. t. } I^{(2)}_{\Delta,\alpha}(L) = 0 \end{aligned}$$

Such a problem may be reformulated using notations introduced in Section 4.10:

$$\begin{aligned} &\min_{L \in \mathcal{P}(\mathcal{L})} && n \\ &\text{s. t. } I^{(4)}_{\Delta,\alpha}(L, n) = 0 \end{aligned}$$

Indeed, there may not be a unique solution to this problem. Algorithmic approach to approximate the solution of this problem: iterative approach, random sampling at every iteration, keeping a localization that minimises the cardinality of the intersection set.

Minimizing the cardinality of a set of localizations while avoiding $(\Delta, \alpha)-$criticality under a partition constraint Assume that the whole set of localizations can be partitioned into several subsets and non-overlapping sets of localizations:

$$\begin{aligned} \mathcal{L} &= \bigcup_{i \in 1}^{n_{\mathcal{L}}} \mathcal{L}_i \\ \forall (i,j) \in \{1, \ldots, n_{\mathcal{L}}\}^2, (i \neq j) &\implies \mathcal{L}_i \cap \mathcal{L}_j = \emptyset \end{aligned}$$

Partitioning such a set may be motivated by the fact that we may want to deploy a minimal amount of capacities in every partition.

Minimizing the cardinality of a set of localizations while avoiding $(\Delta, \alpha)-$criticality under a partition constraint

$$\begin{aligned} &\min_{\substack{L_1 \in \mathcal{P}(\mathcal{L}_1), L_1 \neq \emptyset \\ \ldots \\ L_{n_{\mathcal{L}}} \in \mathcal{P}(\mathcal{L}_{n_{\mathcal{L}}}), L_{n_{\mathcal{L}}} \neq \emptyset}} && Card\left(\bigcup_{i=1}^{n_{\mathcal{L}}} L_i\right) \\ &\text{s. t. } I^{(2)}_{\Delta,\alpha}\left(\bigcup_{i=1}^{n_{\mathcal{L}}}\right) = 0 \end{aligned}$$

Indeed, there may not be a unique solution to this problem. Algorithmic approach to approximate the solution of this problem: iterative approach, random sampling at every iteration, keeping a localization that minimizes the cardinality of the intersection set.

5.2 Generalization of the tool

In this document, we converted wind signals from given locations into power signals using specific wind turbines in order to compute, via the tool we developed, the critical windows that can occur. However, this tool could be used in more general analyzes that include different resources (solar, water, etc) and different technologies (solar panel, hydropower plant, etc).

Hence, we could redefine formally a source signal as follows:

A source signal $\left(s_t^{\sigma,l}\right)_{t\in\{0,\dots,T-1\}}$ of type $\sigma \in \Sigma$ is a non-negative time series associated with a specific location $l \in \mathcal{L}$:

$$\forall l \in \mathcal{L}, \forall t \in \{0,\dots,T-1\}, s_t^{\sigma,l} \in \mathbb{R}_{\geq 0}$$

The type σ specifies the type of resource (e.g., solar or wind), and the set Σ denotes the set of all available types of resources.

Also, a new transfer function and the power time series would be redefined as follow:

Let $f^{\sigma,i}$ denotes a transfer function associated with a given resource $\sigma \in \Sigma$ and given technology for transforming the resource into electricity $i \in \mathcal{I}^\sigma$. This function associates, for any appropriate signal source, a load factor value:

$$\forall \sigma \in \Sigma, \forall (i,l) \in \mathcal{I}^\sigma \times \mathcal{L}, f^{\sigma,i}\left(s_t^{\sigma,l}\right) \in [0,1].$$

For a given type of source signal σ, the set $\mathcal{I}^\sigma$ may contain several technologies (e.g., for wind signal, several types of wind turbines). In the following, we also denote by $\left(u_t^{\sigma,i,l}\right)_{t\in\{0,\dots,T-1\}}$ the time series defined as follows:

$$u_t^{\sigma,i,l} = f^{\sigma,i}\left(s_t^{\sigma,l}\right).$$

6 Conclusion

In this document, we evaluated the *global grid* approach in the case of large electric interconnections between Greenland and Europe. The purpose of such large connections was to take advantage of wind decorrelations to harvest renewable wind energy and to increase the proportion of wind energy in the European power production. To achieve this, our analyzes were based on several optimization problems grounded on the notion of critical time windows, which are series of subsequent time steps during which the load factor of localizations are below a certain threshold.

In order to run our analyzes, we used data generated by two climate models, the MERRA-2 model and the MAR model in the context of Western-Europe and Greenland interconnections. However, since MAR model tends to be more accurate by taking into account katabatic winds, we showed that MAR data from the South-East part of Greenland faced bigger wind speeds than MERRA-2 data from the same region.

The first approach consisted in computing the average proportion of critical windows for wind signals found in a given zone. This approach showed that the average proportion of wind signals from Europe zone is very similar to the proportion of wind signals from the France zone.

However, we noticed smaller proportions when considering the Greenland South coast zone.

As second approach, we computed the average proportion of critical windows that are common to all wind signals from a given zone. For this approach, the average proportion was smaller than the first approach due to the critical windows that have to be common to all wind signals. Hence, we observed a smaller and smaller proportion when considering France, then France and Greenland, and finally Europe (all null proportions). This is also due to the zones that are bigger and bigger and include more and more various wind regimes.

Then, we tried to find the 10 locations in Europe and/or Greenland that correspond to the worst scenario to install wind turbine farms. To do this, we tried to compute the 10 points such that the average number of critical windows, common to all signals in the zone, is maximal. For this approach, we observed that the 10 locations are always selected in less windy regions, which include the Alps in Europe and the offshore side of the South coast in Greenland.

Finally, in order to find the 10 locations in Europe and/or Greenland that correspond to the best scenario to install wind turbine farms, we introduced a last approach. This final approach corresponded to compute the 10 points such that the average number of critical windows, common to all signals in the zone, was minimal. When considering Greenland and Europe separately, we observed that the 10 best locations were situated in the sea, on the boundaries of the region of interest. Indeed, on the boundaries are located the wind signals with wind regimes from the zone as various as possible. Also, when considering Greenland and Europe or Greenland and France together, we noticed that the 10 locations maximizing the total power production were split into both lands.

When analyzing results from these four approaches, we can find that the $(\Delta, \alpha)-$criticality may be too conservative in certain situations (like the second approach showing all null values for a big zone as Europe). Dealing with this problem may be done by defining a criterion able to discriminate among situations based on the proportion of places that are critical (for any time-window).

Further analyzes may also be introduced using the tool we developed. As described in section 5, new optimization problems may be formalized. For example, we defined new problems in order to avoid $(\Delta, \alpha)-$criticality in a given geographical zone.

Finally, as this document only uses the tool in the wind energy context, we proposed to extend our formalization in order to use our tool in more general settings which may include different technologies for harvesting wind energy (i.e., different types of transfer functions), and also different type of resources (for instance, including solar energy).

7 Appendix

7.0.1 Kmeans algorithm pseudo-code

Algorithm 2: Kmeans Algorithm

Data:

k the number of clusters
M the matrix of observations of size I x J
A the matrix of associated clusters for each observation
C[k] the vector for the mean clusters
S^1,...,S^k k sets of observations
$norm^{kmean}(x, y)$ the function that computes the norm between observations x and y
$mean^{kmean}(S)$ the function that computes the mean of the set S of observations.

return:

the matrix A
the vector C

Function *Kmean(M , k)* **:**

```
//initialisation
for c = 1 to k do
    repeat
        x = rand(A);
    until x ∉ C[:];
    C[c] = x;
end
//running algorithm
change = true ;
while change do
    foreach set S^i do S^i = ∅;
    for i = 1 to I do
        for j = 1 to J do
            index = argmax(c = 1 to k) { norm^kmean(M[i,j], C[c]) };
            A[i,j] = index;
            S^index = M[i,j];
        end
    end
    change = false
    for c = 1 to k do
        new = mean^kmean(S^c);
        if norm^kmean(new, C[c]) > 0.001 then
            change = true;
        end
        C[c] = new;
    end
end
return A,C;
end
```

Algorithm 3: $norm^{kmean}$

Data:
a,b two wind speed values

return:
the distance d

Function $norm^{kmean}$ *(a , b)* **:**
```
d = abs( a - b );
return d;
```
end

Algorithm 4: $mean^{kmean}$

Data:
S a set of wind speed values

return:
n the new mean value

Function $mean^{kmean}$ *(S)* **:**
```
n = 0;
foreach signal s in S do
    n = n + s;
end
n = n / |S|;
return n;
```
end

7.1 Wind turbine VESTAS V80-2MW

V80-2.0 MW
Facts and figures

POWER REGULATION pitch regulated with variable speed

OPERATING DATA	
Rated power	2,000 kW
Cut-in wind speed	4.0 m/s
Rated wind speed	16 m/s
Cut-out wind speed	25 m/s
Wind class	IEC IA and IEC IIA
Operating temperature range	-20°C to 40°C low temperature turbine: -30°C to 40°C

ROTOR	
Rotor diameter	80 m
Swept area	5,027 m^2
Nominal revolutions	16.7 rpm
Operational interval	10.8 - 19.1 rpm
Air brake	full blade feathering with 3 pitch cylinders

ELECTRICAL	
Frequency	50/60 Hz
Generator type	4-pole doubly fed generator, slip rings

GEARBOX	
Type	two planetary stages and one helical stage

TOWER	
Type	tubular steel tower
Hub heights	60 m, 67 m and 78 m (IEC IA) 60 m, 67 m, 80 m and 100 m (IEC IIA) 100 m (DIBt)

BLADE DIMENSIONS	
Length	39 m
Max. chord	3.5 m

NACELLE DIMENSIONS	
Height for transport	4 m
Height installed (incl. CoolerTop®)	5.4 m
Length	10.4 m
Width	3.4 m

HUB DIMENSIONS	
Max. diameter	3.3 m
Max. width	4 m
Length	4.2 m

Max. weight per unit for transportation	70 metric tonnes

Figure 36: Characteristics of the V80 2MW.

References

[1] Spyros Chatzivasileiadis, Damien Ernst, and Göran Andersson. The global grid. *Renewable Energy*, 57:372–383, 2013.

[2] Spyros Chatzivasileiadis, Damien Ernst, and Göran Andersson. Global power grids for harnessing world renewable energy. In *Renewable Energy Integration (Second Edition)*, pages 161–174. Elsevier, 2017.

[3] Xavier Fettweis, Jason E Box, Cécile Agosta, Charles Amory, Christoph Kittel, Charlotte Lang, Dirk van As, Horst Machguth, and Hubert Gallée. Reconstructions of the 1900–2015 greenland ice sheet surface mass balance using the regional climate mar model. *The Cryosphere*, 11(2):1015, 2017.

[4] Xavier Fettweis, Marco Tedesco, MR Van de Broeke, and Janneke Ettema. Melting trends over the greenland ice sheet (1958–2009) from spaceborne microwave data and regional climate models. *The Cryosphere*, 5(2):359–375, 2011.

[5] Ronald Gelaro, Will McCarty, Max J Suárez, Ricardo Todling, Andrea Molod, Lawrence Takacs, Cynthia A Randles, Anton Darmenov, Michael G Bosilovich, Rolf Reichle, et al. The modern-era retrospective analysis for research and applications, version 2 (merra-2). *Journal of Climate*, 30(14):5419–5454, 2017.

[6] Melanie Mitchell. *An introduction to genetic algorithms*. MIT press, 1998.

[7] Pradnya A Vikhar. Evolutionary algorithms: A critical review and its future prospects. In *Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC), 2016 International Conference on*, pages 261–265. IEEE, 2016.

[8] Coraline Wyard, Chloé Scholzen, Xavier Fettweis, Jean Van Campenhout, and Louis François. Decrease in climatic conditions favouring floods in the south-east of belgium over 1959–2010 using the regional climate model mar. *International Journal of Climatology*, 37(5):2782–2796, 2017.