Politecnico di Torino

Master degree in Energy and Nuclear Engineering, Renewable Energy Systems

A.a. 2023/2024

Sessione di Laurea Luglio 2024

Impact of Forecast Uncertainty in O&M scheduling for OWF, FCM Machine Learning integration

Relatore: Prof. Giovanni Bracco Candidati: Riccardo Meda, 303636

Contents

Abstract **4**

1. Introduction and Technical outlook **5**

1.1. Introduction 5

1.2. Historical and Current Perspectives on Global Climate Initiatives 6

1.3. Europe and Portugal Energy Outlook 8

1.4. Renewable energy situation 10

2. State of Art of wind energy **13**

2.1. Wind energy production theory 13

2.2. Wake effect theory 17

2.2.1. Gaussian velocity deficit 19

2.2.2. Crespo-Hernandez Wake Turbulence Model 20

2.2.3. Sum of Squares Freestream Superposition 21

2.2.4. Floris: a tool for wind farm performance prediction 22

2.3. Operations and maintenances in offshore wind farm 24

2.4. Weather forecast 27

2.5. Data-driven method – Fuzzy C Means Theory 31

3. Case Study **34**

3.1. Wind farm characteristics 34

3.2. Weather data source 36

3.3. Fuzzy C-mean application 40

3.4. Structure of case study 43

3.4.1. Previous case studies 43

3.4.2. Present case studies 49

4. Results obtained **56**

4.1. AEP for offshore wind farm 56

4.1.1. Validation of FCM integration in AEP weather data 56

4.1.2. AEP results with FCM integration 61

4.2. O&M scheduling optimization 64

4.3. O&M scheduling optimization with FCM integration 70

4.3.1. Validation of FCM data integration in O&M weather data 70

4.3.2. O&M energy results with FCM integration 73

4.4. Sensitivity Analysis 74

4.4.1. Sensitivity Analysis on SWH 74

4.4.2. Sensitivity Analysis on FCM membership limit 76

5. Conclusions and future work 78

Bibliography 81

Abstract

The exploitation of offshore wind farm in energy production necessitates a reduction in cost associated to maintenance and operation that highly contribute to the total cost of the technology. Maintenance activity scheduled as preventive intervention, reduces the probabilities of failure for the turbines, decreasing economic losses due to production loss. Scheduling the operations with priority or urgency considerations, permit to exploit the activities on days with lower wind resources, maximizing the energy production.

Contemplating the perturbations of the wind flow created from the wake effect by upstream turbines for specific wind direction along the farm, is possible to optimize the downtime energy losses. In fact, by a selectively deactivation of specific turbines for maintenance activities, the minimization of energy losses could be pursued, impacting on the overall energy production. For so, the knowledge of accurate weather conditions for the specific wind farm location becomes fundamental. Uncertainties in weather forecast might affect the decisions taken in maintenance scheduling, reaching suboptimal or non-optimal results. The present study evaluates the impact of forecast uncertainties in the optimization of scheduling preventive maintenance operations and quantifies their energetics consequences. The results obtained show lower optimizations than expected, due to inaccuracy of weather forecasted. An integration with a trained data-driven method with historical data is proposed to reduce forecast uncertainties in the estimation of wind directions. The performance of the Fuzzy C Means algorithm used to improve accuracy of the forecast varied across different months and wind speed scenarios analyzed. An improvement in the accuracy of the estimation of wind direction for maintenance purpose it has been reached obtain a small decrement of energy losses. The algorithm shows higher precision in annual energy production estimation. These findings underscore the potential of data-driven methods in mitigating weather forecast uncertainties for remote locations and the possibilities to apply them in maintenance scheduling activities for offshore wind farms.

1. Introduction and Technical outlook

1.1. Introduction

This work aims to reduce the power losses generated from the inactivity of turbines due to maintenance operations.

For an increasing need of sustainable energy production, the reduction of the total cost of the systems has a crucial role to foster the elimination of the fossil fuel technologies in energy generation. To achieve similar objectives is necessary to decrease the cost of energy of the system installed.

Offshore wind farms (OWF) employ the use of a resource that has no cost, for this fact, the initial investment and the maintenance activities represents the most part of the expenditures. Effort to reduce the mentioned costs could increase revenues and the feasibility of new projects.

Preventive maintenance is an important part of O&M that helps to avoid unexpected component failure leading to economic costs and elevated downtime energy losses that increases with the impossibility or slow reactance of organization, to repair the damaged part. To reduce the probability of unforeseen lack of energetic production, these activities are generally planned along the year and are repeated for the entire lifetime of the OWF. The evaluation of O&M needs to consider various parameters such costs, technical availability, failure probabilities and energy generation.

Furthermore, the aerodynamic interactions that are present in a wind farm modify the energy production due to the variation of wind velocities encounter at each rotor of wind turbine (WT). This wind flow perturbation derived from the wind energy extrapolation of the aerogenerator is called wake effect. As O&M requires that the WT need to be shut down for safety reason, the aerodynamic interactions between the downstream WTs is modified. It has been proved that the wake effects magnitude changes mainly due to the wind velocity while the aerodynamics characteristic of such is dependent mainly by the wind direction that flows in the OWF [1].

Taking so into consideration, scheduling preventive maintenance on weather conditions can reduce downtime losses and yield the energy production of the OWF, if forecast weather condition are able to properly preview the meteorological events with a certain precision.

The use of forecast weather data for remote locations could lead to imprecision or non-optimal decision to evaluate the scheduling of O&M. To reduce such risks, a machine learning algorithm (ML) can be implemented for a better understanding of weather conditions in a certain location at different periods of the year.

Simple unsupervised ML such as the Fuzzy C Means (FCM) employed on the only parameter of wind direction can enhance the precision of the forecasted condition with a low computational and resources cost. In this way, improvement in decision for preventive maintenance scheduling of OWF could be achieved.

This study is a further development of [1] [2], introducing forecasted atmospheric weather condition to decision making and integrating a cluster algorithm to refine the identified trend in a daily wind direction. In specific, it investigates different potentials strategies in scheduling preventive maintenance for OWF considering wake effect interaction, different scheduling evaluations and using various lead time forecasted data that could be improved with FCM algorithm.

1.2. Historical and Current Perspectives on Global Climate Initiatives

For the first time in the 1992 during the United Nations Framework Convention on Climate Change also known as Earth summit, in Rio de Janeiro, Brazil, it has been discussed about the limitation of greenhouse gasses (GHG) emission in the atmosphere. 30 years later during the COP27 it has been reached an agreement to create a loss and damages fund to tackle the effect of climate change on developing countries and create a path to implement the gradual phase down from fossil fuels to reach a Net Zero Emission (NZE) carbon by 2050. It is forecast that such scenario should limit the global warming to 1.5 °C comparing to the preindustrial era. Actually, the global average surface temperature is already 1.2 degrees Celsius, leading to an increased likelihood of heatwaves and extreme weather events [3].

The announced ecologic transition to respect the objective set for a decreasing of carbon gas base production will undergo all the sectors that generate GHG emission. In 2019, approximately 34% (20 GtCO2-eq) of total net anthropogenic GHG emissions comes from the energy supply sector, 24% (14 GtCO2-eq) from industry, 22% (13 GtCO2-eq) from agriculture, forestry and other land use (AFOLU), 15% (8.7 GtCO2-eq) from transport and 6% (3.3 GtCO2-eq) from buildings [4].

To ensure a green innovation, different key points are needed, including: increase the use of renewable resources for energy generation, foster electric mobility, improve the availability of energy storages, create smart building strategy, establish a circular economy model and revolutionize the agroecological model. The energy sector, being the most responsive and conducive sector for technological innovation, represents the primary investment opportunity for an ecological transition. Investment in this sector continues to rise, reaching $1.4 trillion USD in 2022, accounting for 70% of the total growth in energy sector investment. Anyway, the fossil fuels still account for the 80% of the primary energy mix showing that the progress to do are still many. To complete a transition to clean energy, it is also essential to establish and enhance the structure of clean energy supply chains where 1.2 USD trillion of cumulative investment are required to fit the NZE Scenario's 2030 target. Actually, only the 60% of the investment required are announced, such need requires a fast intervention by policy makers to be able to establish supply chains in time [3].

The world situation in 2022 sees China as first GHG emitter with more than 29% of the total worldly amount followed by United States and India with respectively 11.2% and 7.3% while the Europe Union 27 is located at the fourth place accountable for the 6.7% of the total [5]. Furthermore, higher number of countries are entering a rapid industrialization and urbanization phase increasing the energy demand. For them the ecological transition is more difficult than developed regions, where for some, the economies are strongly depended on agriculture, making them even more vulnerable to climate change consequences.

Furthermore, each states have different internal situations, where the demography, social inequality, economic crisis, infrastructure and geopolitical factors play important roles in the planning and actuation of the green transition.

Ecological transition represents a potential for economies and countries that are able to seize it, specializing in green energy production. At its hearts is the concept of interdependence between natural, economic and social phenomena, which is fundamental to succeed in embarking on a path to sustainable development. This opportunity represents not only a measure of CO2 reduction but also a chance to increase life quality, reduce human impact, ensure conservation of natural resources and stimulate economies. The green transition is in fact, expected to create new jobs opportunity with estimates suggesting a doubling compared to 2021, reaching 70 million new positions [6].

Given the topic of this thesis, the exploitation of renewable technologies will mainly be taken in consideration, with a deeper focus of Europe Union and Portugal technological situation where this work is set.

1.3. Europe and Portugal Energy Outlook

Before the 17th century, the demand for energy was quite limited, serving only a few needs, typically met through the direct utilization of local natural resources such as wind or water for mechanical power, and wood or coal for heating purposes. With the advent of the industrial revolution and the application of machines in production processes, the need for power began to escalate. The first centralized power plant was established in New York for lighting purposes, utilizing a steam engine powered by coal, marking the begun of the era of electrification. Subsequently, hydroelectric technology was introduced for energy production. However, it was in the late 1800s that a new dominant power source emerged: the petroleum. Since then, when the total energy demand was around 5.65 TWh, petroleum has remained one of the primary energy sources, shaping the modern energy landscape. During the XX century new sources of energy have been exploited as nuclear and renewable energy, leading to challenges in designing and optimizing new and more efficient technologies. Anyway, with a constant rise in global population and constant electrification, the global energy demand keeps increasing by 1% to 2% each year reaching in 2022 a demand of 178,9 TWh of energy [7].

Europe it has been a precursor in the clean energy production diversifying its energy resources, promoting renewable energy technologies and adopting directives for efficiencies improvements. In the 2019 the European leader committed in achieving a climate neutrality by 2050 reducing the emissions to the only amount of GHG that can be absorbed by the environment. To do so, it adopted a pack of rules and policies stipulated in the European Green Deal (EGD) that comprehend interconnected initiatives in climate, environment, energy, transportation, industry, agriculture and financial sustainability. The EGD set goals for not only the 2050 objective but also visualize 2030 as a milestone for representation of the pathway taken. In fact, as each sector addressed is composed by complex parameters and many different interconnected factors, determine a reduction rhythm of emission reduction is fundamental. Each EU member state have so committed to at least 55% of emission reduction to reach in 2030 [8]. In 2023 it has been achieved falls records in Europe for coal and gas power generation. Such resources dropped by 26% (-116 TWh) and 15% (-82 TWh) respectively from their lowest ever value helping in the reduction of 19% (-157 Mt of CO2eq) for the power sector emissions [9].

Portuguese CO2 emission situation kept rise from 1990 till it reached a peak in 2005 with almost 70 million tons of gas released in the atmosphere [10]. Following the European direction promoting the renewable energy resources, Portugal has largely increased its clean energy production thanks also to the huge renewable sources potential that characterize this country. Consequently, to the latest directive applied that made possible a GHG emission reduction in energy sector, in 2018 the transport sector started to become the most pollutant category in the country overtaking the emission from the energy production category. Registering in 2023 a total energy production of 50728 GWh made by more than 60% from renewable sources, the country still has an annual energy import of around 19% of the total [11]. In this context, as established in the PNEC, the goals for 2030 of renewable energy penetration in electricity generation and transport sector are set respectively to 80% and

20% [12]. To do so, prospectives see a major increase of solar and wind energy production technologies in the national energetic mix where a target of 20.4 GW of PV and 10 GW for offshore wind energy installations have been set [13].

1.4. Renewable energy situation

In a situation where the global primary energy consumption is still composed by the majority of fossil fuels sources exploitation, the renewable technologies are trying to be more relevant in power generation system. In 2022 the renewable provided 30% of electricity generated worldwide, with an expectation to rise to 50% by 2030 [3]. Hydropower is keep remaining the largest low-emission source of electricity, with low future installation planned, due to high costs and limitation for favorable site. Solar PV and wind are expected to dominate the future power capacity addition to reduce use of carbon base fuel systems. They are expected to fill almost the 70% of new power installation between now and 2050 in various analyzed scenarios [3], since they are the cheapest source of green electricity available in the market and are already widely known.

Anyway, clean energy technologies supply chains are more geographically concentrated than fossil fuel manufacturing and mining. For system as solar PV, electric batteries and electrolysers, the three largest regions hold more than the 80% of global capacity supply chains. To modify such situation and create a diversification in supply chains it is needed to establish industrial strategies and international co-operation. With these systems use, also the require of rare materials will rise demanding a rump-up in mining activities in diversified regions to avoid deceleration in supply chains. Limiting improvement are instead expected for diversifying in supply resources, where actually China holds half of the planet Lithium chemical plant and Indonesia rule nearly 90% of nickel refining facilities [3].

In 2022 solar PV installation doubled compared to 2019 reaching a record of 220 GW capacity addition rising to over 1000 GW total capacity accounting for almost 4% of global electricity generation. In 2021 the global wind generation capacity achieved 830 GW with an addiction in 2022 decrease by almost a third below the peak in 2020 but it is expected to rise to 175 GW per year [3]. China guides the solar PV market manufacturing representing 70% of silica-based PV production and accounts for almost half of its capacity installation in 2022, followed by Europe and United States [3]. Also for wind power market, China represent the largest capacity addiction for this technology representing around 40% of systems installation.

In order to aim to the target set, the scaling up of solar PV and wind power generation require action in permission and licensing of new project, develop and improve grid connections, expanding wind manufacturing capacity and enhance power system flexibility to rise such technologies integration. Unfortunately, the onshore and offshore wind component manufacturers are experiencing difficulties due to high financing cost, long construction time, upgrading manufactory time in response to technology advancement and fierce competition. Expansion in this field have been announced but are still far below the NZE scenario requirement.

The wind energy technology possesses several characteristics that render it preferable over others energy systems. It harnesses an infinite source of energy, thus bolstering the energy independence of a country. Moreover, it is suitable for remote areas, easily scalable, operates with high efficiency, and represents an economically competitive technology that is already mature and available on the market. This system is under continuous evolution where bigger wind turbines (WT) are being produced. With the increase of the size of the aerogenerator, the power produced rises, with a consequence of less space occupied for same amount of energy produced compared to old type of WTs. Actually, the most power turbine reach 16 MW of capacity with a diameter of 242 meters and a height of 264 meters [14].

Due to the reduction of morphologic limitations and the decrease of intermittency of the wind in the sea that increase the reliability of the energy generation, the offshore wind farm (OWF) represents a promising innovation on energy production [15]. Its capacity projections are almost 270 GW by 2050 [16]. Specially with the advent of floating technologies that allow to exploit this resource in more deep waters. In fact, comparing to offshore fix turbines, a floating farm ensure a higher flexibility in locations selection, can reduce the installation costs by assembling the WT in ports and reduce the interference with the marine ecosystems [17]. OWF entail higher costs as they require significant investments for installation, primarily due to their remote locations and the necessity for platforms or foundations. In fact, the levelized cost of energy (LCOE) for these systems increases from $34 per megawatt-hour (MWh) for onshore turbines to $78 or $133 per MWh for fixed or floating offshore WTs in 2021 [18]. Additionally, they face increased contingencies and maintenance expenses, attributable to the challenges posed by marine conditions and the necessity of specialized equipment to access the farms.

Different typology of floating stations are under study, where for now the most known are the spar-buoy, the semisubmersible and the tension leg platform [19]. They differ in vary aspect as cost, performance, material use, stability and installation strategy. For this work a semisubmersible platform is assumed, as the case study analyzed is based on an existing wind farm with such floating stations.

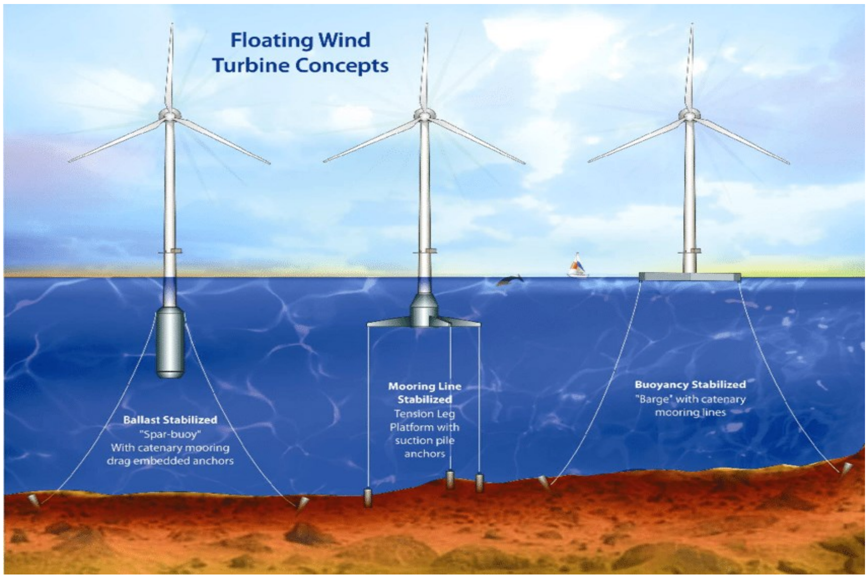

Figure 1: Floating wind turbine platforms retrieved from [20][1]

[1] Every font of the images and tables are referred to the bibliography

2. State of Art of wind energy

2.1. Wind energy production theory

The wind phenomena is created by the heat of the sun that enter the atmosphere of Earth generating gradients of temperature on the surfaces. As different part of the globe receives and emits a non-homogeneous amount of heat, the variation of surfaces temperature generates air pressure variations initiating air mass motions.

The wind turbines operation does not surpass the Atmospheric Boundary Layer (ABL) that characterize the wind behaviors below 1 km height from the surface. The wind in ABL depend mainly on three different aspect: momentum transfer created by terrain elevation and surface roughness z_0, diurnal heat and the presence of moisture.

The velocity profile in ABL is usually modelled using a simple logarithmic approximation:

$$U(z) \propto \ln\left(\frac{z}{z_0}\right) \tag{1}$$

Therefore, knowing the wind velocity at a certain reference height h_{ref} can provide the wind velocity at a hub height U_{hub} through the wind shear equation [21]:

$$U_{hub}(h, z_0) = U_{ref}\frac{\ln\left(\frac{h}{z_0}\right)}{\ln\left(\frac{h_{ref}}{z_0}\right)} \tag{2}$$

where U_{ref} represents the wind speed at a reference height h_{ref}, z_0 represent the surface roughness of the terrain that change as follow:

Table 1: Surface Roughness z_0 for various type of terrain, retrieved from [22]

Terrain Description	Roughness value
Very smooth, ice or mud	0.00001
Calm open sea	0.0002
Blown sea	0.0005
Snow surface	0.003
Lawn grass	0.008
Rough pasture	0.01
Fallow field	0.03
Crops	0.05
Few trees	0.1
Many trees, hedges, few buildings	0.25
Forest and woodlands	0.5

Wind energy production depends on the interaction between the wind and the rotor of the aerogenerator. The production of mechanical power is related to the extraction of the kinetic energy of the air stream flow moving upwind to the rotor where the total available power present in the air stream is represented by:

$$P(v) = \frac{1}{2}\,\rho A U^3 \tag{3}$$

having ρ and A are the air density and the rotor Area, respectively; with U as the wind velocity.

Betz demonstrated the energy generation by calculating the power production from an ideal turbine rotor [23]. This ideal rotor is characterized by several assumptions: it has no hub, an infinite number of blades resulting in the absence of drag resistance, a non-rotational wake, and uniform wind flow through the rotor's cross-sectional area. Additionally, the wind is assumed to have only an axial component with varying velocities related to the distance from the rotor.

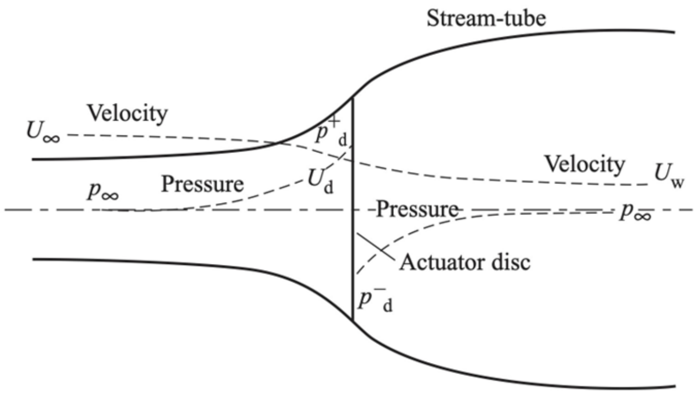

Figure 2: Ideal actuator disk from [24]

In such assumption the maximum theoretical power coefficient reached due to physical limitations is 59.3% of the total energy [25]. The introduction of turbulences behind the rotor, a finite number of blades and an existence of aerodynamic drag reduce the maximum power extraction. So, the evaluation of power production it can be made through the equation:

$$P(v) = \frac{1}{2}\,\rho\pi\,R^2 C_p(\lambda) U^3 \tag{4}$$

where R is the rotor radius, $C_p(\lambda)$ is the power coefficient, which is a nonlinear function of the tip speed ratio (TSR) $\lambda = \omega R/U$, that is translational speed at the tip of the turbine blade to the actual velocity of the wind with ω as rotor speed.

The general value of power coefficient for horizontal 3 blades WT considering mechanical and electrical efficiency energy in conversion is in a range between 0.35 to 0.45 [26]. Control regulations are operated to work in optimal power production and ensure structural safety, the most used method is the Variable-speed variable-pitch. In fact, three different regions of operation exist related to the wind velocity:

- U_{cut-in}: Cut in wind speed is the velocity at which the turbine starts to produce more power than what the electric resistance withdrawn when in operation, generating energy. At an increase of airflow velocity, the rotor spin faster.
- U_{rated}: At rated wind velocity the rotor reaches the tip speed ratio at which the WT has is higher power coefficient. At higher wind speed a pitch regulation modify the angle of attack of the blades maintaining a constant rotor angular velocity and a constant rated power produced.
- $U_{cut-off}$: For velocity of the wind higher than $U_{cut-off}$ the blades are set at an angle that reduce the energy production to prevent the turbine from structural failure. This control could be made with pitch-to-feather or with pitch-to-stall strategy.

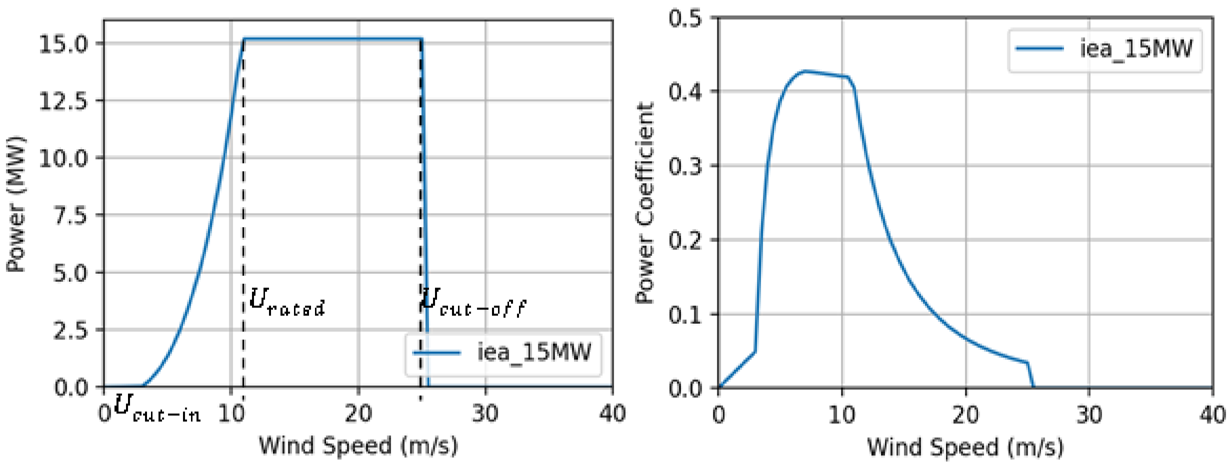

Figure 3: Power curve and C_p curve for IEA 15 MW turbine present in FLORIS library [56]

Regulating in such way the operation of the turbine on different wind speed velocities, is possible to harvest the maximum energy from the aerogenerator. The calculation of energy produced is made by the integration of the power produced along the time at the variation of the inflow wind speed.

$$E(T) = \int_0^T P(t)\, dt \tag{5}$$

The capacity factor (CF) instead, describes the power generation capacity during a given period. It is determined by evaluating the total energy produced over a specified period against the maximum energy that could have been extracted by the rotor if it had operated at its rated power output.

$$CF = \frac{E(T)}{E_{rated}} = \frac{1}{T}\int_0^T \frac{P(t)\, dt}{P_{rated}} \tag{6}$$

Such parameters help in the decision of wind turbine selection and on the localization of the farm. Using statistical methods to evaluate the probability of climate conditions occurrence along a year, it is also possible to forecast the expected annual energy produced (AEP) by a specific farm, in a location set and with a chosen design.

2.2. Wake effect theory

The extrapolation of kinetic energy by the upstream rotors and the presence of element in the airflow affect the stream of air creating turbulences and wind speed reduction. Aerodynamic interactions that are present in a wind farm modify the energy production due to the variation of wind velocities and create a potential increase of mechanical loading in consequence of the generation of turbulences or vorticity. This phenomenon is known as Wake Effect.

The recover of unmodified airstream, the propagation of the wake and its expansion depend by the wind speed, by the distance from the rotor, atmospheric stability and ambient turbulence. As the flow in the wake propagate downstream, it starts to recover its freestream properties by the mixing with unaffected flow around the wake boundaries. However, the recovery is not linear and the wake could be divided in two regions: the near and the far wake. The length of near wake is dependent by various parameter like ambient turbulence and usually the estimation is 4÷5 rotor diameter length [27]. This region behavior is mainly dependent and explain the single turbine energy extraction while far wake region studies are more focused in wind farm interactions.

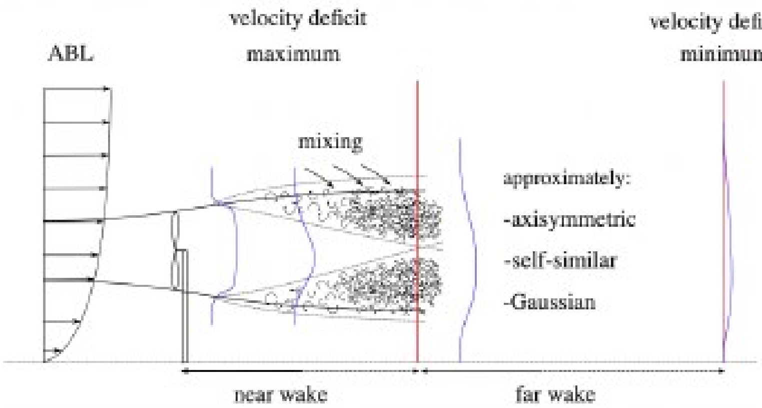

Figure 4: Velocity profile in the wake of a wind turbine for near and far wake regions by the ABL. Retrieved from [57]

Due to the complexity of the physic in the wake effect, an accurate calculation of this phenomena is computational expensive. Furthermore, the interaction of many parameters and their high variability in time make numerical models not suitable for such evaluation as takes long time to evaluate an output response. For this reason, the use of simpler analytical calculations is widely spread in this field. Several models have been implemented to describe the wake effect and minimize power losses. In general, they can be classified as either explicit or implicit.

In explicit method the wake is explicitly modeled, tracked though the simulation, and directly evaluate the dynamic interaction with the surrounding flow. This method is computationally expensive but very accurate in capturing wake interactions and flow phenomena [28]. Implicit calculations are instead most developed, based on approximated Navier-Stoked or vorticity transportation equation resulting in a more efficient computation calculation with a lower accuracy in detailed wake dynamics [29].

2.2.1. Gaussian velocity deficit

Numerical method and wind tunnel measurement have showed that a Gaussian distribution it can be used as approximation of the representation of speed deficit profile along the far wake region [30]. The use of such model represents a good compromise in computational cost and accuracy reached comparing to other simpler wake models like Jensen characterized by a top-hat shape [31] [32]. In their analytical wake model, mass and momentum conservation is applied to a control volume around one turbine where a self-similar Gaussian profile is assumed for the velocity deficit to derive the following equation for the normalized velocity deficit [33]:

$$\frac{\Delta U}{U_\infty} = \left(1 - \sqrt{1 - \frac{C_T}{8(k*x/D+\varepsilon)^2}}\right) * exp\left(-\frac{1}{2(k*x/D+\varepsilon)^2}\left[\left(\frac{z-z_h}{D}\right)^2 + \left(\frac{y}{D}\right)^2\right]\right) \tag{7}$$

where D is the rotor diameter; x, y, and z are streamwise, spanwise, and vertical coordinates, respectively, and z_h is the hub height level.

k denotes the wake growth rate which is a function of thrust coefficient and local streamwise turbulence intensity [34] for a range of conditions ($0.065 < I < 0.15$) calculated as:

$$k = 0.3837\, I + 0.003678\sqrt{\beta} \tag{8}$$

They also proposed the following expression for ε:

$$\varepsilon = 0.2\sqrt{\beta} \tag{9}$$

where β is a function of C_T and can be expressed as:

$$\beta = \frac{1}{2} * \frac{1+\sqrt{1-C_T}}{\sqrt{1-C_T}}, \quad C_T < 0.9 \tag{10}$$

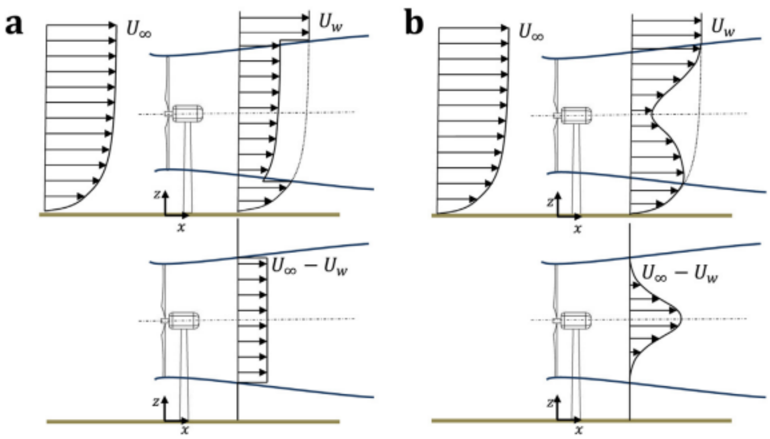

Figure 5: V*ertical profiles of the mean velocity (top) and velocity deficit (bottom) downstream of a wind turbine for (a) a top-hat and (b) a Gaussian distribution for the velocity deficit. Retrieved from [34]*

2.2.2. Crespo-Hernandez Wake Turbulence Model

As already discussed, turbulence intensity can highly affect a power production in a wind farm. For this, a Crespo and Hernandez's empirical equation [35] is used to determine the added turbulence induced by a wind turbine. For the far wake the following expression is used:

$$I_+ = 0.73a^{0.8325}{I_0}^{0.0325}\left(\frac{x}{D}\right)^{-0.32} \tag{11}$$

where a is the induction factor and the range of parameters valid for this equation are $5 < x/D < 15$; $0.07 < I_\alpha < 0.14$; and $0.1 < a < 0.4$. The equation has been tuned through comparisons to high-fidelity computational fluid dynamics (CFD) simulations [36] and several field studies [37] [38] [39] to accurately calculate the added turbulence due to upstream turbine j:

$$I_{+,j} = A_{overlap}\left(0.5{a_j}^{0.8}{I_0}^{0.1}\left(\frac{x}{D_j}\right)^{-0.32}\right) \tag{12}$$

where D_j denotes the diameter of turbine j, and $A_{overlap}$ refers to the fraction of the rotor-swept area of the downstream turbine that intersects with the cross-sectional area of the wake from the upstream turbine. Evaluating the added turbulence of each turbine it is so possible to find the total intensity with the superimposing of the initial ambient turbulency I_0:

$$I = \sqrt{\sum I_{+,j}^{\ 2} - I_0^{\ 2}} \tag{13}$$

2.2.3. Sum of Squares Freestream Superposition

When considering a wind farm that choose a design that minimize the energy losses distancing the turbine, the creation of various combinate wake effect is inevitable due to space limitation and wind resources. It is so needed a superposition method to describe the influence of the various wake effect and calculate the wind velocities at given points. As the air flows in wind turbines operation condition are characterized by high-Reynolds number, their behavior are nonlinear. For this purpose, among the many methods available, the Sum of Squares Freestream Superposition (SOSFS) method it has been used. SOSFS evaluate the velocity in a point considering the energy deficit of all the wake effect present decreasing the value of the unaffected airflow by the sum of square of the velocities:

$$U_1 = U_\infty - \sqrt{\sum_i U_\infty^{\ 2} - U_i^{\ 2}} \tag{14}$$

As showed from wind tunnel measurement, such method is able to represent with good accuracy the integration of more wake effects [40]. Moreover, SOSFS find an equilibrium velocity of the wind speed when calculated for several aerogenerator in a row.

2.2.4. Floris: a tool for wind farm performance prediction

The National Renewable Energy Laboratory (NREL) and the University of TUDelft have developed FLORIS: FLOw Redirection and Induction in Steady State FLOw Redirection and Induction in Steady State (FLORIS). FLORIS is a computationally inexpensive tool that models the steady-state characteristics of wind farm wakes. FLORIS has become widely adopted in the wind energy industry due to its accuracy and efficiency. It uses a combination of engineering models and computational fluid dynamics (CFD) techniques to simulate the behavior of wind turbines and their wakes within a wind farm. FLORIS accounts for complex aerodynamic interactions between turbines, such as wake redirection and induction effects, which significantly impact power production and turbine efficiency.

One of FLORIS's key strengths is its ability to model individual turbines and entire wind farms. This allows for optimization of turbine placement and operational strategies to maximize energy generation while minimizing wake losses. By accurately forecasting wake interactions and turbine performance, FLORIS empowers wind farm developers and operators to make informed decisions about turbine positioning, layout design, and control strategies.
FLORIS has been extensively validated through comparisons with field data and other simulation tools [41]. This demonstrates its reliability and precision in predicting wind farm performance under diverse atmospheric conditions and turbine configurations [42]. Its user-friendly interface and open-source codebase make FLORIS accessible to researchers, engineers, and developers worldwide, facilitating collaboration and innovation in the wind energy industry.

This software has different selections of wake model calculation allowing the user to a higher flexibility. In this study, FLORIS will be utilized to model the wake interactions and performance of the proposed wind farm layout with the following calculations characteristic:

Table 2: Features of wake model implemented in Floris

Gaussian Wake Model		
Velocity Deficit Model	**Wake-Added Turbulence Model**	**Wake Superposition Module**
Gaussian	Crespo-Hernandez	Sum of Squares Superposition Model

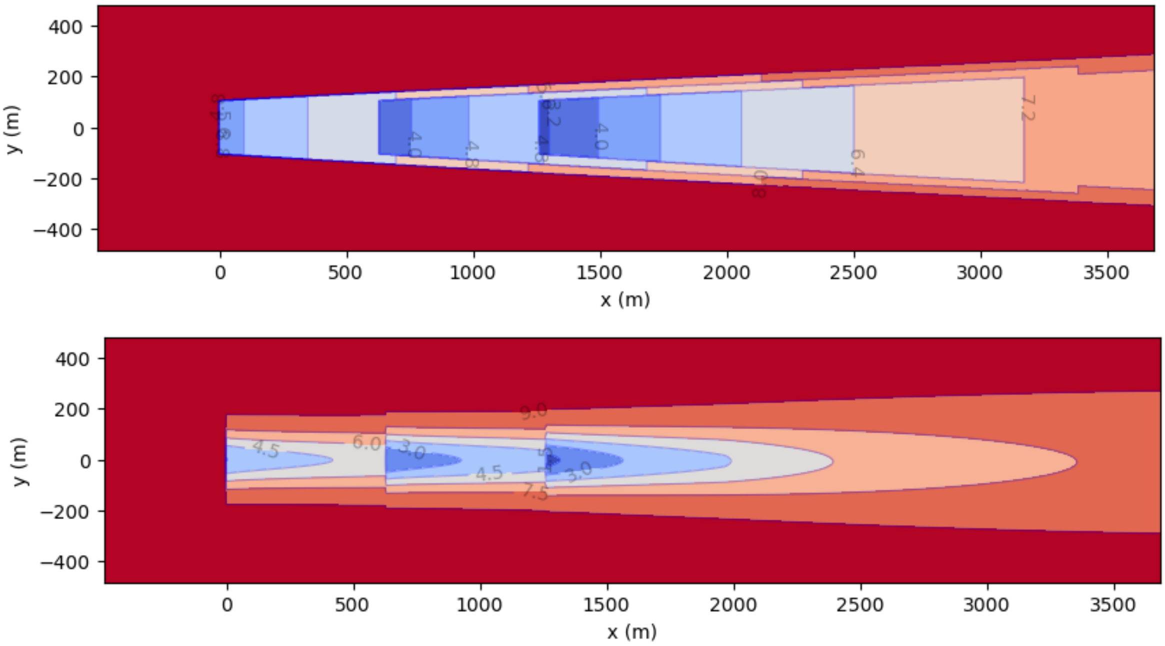

Figure 6: Velocity deficit representation using Jensen model and Gaussian distribution implemented in FLORIS

The present study employs the open software FLORIS simulating different turbine configurations and operational scenarios. The aim is to maximize energy production while minimizing wake losses along preventive maintenance operations evaluating the best set of turbines to operate in certain conditions.

2.3. Operations and maintenances in offshore wind farm

Operation and Maintenance (O&M) costs make up a significant portion, around 27%, of the Levelized Cost of Electricity (LCOE) for offshore wind farms (OWFs) [18]. This highlights the importance of reducing these costs to improve the viability and profitability of new offshore wind projects.

The maintenance costs for OWFs are much higher compared to onshore wind installations due to the complexity of the equipment, specialized personnel, and higher component costs. Accessing offshore wind farms requires specialized vessels to transport equipment in various sea conditions, and the need for trained technicians further increases the costs.

Maintaining offshore wind turbines (OWTs) presents significant challenges. The remote location of OWFs from ports or shores reduces accessibility and increases downtime. Procuring or leasing maintenance fleets and hiring additional technicians also incurs substantial costs. Furthermore, the harsh offshore environment and the need for specialized equipment and vessels contribute to the higher maintenance expenses.

Even with the specialized equipment needed, maintaining offshore wind turbines costs more than similar tasks on land. The tough offshore conditions, like higher winds, waves, and vibrations, cause more components to fail. In recent years, the shift to larger offshore turbines to boost power production requires even bigger and more specialized gear for maintenance and repairs offshore.

The maintenance work depends on the weather conditions and the technical availability of the Crew Transport Vessel (CTV) and its operators. A regular CTV (Old CTV) can operate safely only in wave heights up to 1.5 m and wind speeds up to 15 $m/_s$ at 10 m height, with a capacity limit of 12 to 16 crew members. However, novel CTVs are now available that can safely navigate in wave heights up to 2 m and have a larger capacity, allowing for more turbines to be maintained per day [16] [17].

An in-house data base at WavEC – Offshore Renewables suggest that 2 to 3 technicians are required to perform preventive maintenance on a single turbine. It is also suggested that each turbine undergoes on 80 hours of preventive maintenance per year made by shifts of 10 hours duration [2], counting 2000 total hours of maintenance to reach in each year to be able to maintain all the turbines.

Maintaining an offshore wind farm is vital to minimizing downtime and keeping electricity generation costs low. Regular inspections and repairs are necessary to identify and fix any issues that could prevent the turbines from operating efficiently. However, too many unnecessary inspections can also increase costs. Operators must strike a balance to ensure the wind farm runs smoothly while controlling expenses. Carefully monitoring the turbines and addressing problems promptly is key to maximizing the farm's productivity and profitability.

Different operation and maintenance (O&M) strategies are implemented based on various operational factors and decision-making criteria. The goal is to strike a balance between ensuring reliable energy production and managing the costs associated with maintaining the wind farm over its lifetime [42].
The main distinction of O&Ms approaches are:

- Corrective (Reactive) maintenance effectuate a maintenance only when a failure occurs, this method has low frequency of inspections, but it could lead to main component damage and high downtime losses caused by impossibility of fast intervention.

- Proactive maintenance is a more advanced approach that aim to reduce the failure of components reducing the cost new equipment. This lead to higher operations costs caused by frequents interventions and tests on the systems. This approach could be implemented based on weather conditions, on states of components (Predictive or Condition-based), or on time considerations (Preventive).

- Opportunistic maintenance uses the opportunity to effectuate a preventive maintenance while a corrective operation needs to be carried out due to a failure.

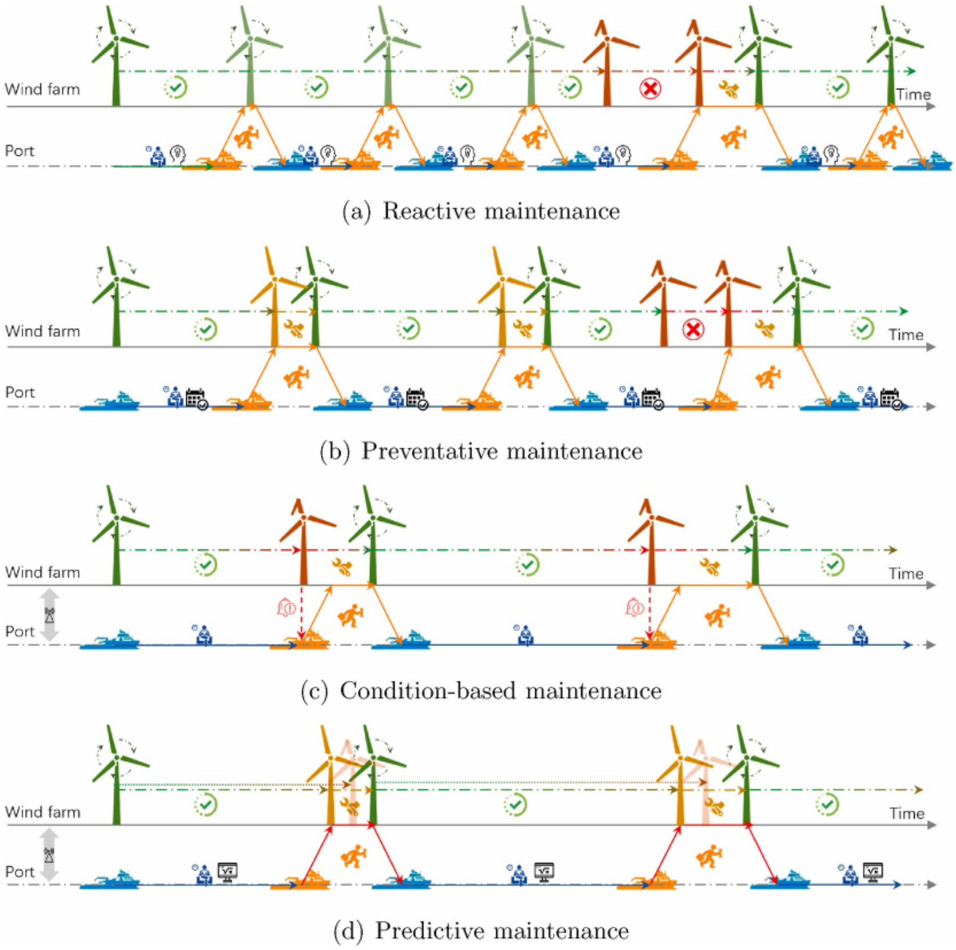

Figure 7: Diagrams of the different maintenance strategies. Green, yellow and red turbines denote normal operation, deactivation due to maintenance and deactivation due to failure, respectively. Blue and orange vessels denote waiting maintenance vessel and vessel in operation, respectively. Retrieved from [42].

Preventive maintenance is an important part of O&M to avoid unexpected component failure that leads to elevated downtime energy losses that increases, with the impossibility or slow reactance of organization, to repair the damaged part. To reduce the possibility of unforeseen lack of availability of energetic production, these activities are generally planned along the year and are repeated for the entire lifetime of the OWF. The evaluation of O&M needs to consider various parameters such costs, technical availability, failure probabilities and energy generation. In this work a preventive strategy for O&M it has been analyzed.

2.4. Weather forecast

Many challenges in predicting weather accurately arise from the complex nature of the climate system. Accurate weather forecasting is crucial for offshore wind farm workers, especially when planning installations operations or maintenance activities. Furthermore, as OWF are often located far from weather observation stations and it can be difficult to obtain precise or high-resolution weather forecasts without a dedicated analysis.

Meteorological predictions are inherently uncertain due to the complex nature of the atmosphere and the various interacting factors that influence weather patterns. The atmosphere's intricate dynamics involve numerous variables, with uncertainties arising from limitations in the quality and coverage of initial data. The difficulty of the numerical model to accurately represent the complex climate system and its many interactions leads to the "butterfly effect," where small variations in initial conditions can significantly affect long-term forecasts.

In the lasts decades important improvement have been made in forecasting weather conditions. Improved models, data-assimilation systems, a higher number of satellite observations and increased computer power have upgraded the predictability of the weather in all the lead time forecasts and increased its accuracy.

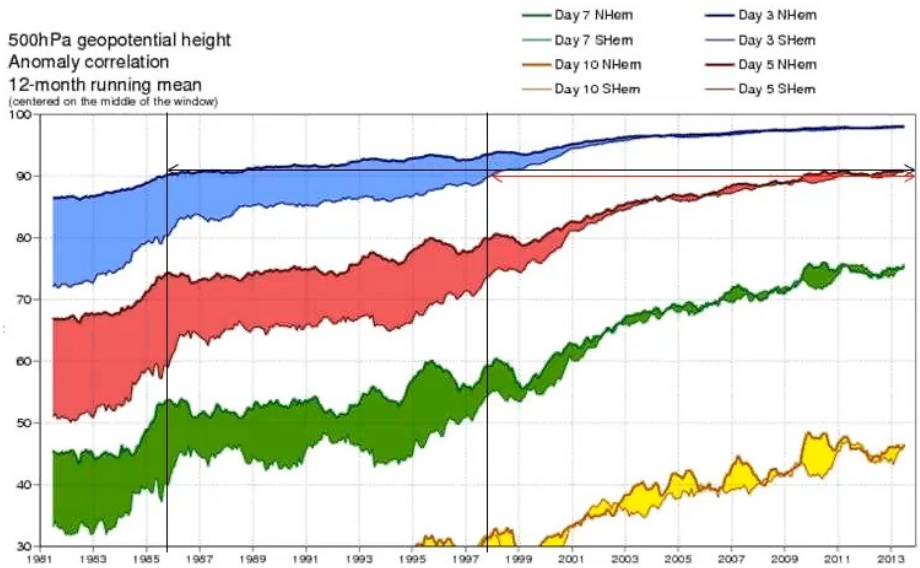

Figure 8: Predictability improvement for Z500 Retrieved from [43]

Traditional weather forecasting systems, like those used by ECMWF or IPMA, employ an Integrated Forecasting System (IFS) that combines various components in different ways. For medium-range forecasts, an atmospheric model is run at different resolutions depending on the forecast length: HRES to investigate detail in the forecast during the first 10 days and an ensemble (ENS) model to investigate detail and uncertainty in the forecast during the first 10days or 15days. These forecasts are then coupled with models for the ocean (ECWAM), waves (NEMO), and land surface (HTESSEL), which interact using a 4D-VAR assimilation analysis [44].

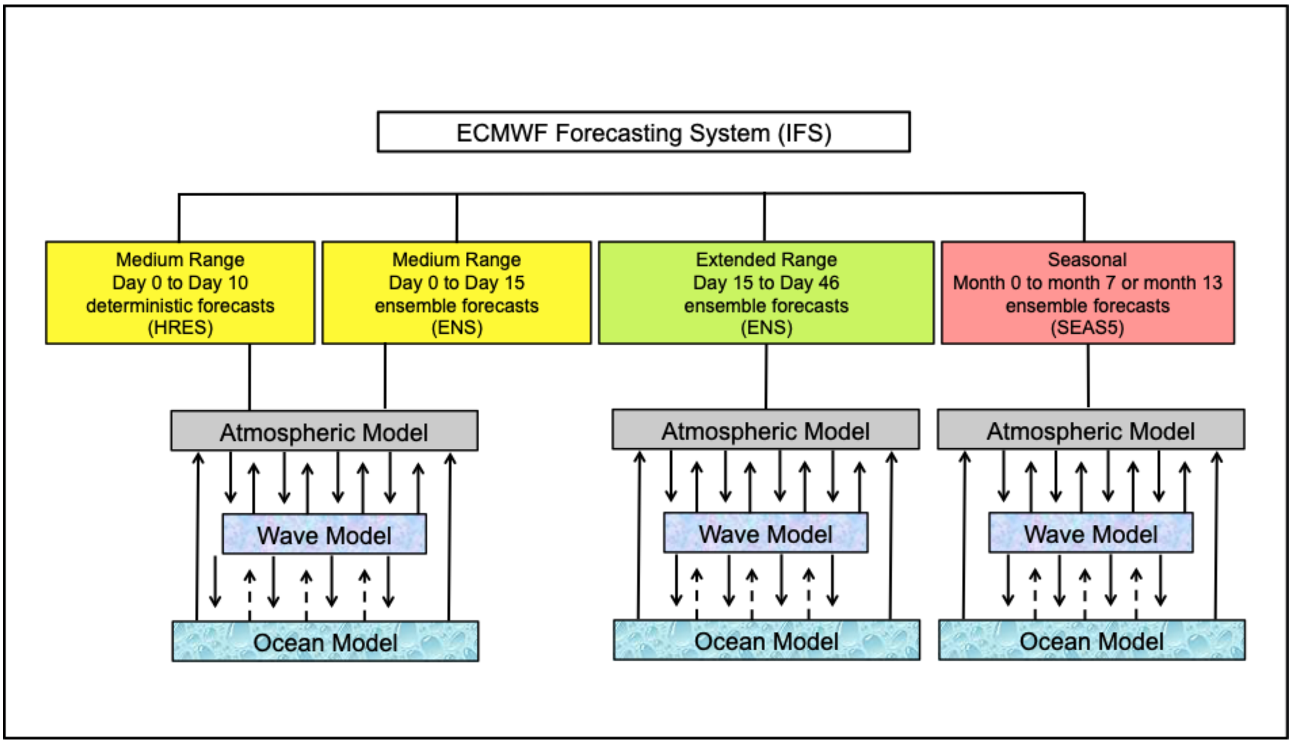

Figure 9: Structure of ECMWF IFS Retrieved from [44]

The outputs of the atmospheric analysis include a best estimate for a single analysis for weather parameters from the HRES model and a range of potential future weather states shown by the ENS analysis.

The HRSE has horizontal resolution of 9 Km and vertical resolution of 137 model levels, it is run four time a day, twice to evaluate a short range forecast analysis with 3.75 day forecast (T+0 to T+90) and twice to estimate a medium range forecast with 10 days forecast (T+0 to T+240).

ENS analysis is obtained by running the numerical prediction model multiple times, introducing small changes to the initial conditions to account for initial uncertainties. The ensemble of forecasts provides a range of future scenarios consistent with our knowledge of the initial state and model capability. It is structured as the HRES with the exception that has longer forecast range time (6 days and 15 days) It is used an unperturbed control member (CTRL) corresponds to the HRES to evaluate the uncertainty related to the best estimate. The evaluation of different possibles evolutions of a weather parameter shows the uncertainties related to the output obtained and its frequency of occurrences demonstrate the probabilities that the condition has to realize. If the forecasts are coherent (small spread) the atmosphere is in a more predictable state than if the forecasts diverge (large spread).

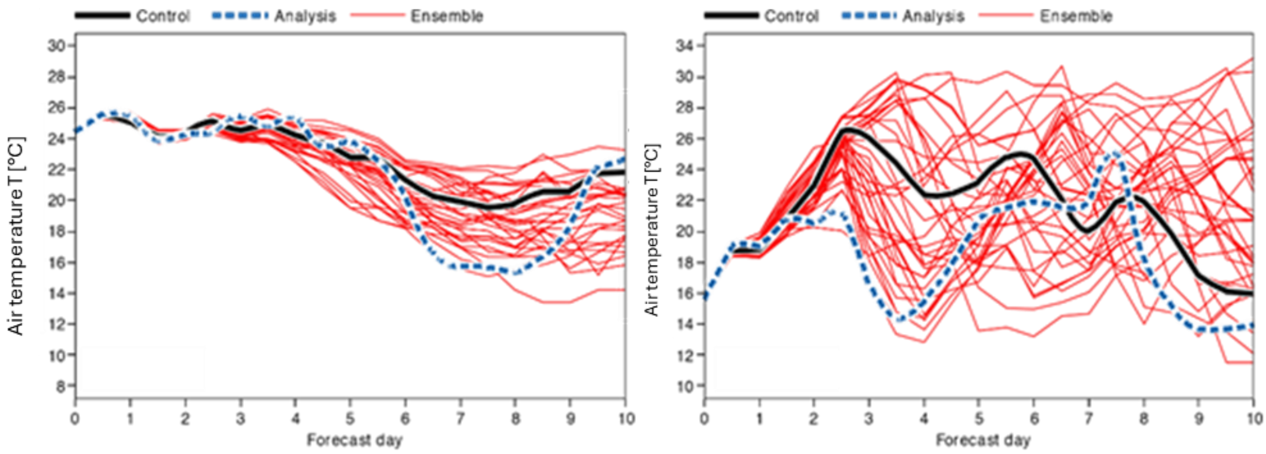

Figure 10: Ensemble forecast for temperature of London in 26th June 1995 and 26th June 1994 Retrieved from [45]

Higher lead time of forecast gives more time to the perturbated parameters to evolve and spread in a wider range of possibilities. Increased uncertainties arise when attempting to forecast weather conditions several days in advance. For a short time prevision the HRES forecast shows to be the best estimate of weather forecast except for few days. Increasing the lead time over 3÷4 days in forecasting weather conditions the anomalies encountered in the HRES model rise and the ENS mean start to become more reliable for any parameters.

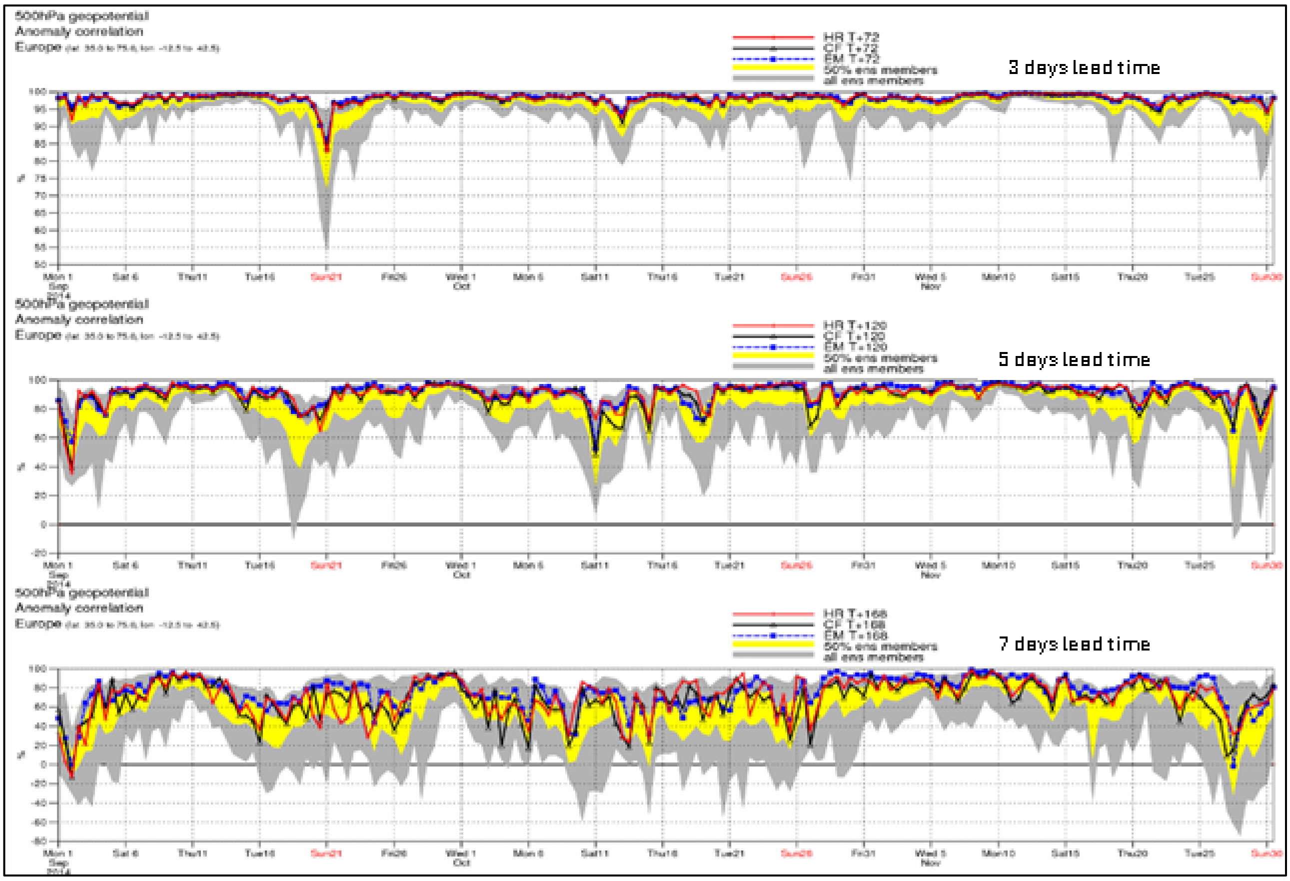

Figure 11: Anomalies correlation of London in 26th June 1995 and 26th June 1994 Retrieved from [45]

The need of precise weather data for OWF are fundamental to decision taking in O&M due to operability limit of the CTV needed to sustain such procedure. An inaccurate forecast could predict a workable day for the vessel that might not be possible leading to an economic loss for default of mission conclusion or in a possible opportunity of maintenance lost. Using data-driven approaches could reduce the inaccuracies of traditional forecasts with low geographical resolution and the use of economic metrics it could improve the economics losses [46].

The months taken in consideration for preventive O&M in this study are from April to October for a range of time that goes from 8 am to 18 pm as, the weathers conditions in this period, are more suitable to such activities.

2.5. Data-driven method – Fuzzy C Means Theory

Wind speed and wind direction can be quite unpredictable and fluctuating over time and space. This makes it challenging to forecast accurately the weather and make informed decisions for offshore wind farms. Many researchers have explored using data-driven algorithms to enhance the accuracy of short-term and medium-term weather predictions, aiming to reduce uncertainties in these forecasts, as seen in studies like [47] [48] [49].

Different data-driven data exist, and they differentiate from method of calculation, output given and methodology applied. Mainly the machine learning methods (ML) can be divided in two main categories: Supervised Learning Machine and Unsupervised Learning Machine. Supervised methods are trained on labeled data and try to learn from the input and output data given the relation present. Their object is to correctly predict outputs for new data that have not been seen in the training of the model.
Unsupervised algorithms use datasets that are not labeled and have no examples of input-output relations. In these cases, the model tries to find pattern or intrinsic correlation hidden between the data. For these algorithms, the validations of the models are more difficult and less straightforward than the previous type.
A further subdivision of these models could be done regarding the aim of the algorithms. Cluster algorithms regroup data under families of similar characteristic while regressive models try to predict a specific continuous value from a given input.

This paper introduces a straightforward data-driven model called Fuzzy C Mean (FCM) to evaluate wind direction. This model it has been chosen for its high simplicity, the low amount of data needed and its reduced calculation time. The goal is to better understand the patterns of incoming winds direction at the location of the analyzed offshore wind farm and optimize the precision of the forecasts.

The Fuzzy C Mean algorithm is an unsupervised clustering method that aims to identify patterns in a dataset by dividing it into groups or clusters. It is based on geometrical distance between the data and the clusters evaluated introducing a similarity weight derived by the Fuzzy Logic [50]. Unlike the K-means algorithm, which assigns each data point to a single distinct cluster, FCM allows data points to have a probability of belonging to multiple clusters. This means that each data point is assigned a membership value between 0 and 1, representing its degree of similarity to each cluster. This approach provides a more nuanced understanding of the dataset, as elements can be associated with multiple clusters simultaneously.

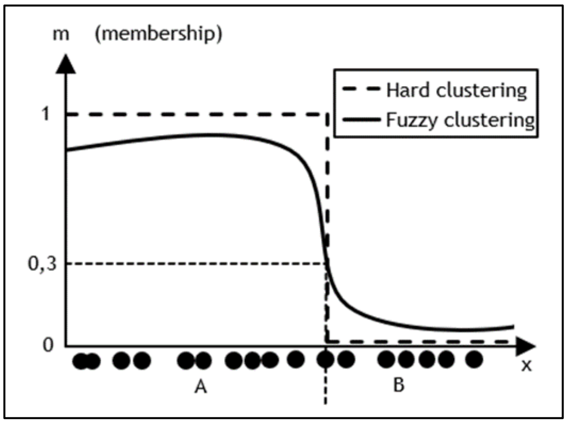

Figure 12: K means and Fuzzy K means membership

A theoretical demonstration could be found in [51], where a short explanation is provided. To operate the Fuzzy C-Means (FCM) algorithm, it requires selecting a threshold value ε as a limit for the objective function. It also requires deciding the number of partitions or clusters present in the dataset beforehand. The optimal number of partitions can be determined using different methods, such as the 'Elbow method,' which aims to assess the reduction in the Within-Cluster Sum of Squares (WCSS), also known as inertia, across various partition numbers. This analysis helps identify the minimum number of clusters that effectively minimize variance. Other useful indices, like Silhouette analysis, Davies-Bouldin Index (DBI), and Calinski-Harabasz Index (CHI), can provide insights into the similarities between the partitions and the data, leading to a better understanding of the dataset's partition.

After randomly initializing the membership matrix $U = \{u_{i,j}\}$ calculate center of mass of the Centroids:

$$c_j = \frac{\sum_{i=1}^{N} u_{i,j}^{m} x_i}{\sum_{i=1}^{N} u_{i,j}^{m}} \tag{15}$$

having N as number of data points, C as number of clusters, m as the fuzzification exponent, x_i the data point i and c_i the centroid value j. Updating than the membership matrix:

$$u_{i,j} = \frac{1}{\sum_{k=1}^{C}\left(\frac{\|x_i - c_j\|^2}{\|x_i - c_k\|^2}\right)^{\frac{2}{m-1}}} \tag{16}$$

The algorithm is iterated till the objective function J is lower than the threshold ε set:

$$J = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{i,j}^{m} \bullet \| x_i - c_i \|^2 < \varepsilon \tag{17}$$

Following this scheme the FCM divide the data points in partition evaluating the center of mass of the clusters used to characterize the patterns found. The centroids calculated of each clusters represent the geometrics means values of the partitions.

3. Case Study

This chapter presents the case study analyzed in all its details and structure implementations. Firstly, an introduction of the OWF taken into consideration is presented, then are deeply illustrated all the data used and their utilization strategies to end in the implementation of the methodology methods.

3.1. Wind farm characteristics

The presented case study closely matches the ones described in previous research [2] [1]. It analyzes an offshore wind farm located off the coast of Viana do Castelo in Portugal that actually, has 3 turbines of 8 MW power installed [52]. The wind farm under investigation for this work has been implemented with 25 IEA 15 MW turbines arranged in a square layout, with each turbine spaced eight diameters apart. Technical characteristic of the aerogenerators considered with the features of FLORIS library used are showed in the Table 3 and 4.

Table 3 and 4: Wind Turbine Characteristics used in Floris and Wind Farm Characteristics and Floris features used

Wind Turbine Parameters	
Model	IEA 15 MW
Cut-in wind speed	3 $[m/s]$
Rated wind speed	10.59 $[m/s]$
Cut-out wind speed	25 $[m/s]$
Hub height	150 m
Rotor diameter	240 m

Wind Farm Modelling & Floris features	
Location	41º40’ N ,09º 03’ W
Farm Layout	5x5 Square
Spacing	8 D
Wake deficit	Gaussian
Wake Combination	Sum-Least-Squares
Turbulence model	Crespo-Hernandez

This work aims to demonstrate how the weather forecasts imprecisions can impact the annual energy production (AEP) and preventive O&M planning analysis. Specially, this study concentrates in the downtime power losses that are created shutting down a group of turbines for maintenance activities optimizing them selecting the optimal set of turbines that have the lowest production due to wake effect present in the farm. To minimize uncertainties in weather forecasts, also it has been explored the potential of using a simple data-driven method, such as Fuzzy C-Means, to assess the wind direction parameter.

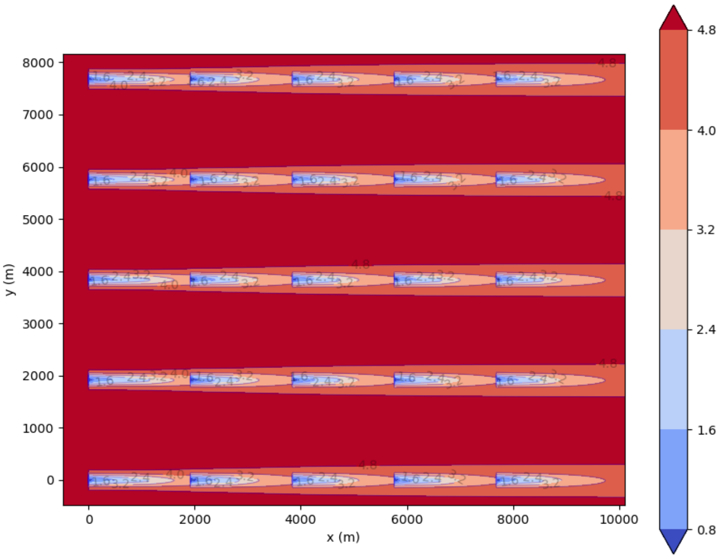

Figure 13: OWF wake effect phenomena implemented from Floris $u_{speed} = 5\frac{m}{s}$; $u_{dir} = 270°$

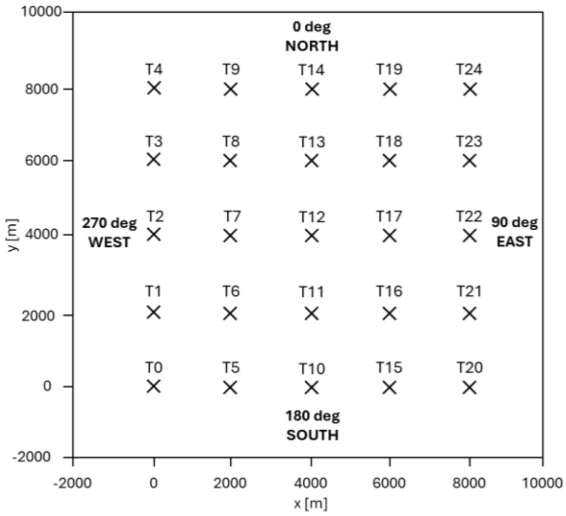

Figure 14: OWF 5X5 squared design

Figure 15: OWF location at around 20 km from Viana do Castelo, Portugal

3.2. Weather data source

The importance of accurate weather data is fundamental for specific location of OWF where there are low observation of weather condition and where a high resolution of weather behavior is required. For this work, two main sources were considered for this purpose that show the weather conditions expected and occurred along a period that goes from September 1, 2021 to December 31, 2023.

Firstly, a forecast weather data with a specific focus on the location it has been used to simulate the available estimation of weather condition when a strategy needs to be chosen. Such data, containing uncertainties, are usually the only available information that an operator of OWF possess to foresee the future weather condition and so, to estimate energetic plant productions. For this reason, weather data from the Portuguese Institute of the Sea and Atmosphere (IPMA) have been requested and obtained. Two different model of ECMWF have been received to understand and to interpret the estimation of future climate conditions. An hourly forecast from the deterministic model (HRES) with atmosphere resolution and sea resolution of 0.1° and 0.123° and an hourly forecast from the Ensemble model (ENS) with atmosphere resolution and sea resolution of 0.2°. The parameters, updated each day from IPMA forecast are the following:

Table 4: IPMA weather forecast HRES model parameters

Code	Weather parameter	Unit
msl	mean sea level pressure	*hPa*
prec	Precipitation, accumulated	*mm*
SWH	sea wave height	*m*
pp1d	peak period	*s*
mwd	mean wave direction	*degree*
mwp	mean wave period	*s*
U10	wind speed at 10 meter height	*m/s*
U10d	wind direction at 10 meter height	*degree*

Table 5: IPMA weather forecast ENS model parameters

Code	Weather parameter	Unit
per10SWH	10th Percentile of sea wave height	*m*
Per50SWH	50th Percentile of sea wave height	*m*
Per90SWH	90th Percentile of sea wave height	*m*
per10prec	10th Percentile of precipitation	*mm*
Per50prec	50th Percentile of precipitation	*mm*
Per90prec	90th Percentile of precipitation	*mm*

In particular, the data from HRES model have been used for energy calculation, like wind speed and direction. For strategic decisions have been considered mainly the sea wave height available from the ENS model, considering a probable scenario and a more conservative scenario represented by the 50th and the 90th percentile value of the dataset.

Secondly, a comparison of forecasted data with real occurred weather conditions is needed to understand and assess the mismatch of results obtained from the expected foresees generated from forecasts uncertainties. Due to the lack of accessibility to weather observation in the specific location under study, a hindcast dataset from ERA5 was considered for this purpose. Two different time periods have been used from ERA5 database. One that serve the purpose explained above for data validation and aligns with the IPMA forecast used, from September 1, 2021 to December 31, 2023. The other time period has been used for the training of the data-driven method implemented for improve the estimation of the wind direction calculated from the IPMA weather forecast. For this task, the condition of the wind direction and wins speed it has been taken from a period of 20 years that goes from the year 2000 to the 2020.

The hindcast process involves the application of sophisticated models, historical data observation and computational simulations to recreate and evaluate how weather patterns or climatic events behaved in the past. A reanalysis combines model data with observations taken from across the world into a globally complete and consistent dataset using the laws of physics. The principle of integrate newly available observation to the method used by numerical weather prediction to produce a best new estimate of the state of the atmosphere is called data assimilation. Data assimilations are used in reanalysis models with a reduced resolution to allow for the provision of a dataset spanning back several decades. As a hindcast does not have the constraint of issuing timely forecasts, the use of more time to collect observations allow the ingestion of improved versions of the original observations, which benefit the quality of the reanalysis product [53].

A comparison between the two data sources for the time period of 1st September 2021 to 31st December 2023 has been computed to evaluate the discrepancies. The IPMA forecast data taken at hour 0 have been considered for such case, as the firsts outputs of a forecast model are considered very precise and similar to in time observational data due to the low influence of time parameter. The comparison revealed that the forecasted data at a 0-hour lead time has a strong but not perfect correlation for wind speed and significant wave height, though modest resolution for the wind direction, as showed below.

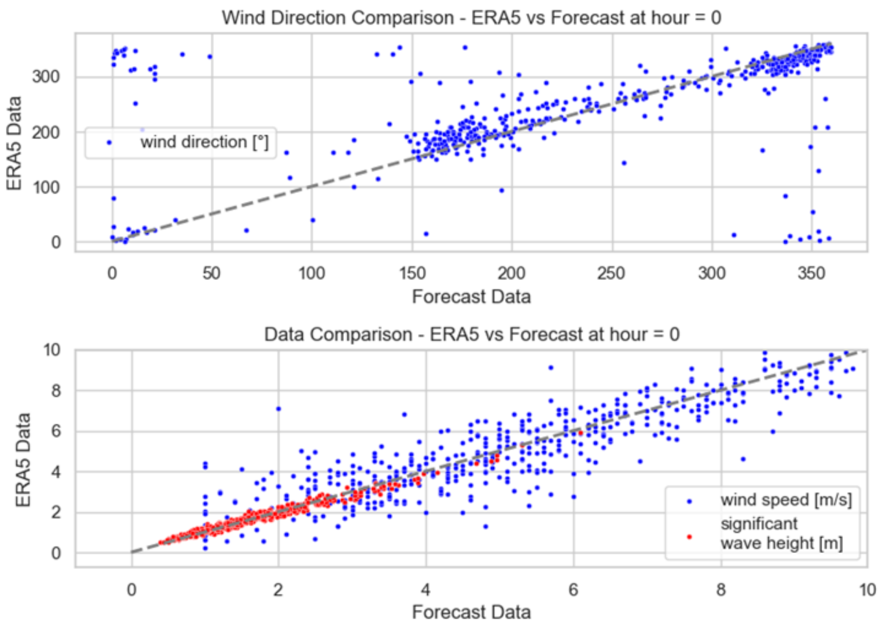

Figure 16: Comparison of weather dataset for months from April to October of ERA5 with IPMA forecast at hour zero lead time

These comparisons highlighted the different implementation methods employed by the data sources, with their intrinsic uncertainties. The minor discrepancies mentioned above can result in biased errors in subsequent analyses, representing the treat of the ERA5 hindcast data as actual observations as the most significant assumption in the analysis.

Another analysis of the used ERA5 weather dataset is conducted confronting the physical parameters in the data obtained in the first 20 years of the XXI century with the time period from 2021 to 2023. The aim is to evaluate the trend and the similarities of the weather parameters for the two different time period taken into account, in order to visualize hypothetical significant pattern variations.

In the Table 6 are presented the differences in data distribution for the two datasets. The results underline a less severe weather conditions occurred in recent years respect of the oldest period having a lower mean difference along the years for SWH and U10 .

Table 6: Difference of data distribution ERA5 2000-2020 vs 2021-2023

Variable:	SWH [*m*]	U10 [*m/s*]
Mean Difference:	0.110	0.178
Std Difference:	0.007	0.044

3.3. Fuzzy C-mean application

A data-driven method it has been integrated in this study to assess the wind direction from weather forecast with historical data. The use of such machine learning it has been implemented for two different goals. Firstly, when considered, it is used in substitution of the wind direction obtained from IPMA in the AEP to evaluate the possible improvement from this implementation. Secondly, when opportune, the outputs of wind direction received from the FCM in the O&M analysis are used to select the set of turbines under maintenance, with no use in the energetic results.

The FCM algorithm has been trained on 20 years of historical climate data and tested using the weather forecasts from IPMA for the period 2021-2023. This allows the algorithm to recognize patterns and assign membership degrees to different clusters. For data preparation, the dataset is divided into monthly categories. Each category is further divided into three speed ranks based on the average wind speed calculated for different times of the day (24 hours for AEP, and 6 am to 7 pm for O&M). The optimal number of partitions for each category are determined by minimizing the inertia of each member classification, while ensuring a balance between cohesion and separation of the clusters expressed with the use of Silhouette index, CHI and DBI.

It has been found that the lower wind speed group of that goes from 3 to 6 $m/_s$, has a higher directional variation among a day respect to the medium speed class characterized by velocities from 6 to 10 $m/_s$. Also, for the high wind speed categories represented by velocities higher than 10 $m/_s$, are needed low number of partitions to define the entire dataset due to reduced presence of data and due to the similarities among the patterns. For days with representative wind speed lower than 3 $m/_s$, the FCM integration it has not been applied.

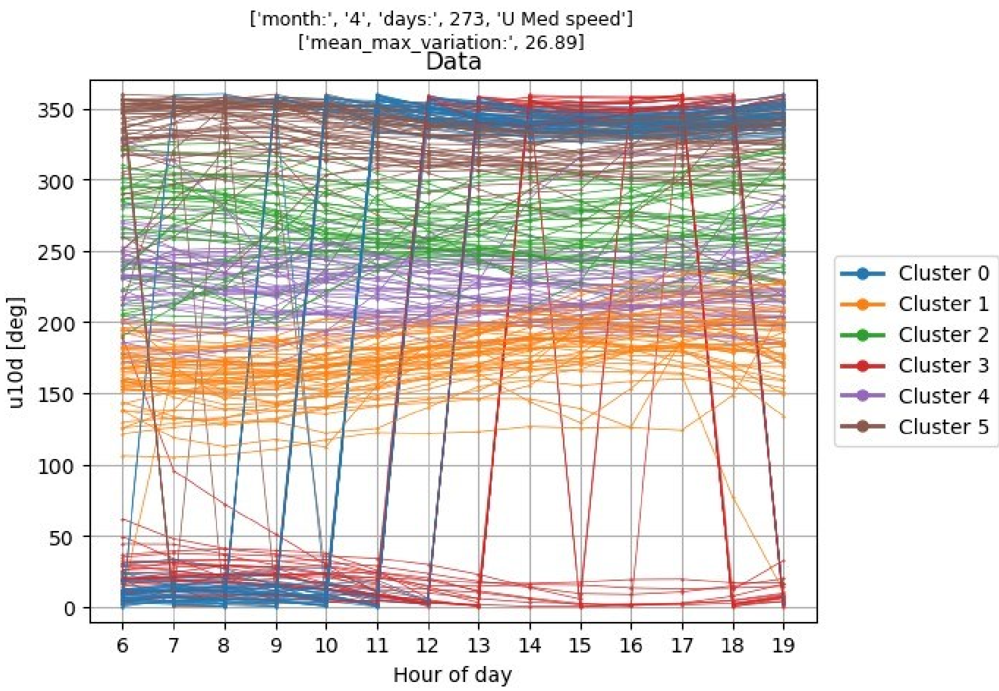

Figure 17: Training of FCM model for O&M purposes, clustering of 20 years of wind direction data, April, medium U10

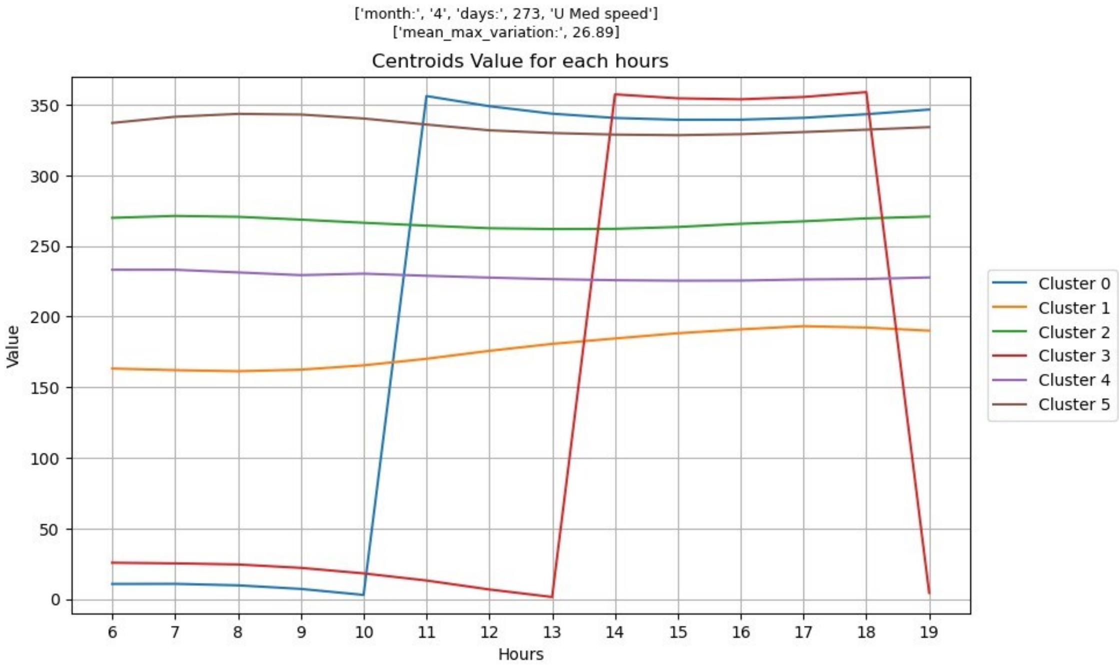

Figure 19: Centroids of the cluster obtained for O&M purposes, month of April, medium wind speed U10

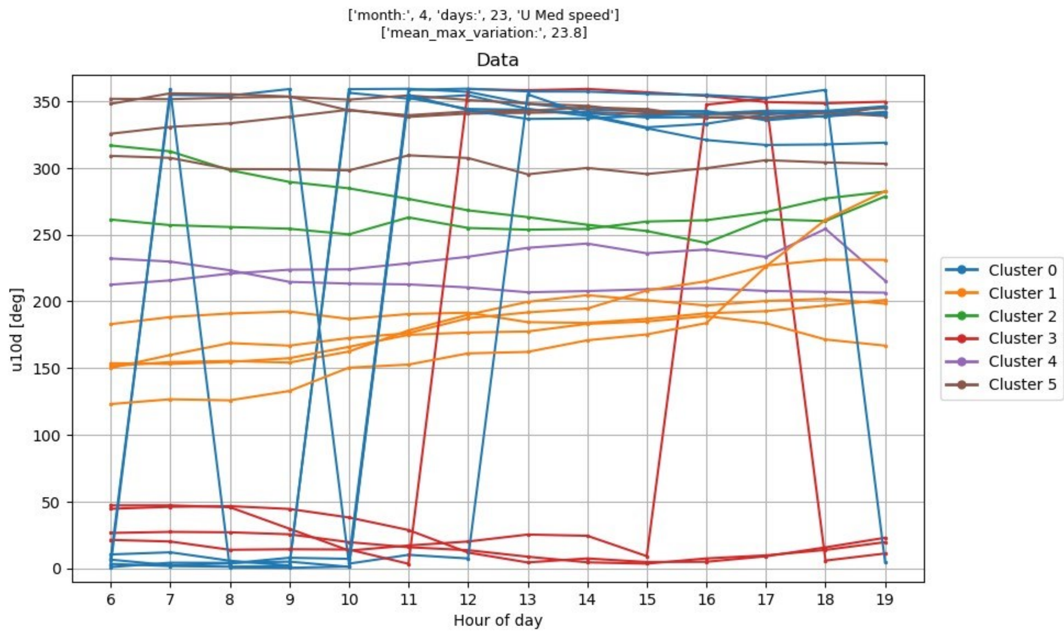

Figure 18: Test of the FCM model for O&M purposes, recognizing of weather forecast data into clusters

As the FCM is a Fuzzy algorithm it assigns a membership value to each day analyzed for each partition presents. The centroid of the cluster which has the highest value of membership received for a data, is assumed to be the most accurate representation of the wind direction along the day. To filter outlier data of rare wind direction two different methods are applied and showed below:

- Membership limit (*memb_50%*): a threshold on the maximum membership value received from the FCM it has been set to $memb_X\% = $ X. Below such parameter, the daily variation of wind direction is considered not similar enough to any partition found in the training of the model. For a similar case, the integration of the data-driven model in the analysis is not applied. Furthermore, a sensitivity analysis has been made variating the membership limit from a range to 0.3 to 0.7 with a step of 0.1 to better understand the parameter and its influence.
- Occurrence limit: set to consider a cluster consolidated and not a group of outlier data. In the training of the model, the occurrences of data considered most similar to a partition must exceed the value of 7.5%. The cluster with a lower occurrence will not be used in the analysis and the integration of the ML will not be applied for daily wind direction represented by the mentioned partitions.

If the previous criteria are respected, the centroid of the cluster membership with higher value received from FCM model is used for the AEP calculation and for the O&M selection of turbines. Otherwise, the IPMA forecast wind direction it is used instead for these purposes.

3.4. Structure of case study

3.4.1. Previous case studies

This work is a continuation of the studies previously done by WavEC Offshore Renewables showed in the thesis [2] [1]. The objective of this investigation is to enhance the efficiency of an offshore wind farm by minimizing power losses resulting from turbines inoperability for preventive maintenance activities, often referred to as downtime losses. This optimization is achieved throughout a strategic scheduling of preventive maintenance operations, with a specific emphasis on mitigating the impact of wake effects selecting the optimal combination of aerogenerators to maintain. A comparation for the evaluation of workable days available for O&M it is also presented, showing how the planning of preventive maintenance operations could affect the energy production of an OWF.

In [1] it has been analyzed the wake effect on the OWF presented in the previous chapter. The evaluation of the energy production in consequence to the non-operation of a turbine is found for each aerogenerator present in the OWF to find the minimum power losses occurred. Doing so, it is possible to find the optimal turbine to switch off for preventive maintenance operation that will decrease the power losses. Such downtime losses have been firstly calculated for each wind direction with different step resolution on a fixed wind speed of 5 $m/_S$ at a reference height of 10 m. Secondly, fixing the wind direction on a single value, the power losses have been calculated for different wind speed range from the U_{cut-in} wind speed to a value well above the U_{rated} velocity with an increment of 0.1 $m/_S$. Interesting results have been found, showing that the power losses for not maintaining the optimal turbine for short ranges of wind directions are small, as showed in the following table.

Table 7: Power differences occurred from wake effect [1]

	Wind farm's power losses	
Resolution (deg)	**Highest power difference**	**Mean power difference**
1	- 0.02 %	- 0.00 %
5	- 0.14 %	- 0.05 %
10	- 0.32 %	- 0.18 %
15	- 0.41 %	- 0.18 %

Furthermore, it has been discovered that, a variation on the wind speed magnitude, has a very limited impact on the on the wind turbine that yields the lowest farm production losses when deactivated as showed.

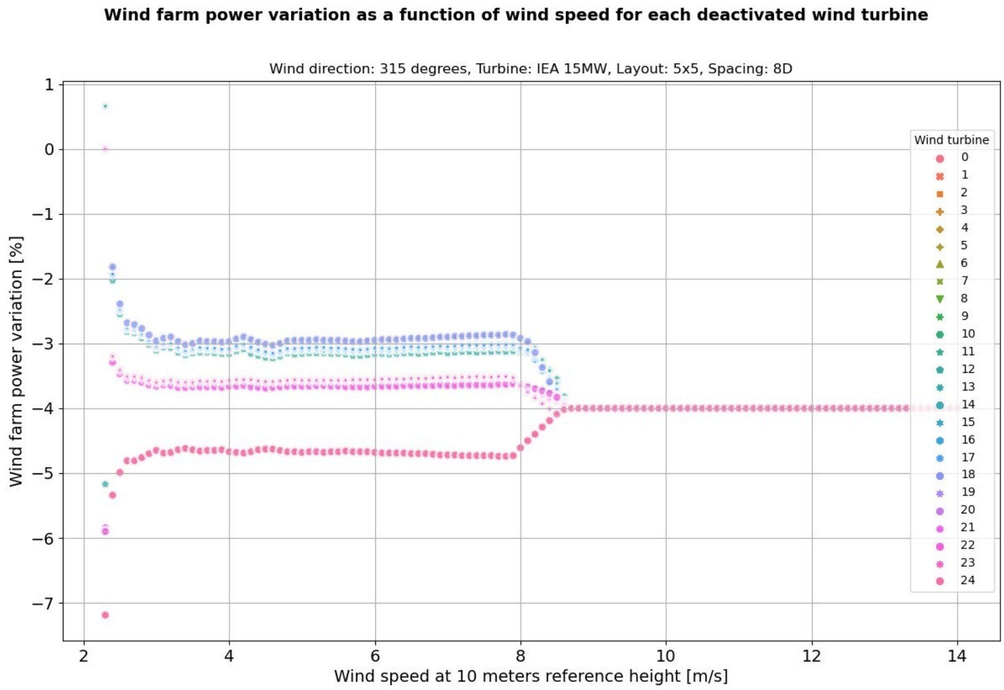

Figure 20: Wind farm power variation as function of wind speed for each deactivated turbine from [1]

Considering the geometry of the analyzed offshore wind farm (OWF), the calculation was specifically performed for 45 degrees of the entire wind direction range. A symmetrical solution was then used to derive the rankings for the remaining directions. For these two main reasons, the combinations on the optimal turbines to shut down for maintenance purposes have been calculated only on wind direction variations with a step range of 5° showed in the .

Table 8: Table 8: Rank combination of 5 turbines maintenance along a dayshift [1]

	Wind direction (degrees)				
Rank	**0°**	**5°**	**…**	**350°**	**355°**
1	[1, 6, 11, 16, 21]	[3, 16, 18, 21, 23]	...	[22, 17, 12, 7, 2]	[16, 6, 18, 13, 8]
2	[1, 6, 16, 18, 21]	[13, 16, 18, 21, 23]	...	[17, 12, 7, 2, 23]	[11, 6, 18, 13, 8]
3	[1, 11, 16, 18, 21]	[1, 3, 18, 21, 23]	...	[22, 12, 7, 2, 18]	[16, 6, 23, 18, 8]
4	[6, 11, 16, 18, 21]	[1, 3, 8, 21, 23]	...	[22, 17, 12, 2, 8]	[11, 6, 23, 13, 8]
...	...	...	...	...	...
53127	[20, 21, 22, 23, 24]	[20, 21, 22, 23, 24]	...	[0, 1, 2, 3, 24]	[5, 6, 7, 8, 9]
53128	[15, 16, 17, 18, 19]	[15, 16, 17, 18, 19]	...	[20, 21, 22, 23, 24]	[15, 16, 17, 18, 19]
53129	[5, 6, 7, 8, 9]	[5, 6, 7, 8, 9]	...	[0, 1, 2, 3, 4]	[0, 1, 2, 3, 4]
53130	[10, 11, 12, 13, 14]	[10, 11, 12, 13, 14]	...	[0, 21, 22, 23, 24]	[20, 21, 22, 23, 24]

In [2] the study continued validating the different impact of wake effects on OWF power losses by variating the wind direction. Other than prove the symmetry on the layout of the OWF selected, it shows that for specific values of wind directions, the energies gain in the selection of turbines under maintenance could be negligible.

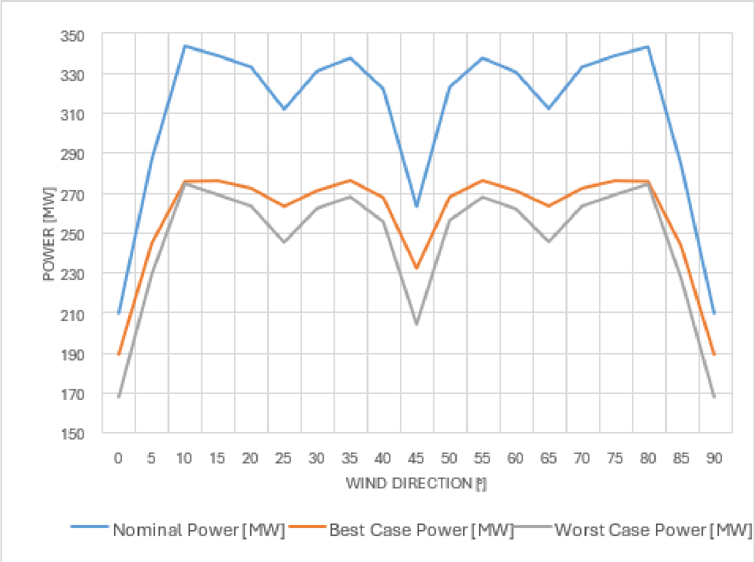

Figure 21: Comparation of power production with different direction [2]

Furthermore in [2] it has been introduced and compared different selections of criteria for the evaluation of workable days to effectuate preventive maintenances. The study is conducted for a period of time that goes from 1992 till 2012 considering only the months from April till the end of October. More specifically two different strategies have been analyzed: a *binary criterium* and a *non-binary criterium.*

- Binary criterion, determinate whether a day is available for the CTV to operate whenever the weather conditions allow it. It means that a maintenance activity is conducted for each day where the significant wave height (H_s) and the maximum wind speed magnitude (u_{lim}) is lower than the threshold imposed by the CTV limits.

- Non-binary criterium impose instead more strict limitations on maximum wind velocities to reduce the days of work where occur high power losses due to higher wind resources. Such limitation on the wind speed is flexible. It variates on different months (mth) taken into consideration, prioritizing the urgency of reaching the total annual hours planned (h_{tot}) considering the time left. The non-binary criterium, also involve an abstention from operational workdays with $u_{lim} > 3\ [{}^{m}/_{s}]$ if the cumulative hours of preventive maintenance effectuated is higher than a threshold limit (h_{lim}). Such value is selected comparing the amount of hours of maintenance achieved with h_{tot} planned for the year. The evaluation of u_{lim} in [2] it is made as follows:

$$\begin{cases} mth < 6: \quad u_{lim} = 3\left[\frac{m}{s}\right]; \quad h_{lim} = h_{tot} \\ 6 \leq mth < 8: \quad u_{lim} = 4\left[\frac{m}{s}\right]; \quad h_{lim} = \frac{1}{3}h_{tot} \\ mth = 8: \quad u_{lim} = 5\left[\frac{m}{s}\right]; \quad h_{lim} = \frac{2}{3}h_{tot} \\ mth > 8: \quad u_{lim} = 6.5\left[\frac{m}{s}\right]; \quad h_{lim} = h_{tot} \end{cases} \tag{18}$$

The use of different CTV comparing the regular with novel vessel shows disparate methods of O&M planning that led to variations of downtime losses having the possibilities to operate on 6 turbines along a day instead of a maximum of 5. Moreover, the selection of optimal turbines combination over a sequential group or worse case options with the highest power losses, highlight the potential of such study.

Finally, the results obtained shows high potentials for reducing downtime losses using a non-binary criteria respect to a fixed u_{lim}. In addition, taking into consideration the selection of turbines on the wake effect losses demonstrates the improvement possibilities that such strategy has. The energy gains per year are showed in the Table 9, evaluating the energy losses and the AEP along the overall period taken into consideration of the OWF. The increase of households that could be alimented per year is evaluated considering the data retrieved from [54]

Table 9: Energetic results of the analysis of [2]

	Optimized, Regular CTVs	**Optimized Binary New CTVs**	**Optimized 5T, Non-Binary**	**Sequential 5T, Non-Binary**	**Optimized 6T, Non-Binary**
Energy Loss [%]	-2.74 %	34.09 %	-36.53 %	-34.90 %	-48.18 %
AEP [%]	0.01 %	-0.13 %	0.14 %	0.13 %	0.18 %
Households	9	-114	122	117	161

3.4.2. Present case studies

Continuing from the results obtained in the previous work, the structure of this investigation is then explained. Along this study, the Python language with the application of its different libraries it has been employed. Specially the application of the library FLORIS it has been widely used to calculate the AEP and the consequences of O&M decision taken on the OWF energetic productions.

A maintenance activity requires shutting down the wind turbine during the operation, causing downtime losses and reduction in energy production. The energy lost is the lack of power that could have been generated given the wind speed and direction on the turbine's rotor during the maintenance hours. Selecting the optimal set of turbines to operate aims to the reduction of the loss occurred. To do so under specific weather conditions requires the knowledge of such losses and the makes of assumptions. As the wind direction and speed can vary significantly over time, it's important to base this choice on a proper representation of these parameters to evaluate the impact of O&M on the overall time of the operation.

The method selected in this paper to evaluate the wind direction on a workable day consist in evaluate each point of the day obtained dividing them by directional groups with a step value of 45° from 0° to 360°. The mode of the class mostly occurred along a day is selected and finally, is evaluated the median of the modal class as representative wind direction.

Instead, the calculation of a constant value for the wind speed along a day consider the cubic influence that the wind speed has on the power production. For so, it is determined the cubic root of the average cubic value of the wind speed along the time under study.

$$\boldsymbol{v_{speed,\ repr}} = \sqrt[3]{\frac{\sum_{i=0}^{N} v^3}{N}} \tag{19}$$

As the weather data available has a resolution of one hour, the representative values for these parameters are made on 24 points for the AEP analysis and on 11 points for the O&M analysis (from 8 am to 18 pm). The results showed more further, display how this method is a far estimate to evaluate the OWF weather condition leading to an unprecise AEP respect to the application of a variational wind speed with an hour step.

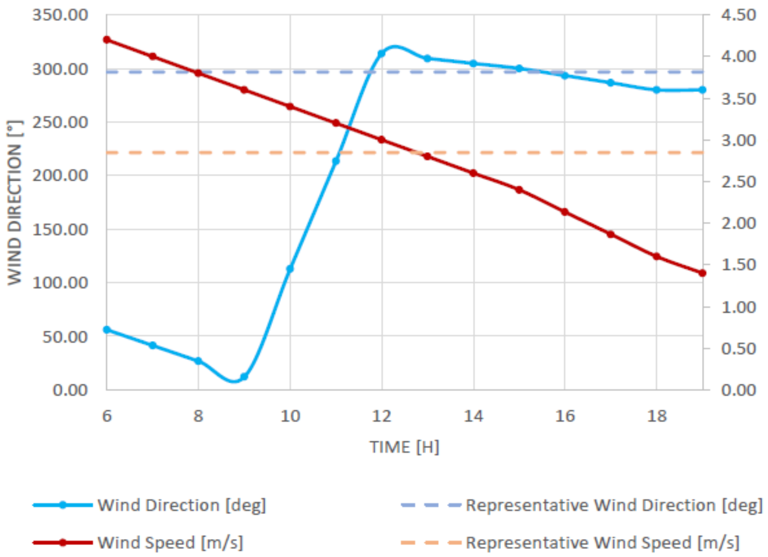

Figure 22: Representative and hourly wind condition for 19th October 1992 [2]

For the AEP investigation, the weather data from IPMA source are used for the period analyzed. The use of this dataset is used to evaluate in Floris the AEP of the OWF confronting it with the hindcast of ERA5.

At start of the O&M analysis, the evaluation of a feasible day for O&M is conducted, considered so if the forecasted wind speed from HRES forecast and significant wave height from ENS model do not exceed the safety limits for the CTV throughout the day (8am-6pm).

Preventive O&M activity require availability of equipment and operators. A trade-off in advance time planning operations and lead time forecast accuracy is needed. To evaluate the different impact that lead time in weather forecasting could have on the O&M planning, 3 separated scenarios of advance time decision have been implemented:

- 1-day lead time probable decision (_); the forecast considered have 1-day lead time to the O&M activities and the SWH is taken from the 50th percentile (P50) of the ENS model.
- 1-day lead time conservative decision (1 day); the forecast considered have 1-day lead time to the O&M activities and the SWH is taken from the 90th percentile (P90) of the ENS model.
- 3-day lead time conservative decision (3 day); the forecast considered have 3-day lead time to the O&M activities and the SWH is taken from the P90 of the ENS model.

The consideration of the 90th percentile over the P50 in SWH ensures a more conservative strategy for workable day decision as examine higher values of wave height that could occur along a day due to a wider consideration of parameters results from the ENS forecasts.

Once a day is considered workable, the decision to maintain turbines is made based on the representative values of the HRES forecast data choosing a combination of aerogenerators with a non-binary intervention strategy mentioned above. Knowing the forecasted weather data condition and setting the design of the OWF, through the FLORIS model, is calculated the annual energy production and the expected O&M energy losses. The actual energy production of the OWF is evaluated in a similar way, using the hindcast data from the ERA5 dataset as input for the weather conditions. In such way, it is possible to understand the discrepancy of the expected results from the occurred outputs due to the inaccuracies of weather forecasts.

The FCM data-driven method has been integrated and a validation on the accuracy it has been made for the two analysis conducted. The validation of the output data for the unsupervised method has been made by the Root Mean Square Error (RMSE) method showed below, from the considered historical data (ERA5 dataset), and compared to the RMSE of the IPMA weather forecast.

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}} \tag{20}$$

With $\hat{y}_i$ the predicted i value (FCM centroid or IPMA wind direction) and with y_i the ERA5 observation of the wind direction. Then, the energetic calculations have been corrected with the influences of the new weather input data from the data-driven method.

A comparation between the different possibilities of combination of turbines under maintenance has been made to understand the efficacy of the optimization. Specifically, in this study, we explore three distinct strategies for turbine selection:

- Sequential selection, which does not consider the wake effect;
- BEST selection, aimed at minimizing downtime energy losses due to wake effect;
- WORST selection, designed to maximize losses.

The ranks of turbines combination are selected based on previous research findings [2].

Table 10: AEP cases analyzed

forecast delay	weather data used	code
1 Day P90	Representative values	*Representative*
	ERA5	*Real*
	IPMA	*Forecast*
	FCM integration memb30%	*ML_memb_30%*
	FCM integration memb40%	*ML_memb_40%*
	FCM integration memb50%	*ML_memb_50%*
	FCM integration memb60%	*ML_memb_60%*
	FCM integration memb70%	*ML_memb_70%*
3 Day P90	Representative values	*Representative*
	ERA6	*Real*
	IPMA	*Forecast*
	FCM integration memb30%	*ML_memb_30%*
	FCM integration memb40%	*ML_memb_40%*
	FCM integration memb50%	*ML_memb_50%*
	FCM integration memb60%	*ML_memb_60%*
	FCM integration memb70%	*ML_memb_70%*

Table 11: O&M cases analyzed

forecast delay	selection of rank turbine	code
1 Day P50	best	*OLD_BEST*
	sequential	*OLD_SEQ*
	worst	*OLD_WORST*
	best + FCM integration	*Normal ML*
1 Day P90	best	*OLD_BEST*
	sequential	*OLD_SEQ*
	worst	*OLD_WORST*
	best + FCM integration	*ML 1 Day unc*
3 Day P90	best	*OLD_BEST*
	sequential	*OLD_SEQ*
	worst	*OLD_WORST*
	best + FCM integration	*ML 3 Day unc*

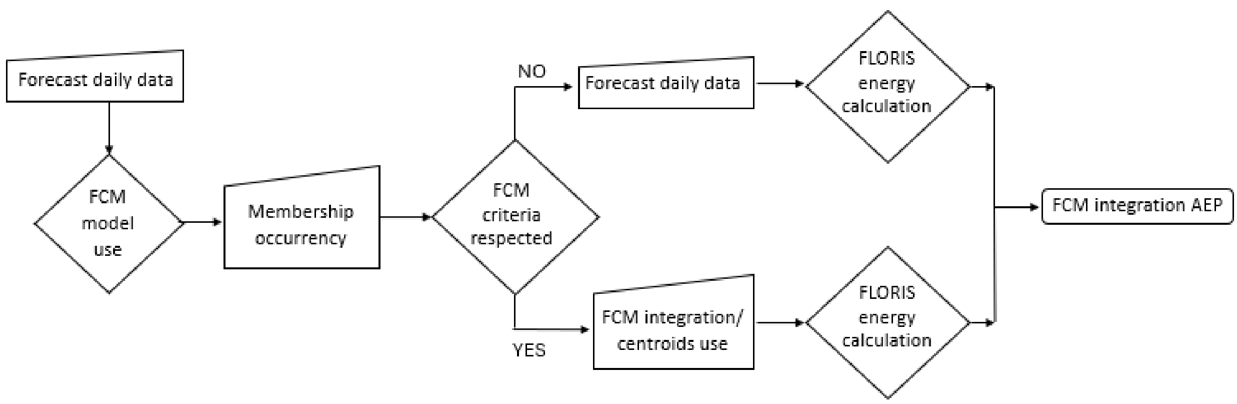

Figure 23: flowchart of structure for AEP analysis

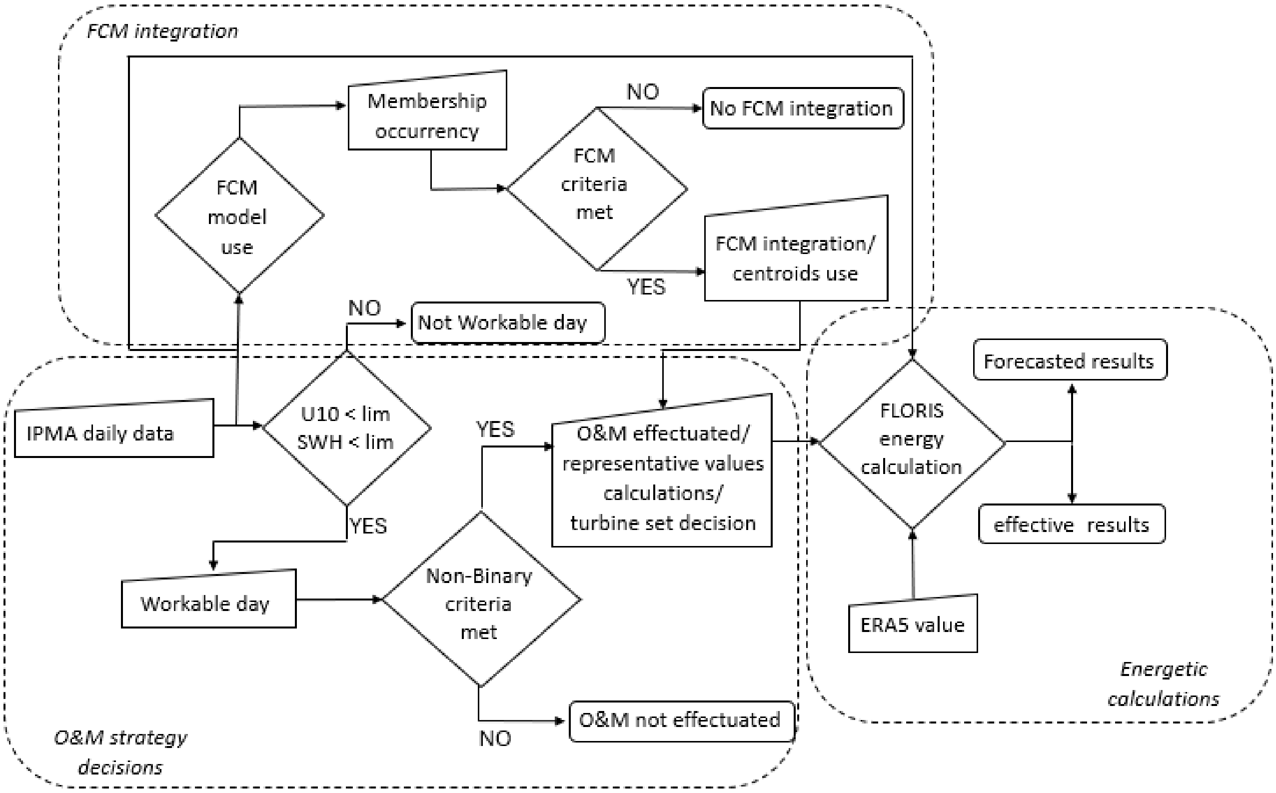

Figure 24: flowchart of structure for O&M analysis

4. Results obtained

This part of the work shows the result obtained in the study conducted describing in detail the consideration made from such outputs. The presentations are structured dividing the different analysis conducted, analyzing firstly the AEP with the data-driven integrations. Subsequently the O&M results are analyzed evaluating the impact of the forecast uncertainties to end with the optimized results obtained with the use of the FCM methods. Finally, a sensitivity analysis is presented to show the different possibilities to implement the ML and to have a better understand of its features.

4.1. AEP for offshore wind farm

4.1.1. Validation of FCM integration in AEP weather data

Before the illustration of the AEP calculation conducted for the OWF taken in consideration, are shown the validations of the weather data obtained integrating the FCM method. To establish the improvement of the FCM accuracy on the wind direction it has been used the RMSE of the model from the ERA5 data. Confronting it with the RMSE of the wind direction taken from IPMA forecast, it is possible to recognize if a higher precision for such parameter is reached.

The data-driven method behaved differently across the months and the wind velocities categories in which has been divided. The accuracies obtained for the wind direction are not uniform, showing higher precision in the analysis made with three-day delay compared to the IPMA forecast evaluation of the parameter. Such differences in precision are dependent not only on the abilities that the FCM model has to adapt better in particular situations, but also on the decision made for the number of partitions in the categories analyzed. In fact, the variations are caused also by the discrepancies in how the model was trained on different datasets. This underscores the importance of conducting tailored analyses that are specific to the context at hand. An example of graphical representation for the wind direction evaluation is presented below in Figure 25 where are compared the forecast data with the output obtained from the FCM integration.

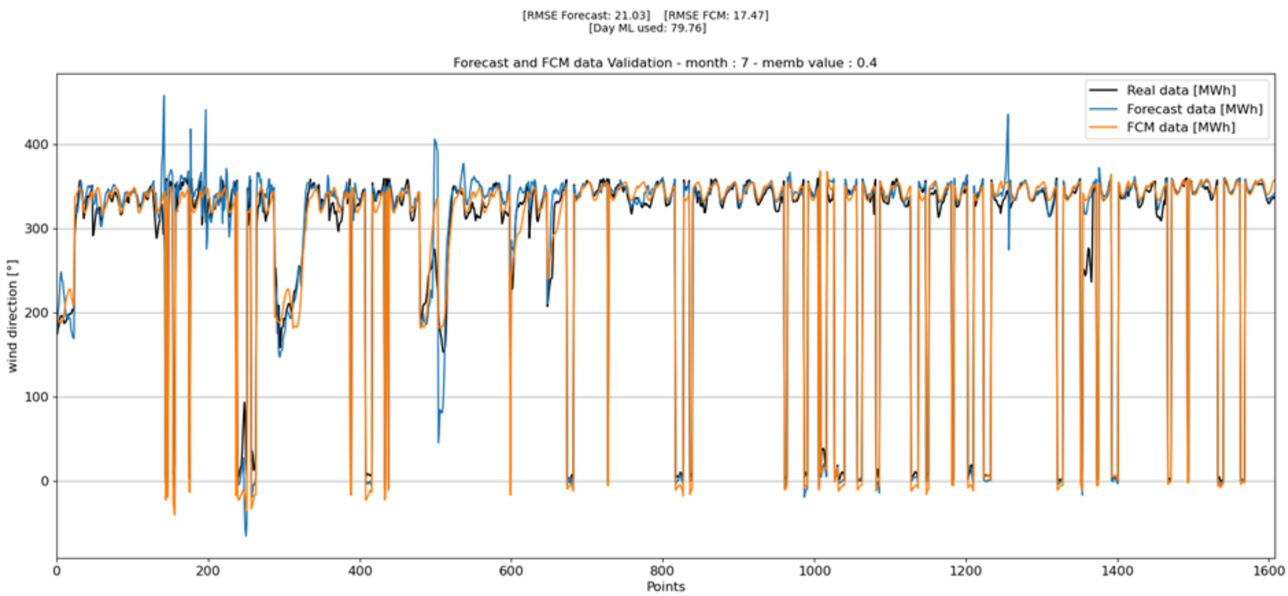

Figure 25: Wind direction for 3-day delay in July from different datasets

To deeply understand the evaluation of the precision for the wind direction compared to ERA5 dataset, below are presented the FCM RMSE compared with the IPMA RMSE estimation separating the different categories of wind velocities considered. The graphs shown represent the results obtained for the various membership limit used for the FCM integration followed by the overall categories. A value below 1 means an improvement of accuracy with the FCM method.

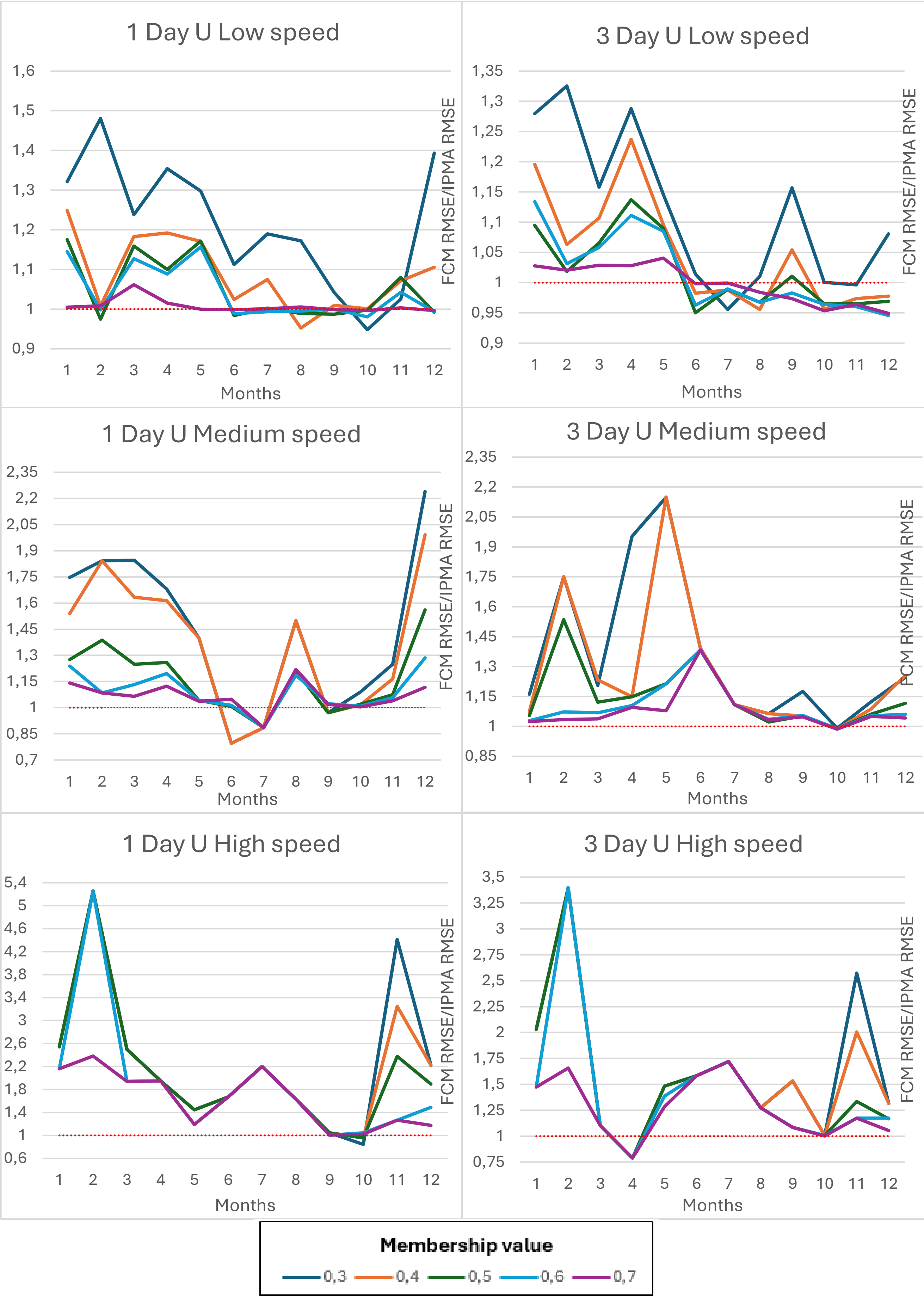

Figure 26: FCM RMSE over IPMA RMSE along the months divided for wind speed and membership limits for AEP data validations

Except some particular cases, it is possible to notice that in 3 days lead time forecast the RMSE of FCM found has generally a lower value compared to the integration of the machine learning on a single day lead time forecast. This occurs since higher lead time in weather forecasting increase the inaccuracy of the prevision. So, integrating a data-driven method trained on historical data for longer lead time forecast for certain months increase the accuracy of the data. Specially, improvement in precision have been found for low wind speed days. This might be reconducted to the fact that low wind velocities are characterized by frequently variations of wind directions, making it hard from forecast model to precisely estimate such parameter in these conditions. Consequently, the data-driven model integration that rely on historical data is found more accurate for these cases.

Another important consideration is highlighted by the increases of membership limits for FCM outputs. Using such parameter as decision if a weather estimation along a day it could be substituted with data-driven data, generally bring higher accuracy for the machine learning estimation. Such fact is not always respected. In fact, for some months analyzed, consider wind data with lower membership assigned, increase the precision of the estimations. This might be related to the fact that, for some days with high variation in wind direction, an output with lower certainty in cluster membership have higher potential than weather forecast.

Considering the overall results applying the FCM in the estimation of the wind direction along the various months is confirmed a notable improvement on accuracies. Lower uncertainties are present for some months specially for the summer season where the wind speed is mostly lower. The highest improvement reached reduce the RMSE of almost 15% for the month of July showing a good potential of the method applied.

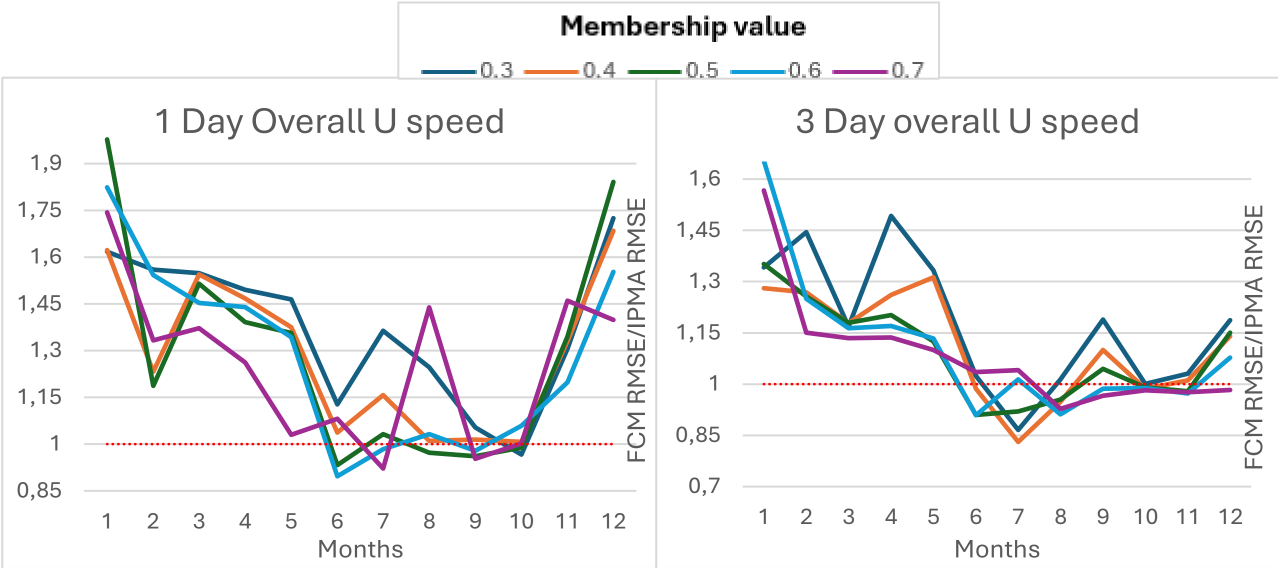

Figure 27: FCM RMSE over IPMA RMSE for the overall data along the months in different membership limits

The increase of membership limit affect the FCM use, as it is shown for 1 day lead time:

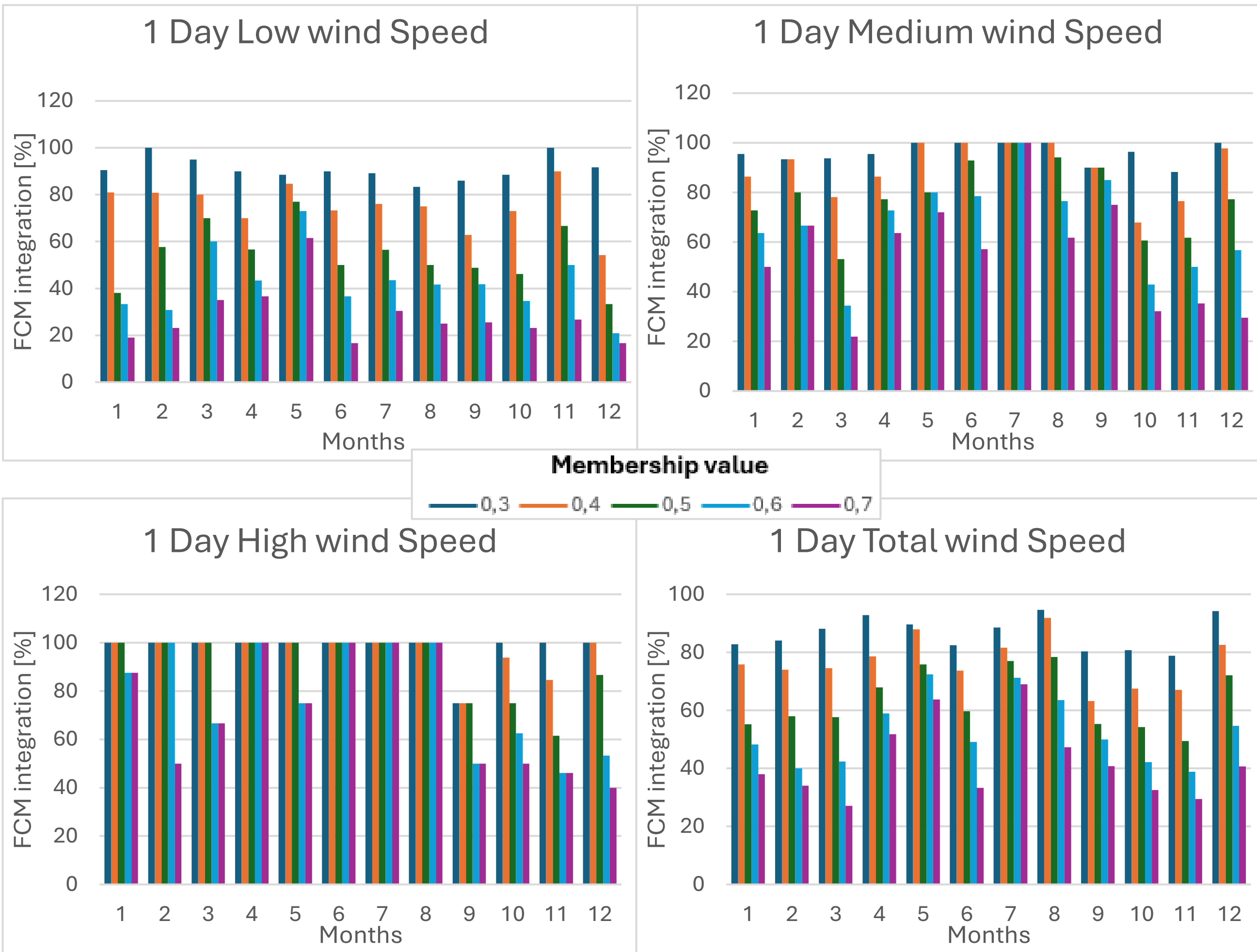

Figure 28: Percentage of FCM integration along all the time period analyzed for 1 day lead time forecast for AEP data validation

Is it notable that days with lower wind speed are characterized by a reduced number of memberships value. An increase in the limit for such cases, lead to a consistent reduction of FCM integration. Instead, when high wind velocities occur, is easier for the data-driven method to recognize a daily wind direction pattern, assigning high value of membership to the outputs due to the less variability of wind directions along a day.

In final, the integration of the FCM model in forecast weather data for the calculation of AEP along all the period analyzed and on the various speed categories are shown in the table below.

Table 12: FCM final validation results in AEP analysis expressed in RMSE FCM/ RMSE IPMA

Membership	**0,3**	**0,4**	**0,5**	**0,6**	**0,7**
1 Day lead time	1,3089	1,2419	1,2272	1,2454	1,2040
3 Day lead time	1,1427	1,0866	1,0611	1,0552	1,0337

The use of the data-driven model on all the data did not achieve a total improvement for the RMSE comparing with the forecast uncertainties, where it could be highlighted a similar uncertainties value for the 3-Day lead time case specially using higher memberships values.

4.1.2. AEP results with FCM integration

After the data validation that brought a deeper understanding for the features used for the FCM method, it has been calculated the effective AEP from the various set of data available. The AEP is computed for one- and three-day forecast lead time, utilizing IPMA data for this wind direction while keeping U10 and SWH correspondent to the ERA5 dataset. This decision aim to isolate results around wind direction, excluding uncorrelatable differences in energetic production derived from modified wake effects with wind speed variations, in order to better comprehend the influence of the FCM integration. As the energy outputs of the OWF are not available due to the actual installation of only 3 turbines in the location, the comparison of the AEP will be effectuated on the hindcasts data of ERA5 (*Real AEP*). In fact, the energetic results evaluated are normalized on such estimation as follow:

$$Estimated\ AEP\ normalized = \frac{Estimated\ AEP}{Real\ AEP} \tag{21}$$

A value closer to 1 means a closer estimation to the assumed *Real AEP* occurred along the time period analyzed. In the figures Figure 29Figure 30 are also shown the percentage of cases where the integration of the FCM method with the forecast data it occurred.

A parallel approach is involved using the representative wind value calculated to highlight the limited accuracy reached evaluating a daily representable wind parameter.

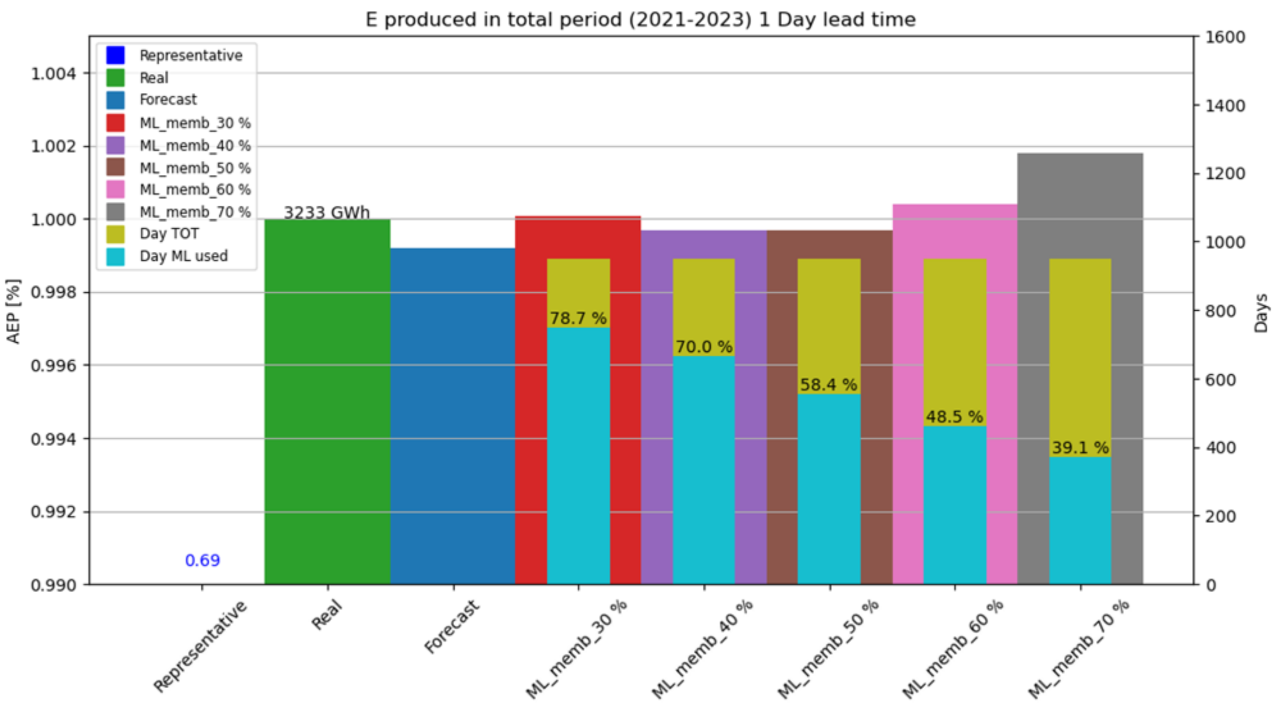

Figure 29: AEP 1-day lead time

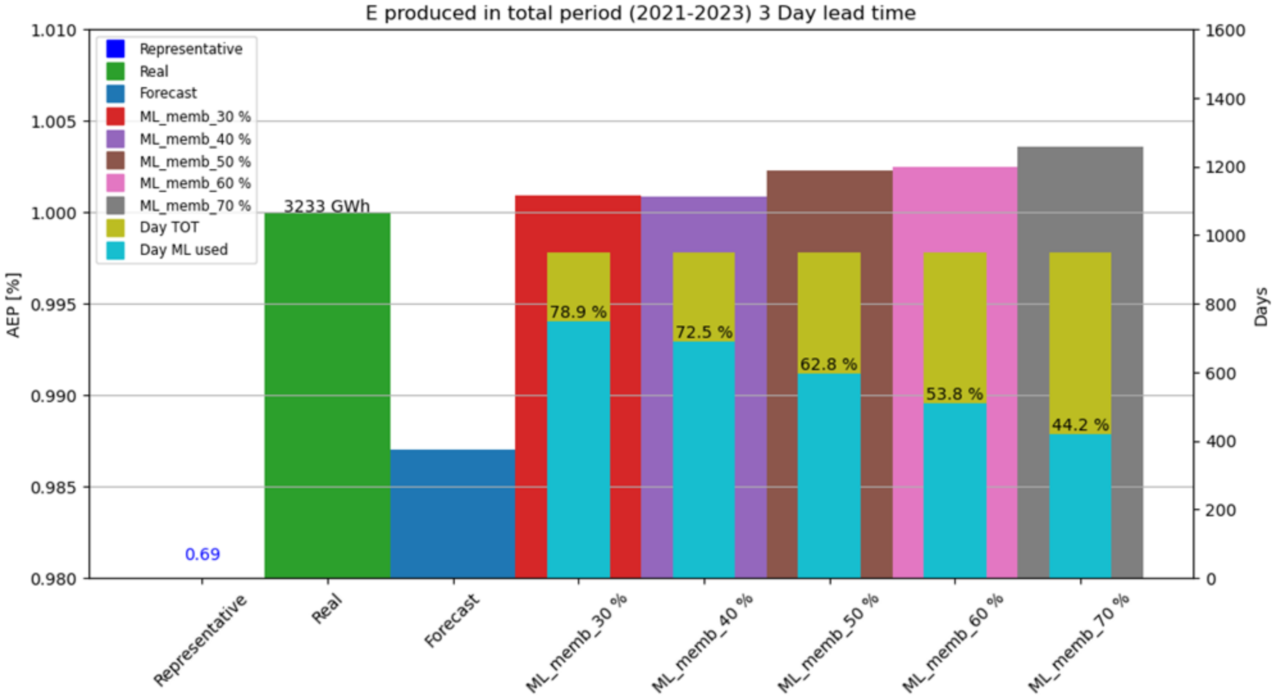

Figure 30: AEP 3-day lead time

Comparing the results of the energy analysis and incorporating the FCM algorithm for wind direction correction, is demonstrated an improved precision in AEP calculations for most cases, except for the case of 1 day-lead time characterized by a membership limit higher than 70%. These consequences could be attributed to the fact that, the achievement of higher accuracy in wind direction estimation does not necessarily result in a more precise AEP calculation, as various parameters interact differently with distinct weight contributions.

4.2. O&M scheduling optimization

The study found that when the exact knowledge of wave and wind conditions is not available, the results of the optimization of maintenance strategy could be highly affected. This can lead to suboptimal choices due to unprecise information. The study also found that the advance of days in decisions based on forecast model impact the availability of workable days to operate maintenance activities.

Three different input of forecast value with different lead time are used :

- *"_"* : Values with 1 day lead time and SWH parameters from 50 percentile (P50) of ENS
- *"_1 Day"* : Values with 1 day lead time and SWH parameters from 90 percentile (P90)of ENS
- *"_3 Day*" : Values with 3-day lead time and SWH parameters from 90 percentile (P90)of ENS

Increasing the time in advance in planning the maintenance operations rise the uncertainties of precise weather condition forecasts. Specially, when the SWH is considerate an important factor for the decision maker that rely on the ensemble percentile for possible sea weather conditions. In fact, an output of ensemble forecast with higher lead time has a higher variability of parameter's evolution. This cause an overestimation of waves height leading to a reduction of workable days available, with consequences of missed opportunity to O&M. An excessive reduction of days with conditions that allow a maintenance operation, could result in an unsuccess to respect the total amount of maintenance hours set per year, as occurred in the cases of regular vessel with 3-day lead time forecast (*df_OLD_BEST_3 Day*). The consideration of the two types of vessels has been taken only for the outcomes of the schedule analysis and are shown below.

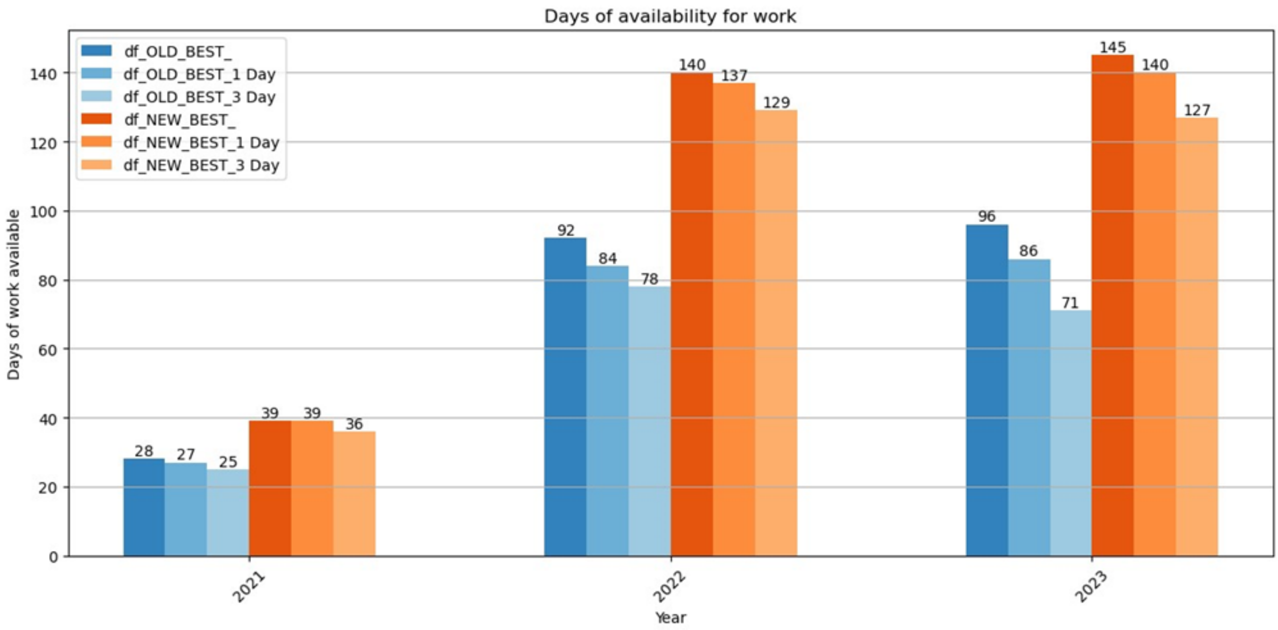

Figure 31: Workable days per year for different lead time forecast data

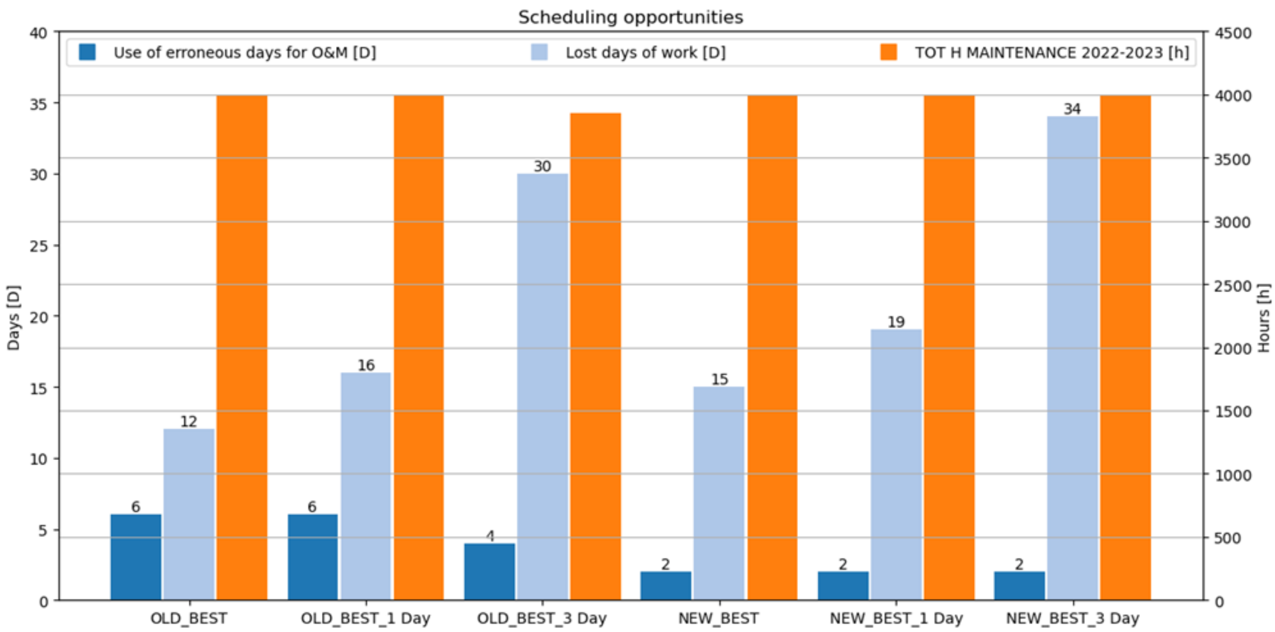

Figure 32: Outcome of different lead time forecast planning of O&M

The use of novel CTV for maintenance purposes highly increases the availability of days with favorable conditions to effectuate O&M operations. This ensures a larger possibility to achieve the amount of maintenance hours set along the year and create a wider set of decisional possibility for the operations. Furthermore, is it possible to highlight the increases of lost opportunities for O&M relying on higher lead time forecasts. It is also possible to notice that are encountered less erroneous days evaluated from the strategy applied. In fact, due to the same consideration that higher weather variability rises the ensemble percentile values, less events are evaluated as workable days where the weather conditions are not suitable for the conduction of a safe operation.

Considering the energetic output obtained from the research, the power losses occurred along the time analyzed are evaluated for the different turbines selections strategies. The results shown in *Figure 33* and *Figure 34* are normalized on the worst case scenario where the combinations of turbines under maintenances generate the higher energy losses that could occur along the operations. This comparison highlights the potential that an optimized strategy (*Best*) could have on a systematic decision (*Sequential*) evaluating the range of results that could be encountered.

The results shown are divided in *Forecasted Energy Losses* and in *Real Energy Losses*. The first are the consequences expected on the energetic productions that rely on the forecast data available, on which the decisions have been taken. Then the effective energetic losses are encountered and could be registered only after the occurrence of the event. For our case study these calculations are evaluated with the *Forecasted energy Losses* utilizing the ERA5 weather data for the effective energetic results in consequences by the previous decisions taken.

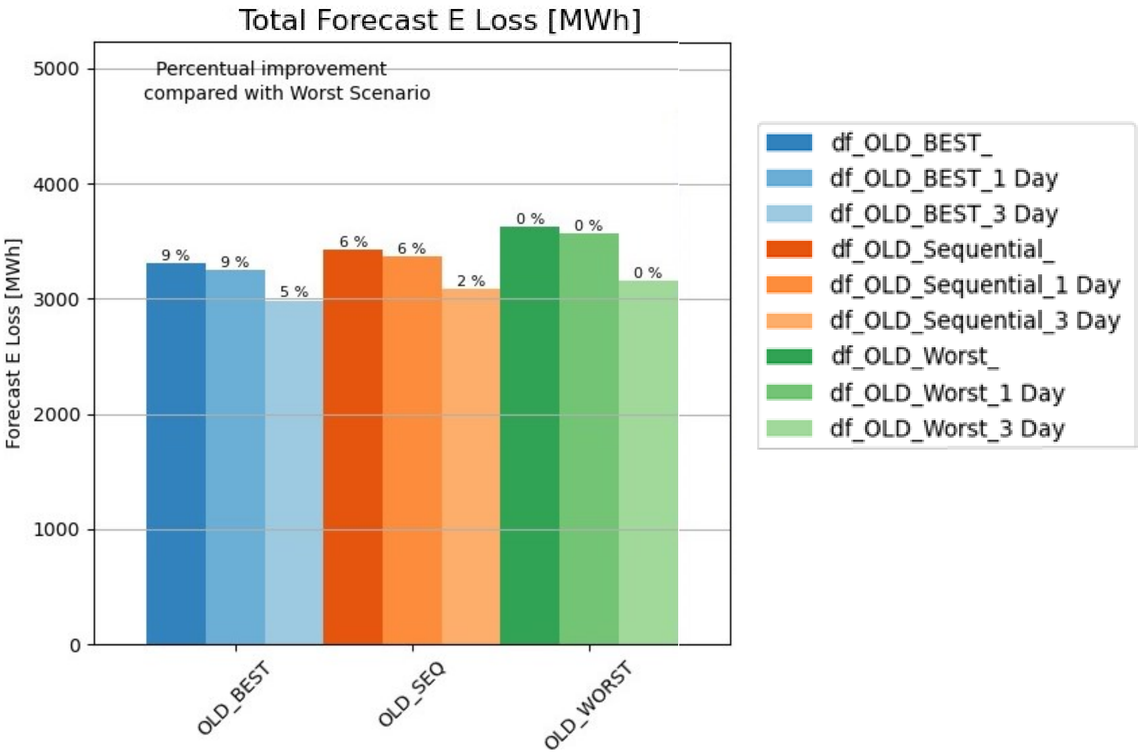

Figure 33: O&M Energetic Losses expected due to maintenances activities

As expected, the uncertainties in the weather forecast have reduced the optimization in the analysis conducted. As demonstrated in [2], the strategy applied minimize the downtime energy losses considering the estimation of wind conditions, with a maximum reduction of 9 % respect to the worst case turbines selections for a lead time of 1 day decision and a decrement of 5 % on 3 days lead time.

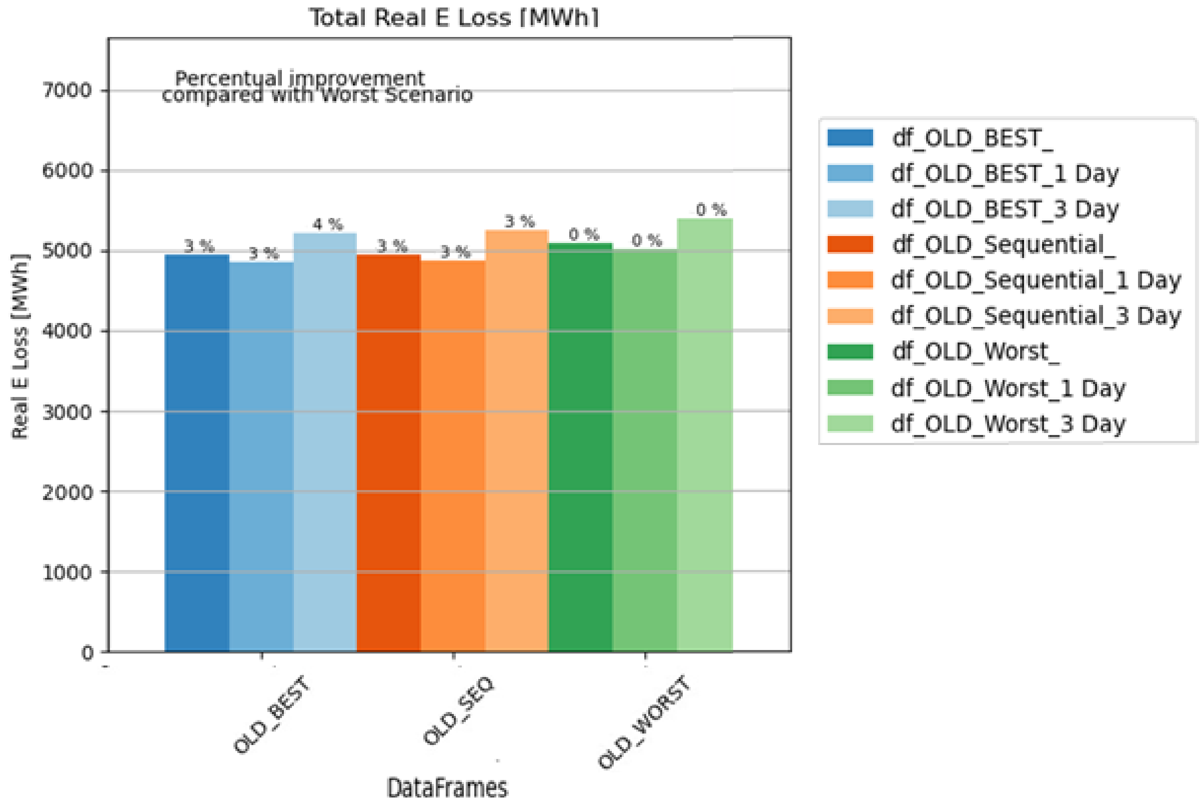

Figure 34: O&M Energetic Losses encountered due to maintenances activities

Focusing on the effective energy losses occurred when carried the O&M activities, the uncertainties on the forecast model lower the potential of the approach used. In fact, the optimized scenario does not achieve significant higher results respect the sequential decision on turbines selections, but still improve the energy generation of the farm considering the worst case possible, validating the sense of the study conducted. The small difference in energy losses between the best and sequential choices can be attributed to wind direction uncertainties, which make challenging find the best combination of turbines for power generation besides, fluctuating wind speed magnitude leads to varying amounts of power outputs.

Such low results are due to the fact that, in lack of observed weather data for the location, as method of comparison it has been used an hindcast method. As shown from the data comparation in the previous section, shows a not perfect correlation with the presence of bias errors, leading to a non-optimal decision taken along the evaluation of turbine selection.

The discrepancies of the results obtained from the expected and occurred downtime energy losses are mainly due to the wind velocity imprecisions generated by IPMA forecast. In fact, from the figure below is possible to see that the ERA5 dataset is characterized by a higher mean value of wind speed, leading to a substantial higher energy production. In consequence of this, along the days selected for the O&M operations the effective energy losses are higher than expected and subsequentially, the uncertainties in wind direction impact more on the results.

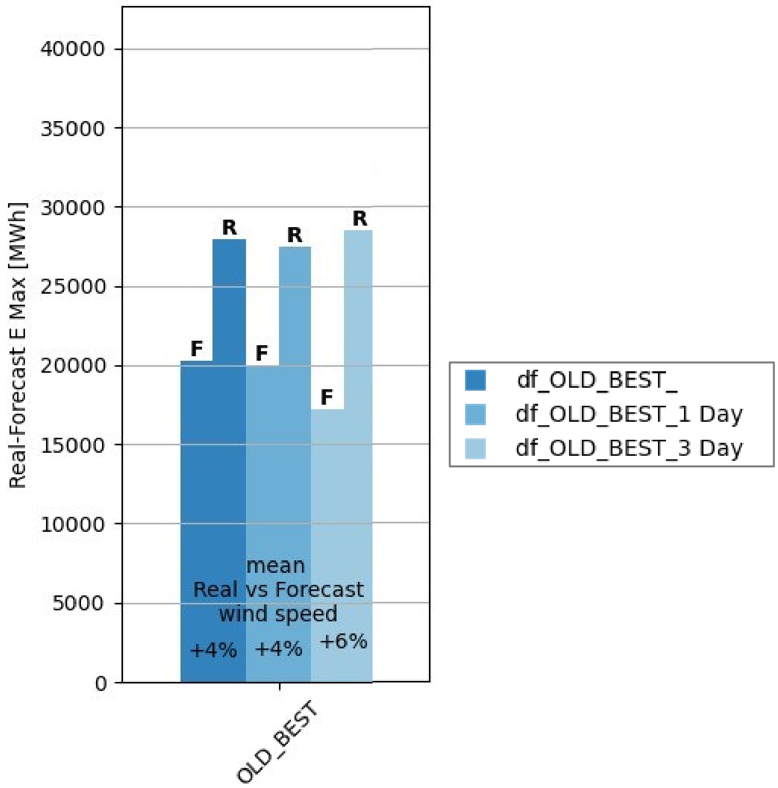

Figure 35: Total Real vs Forecast energy production if no O&M activities are conducted for different lead time forecast

4.3. O&M scheduling optimization with FCM integration

4.3.1. Validation of FCM data integration in O&M weather data

As showed in the previous paragraph, the optimization of turbine selection has a potential if the weather forecast precisely align with the real conditions. In case the imprecisions of the forecasts start to increase, the decisions for the combinations of the aerogenerators to maintain differ to the optimal set. In particular, the calculations for the daily representative wind speed and wind direction evaluate suboptimal assumptions of the weather conditions.

To improve the uncertainties of the wind previsions it has been used the FCM algorithm instead of the forecasted data obtained from IPMA models. Such improvement is made replacing the dates at which the decisions is being made by the centroids with the higher membership clusters during the selection of wind turbines to be operated. So is done in a bid to involve uncertainties in enhancing data accuracy for decision-making purposes.

Comparable results with AEP were seen during O&M data validation using the FCM algorithm. The data-driven method has exhibited different behavior based on the selection of months, as indicated in the table below. The FCM integration reduced the RMSE sum of all lead time forecast cases, better data acquisitions for O&M scheduling were made when it showed a reduction of errors of wind direction.

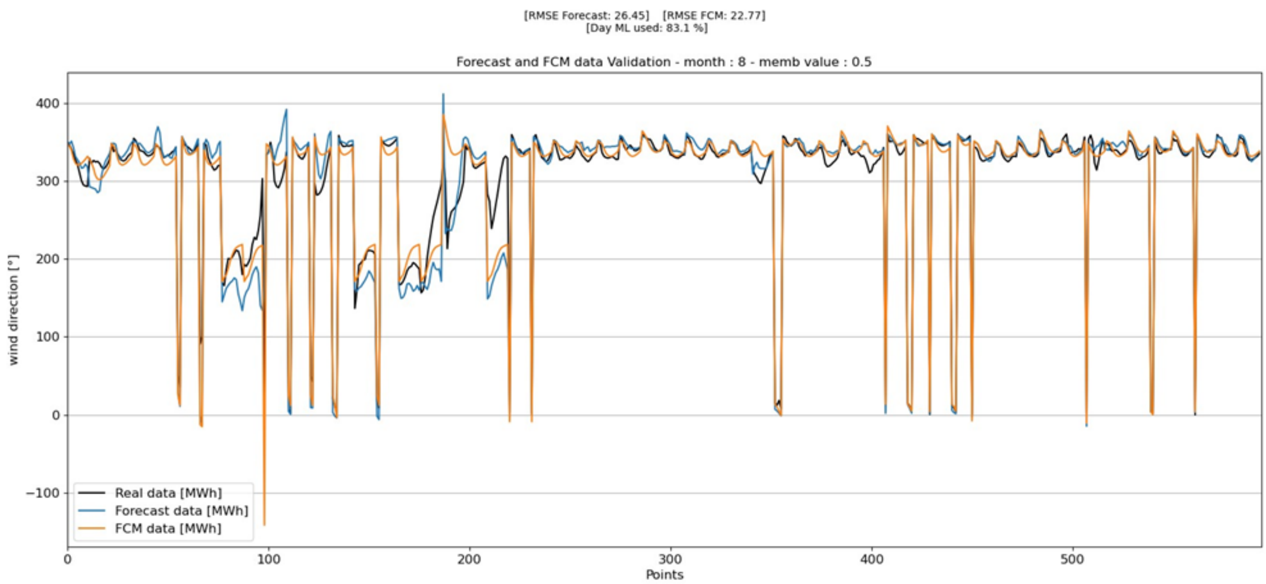

Figure 36: Wind direction for 3-day delay in August from different datasets

To further understand the performances of the FCM on this analysis, are shown the comparisons between the RMSE of the data-driven model with the RMSE of the IPMA forecast. Such validations are made only for the daily hours (6 am to 7 pm) of the months most suitable for the maintenance activities (from April to October). The verifications of the results are divided for the various memberships limits used on the ML to recognize the most probable wind pattern direction (Figure 38). In addiction, the percentages of the FCM integrations are presented over the total available days for O&M operations for each months in the Figure 37.

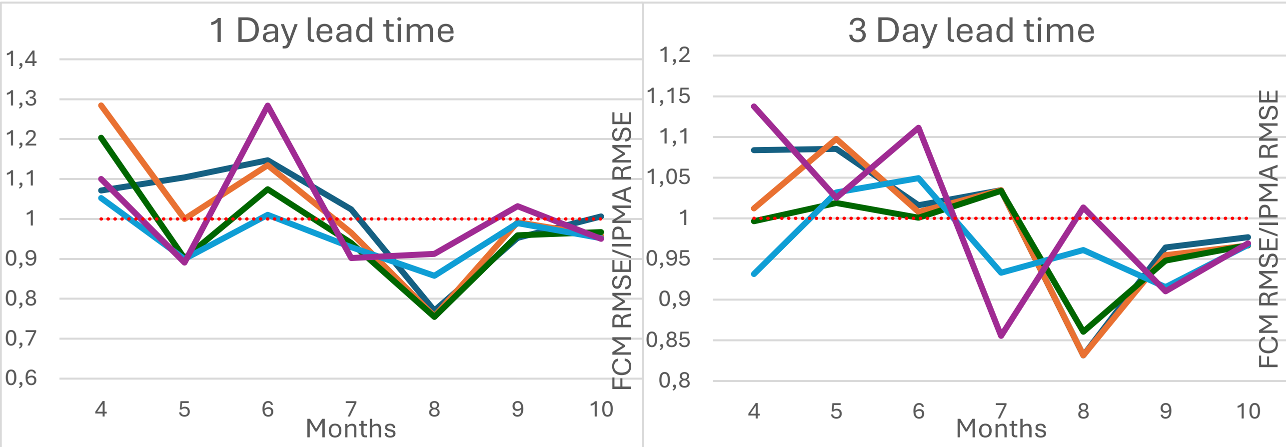

Figure 38: FCM RMSE over IPMA RMSE for months divided in membership limits for O&M data validation

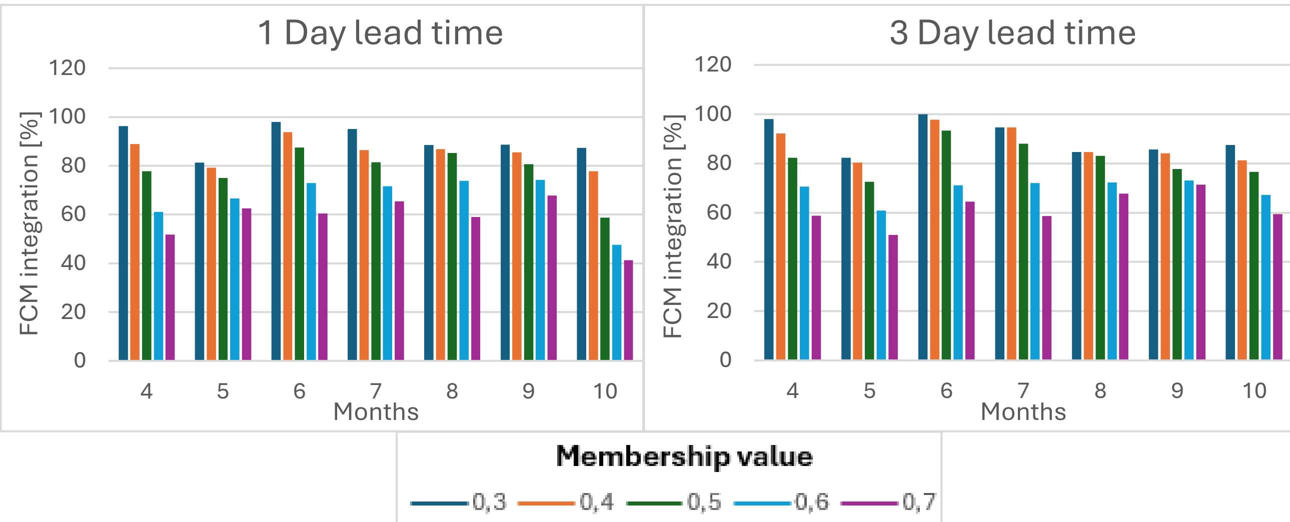

Figure 37: Percentage of FCM integration for all the time period analyzed for 1- and 3-day lead time forecast for O&M data validation

The results display the various improvements achieved integrating the FCM algorithm on the weather forecast. The utilization of the data-driven method in the present analysis obtains higher accuracy in the estimation of wind directions compared with the AEP study. In particular, for remarkable results are captured for the months of August for both 1 day lead time and 3-day lead time forecasts.

The increase of higher membership limit to accept the estimation of the model in the decision making reduce the utilization of the ML algorithm, as occurred in the previous AEP test.

At final, the RMSE values normalized that consider the total period of time in which the study has been conducted are shown in the table below.

Table 13: FCM final validation results in O&M analysis expressed in RMSE FCM/ RMSE IPMA

Membership	0,3	0,4	0,5	0,6	0,7
1 Day lead time	1,0104	1,0060	0,9746	0,9621	0,9983
3 Day lead time	0,9935	0,9818	0,9726	0,9675	0,9915

The integration of the data-driven model in the 3-day lead time forecast enhances the precision of the total wind direction estimations reaching a maximum of 4 % improvement in RMSE. In general, integrating such method in a longer lead time weather forecast show higher potentials due to the fact that rely on historical patterns for high variational parameter might be more accurate than long time predictions.

The results obtained not only show a reduction in forecasts imprecisions even if not all months considered were well suited and trained properly, but also demonstrate the potentiality that ML methods could help the forecast model in the weather predictions.

4.3.2. O&M energy results with FCM integration

After the proof that the method used closely align with the obtained forecast data from IPMA followed by the verification that its application could improve the precision of the weather estimations, is possible to analyze the impact that such strategy could bring to the study conducted. In fact, in this section will be presented the energy outcomes with the integration of the FCM model with the case of membership limit set to 0.5, compared with the scenarios exposed in the paragraph 4.2 in which such strategy is not applied.

This study noticed less progress when the set of wind turbines for O&M was evaluated. So is occurred as most of the emphasis was put on the reduction of wake losses variating turbines combinations under maintenance, thus little room for enhance performances were left. Remodeling the training technique as well as data pre-processing could enable the attainment of higher levels, while also excluding months that did not well represented the weathers conditions.

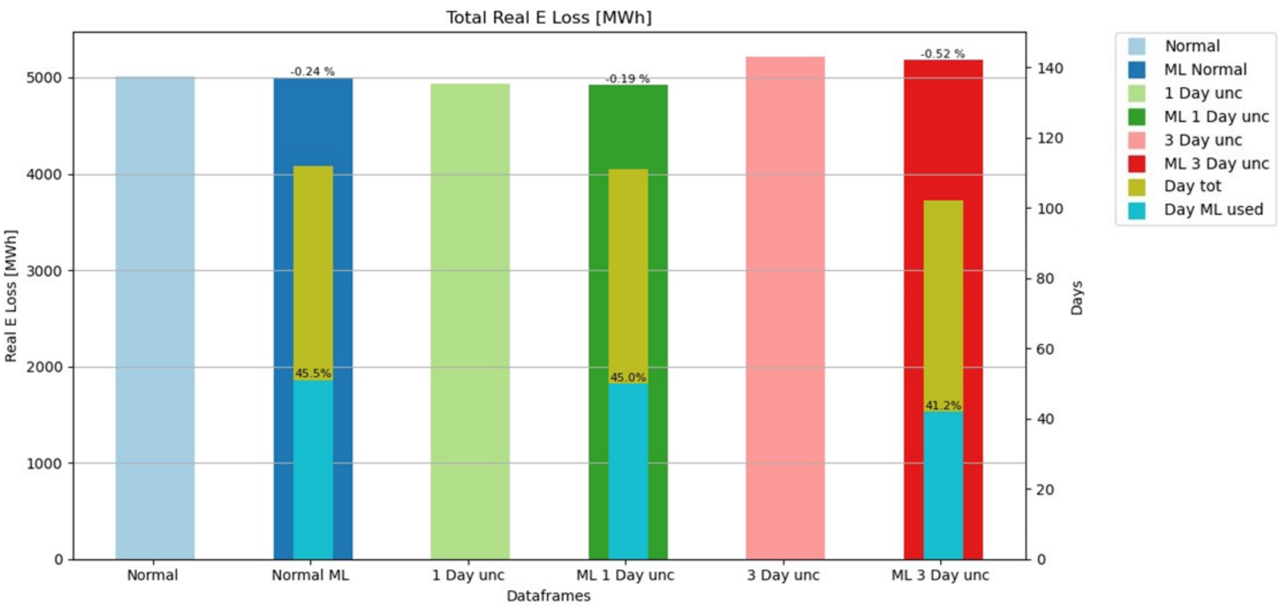

Figure 39: O&M scheduling with FCM algorithm membership limit = 0.5. Comparation with results in 4.2

The scheduling of O&M activities with the selection of aerogenerators under maintenance made with the data-driven model slightly reduced the downtime losses generated. Such low results are obtained as various assumptions have been made along the study, especially the imprecise daily representations of wind conditions and the incorrection of wind speed prevision affected the outcomes of the research. Furthermore, the ML algorithm has been applied to all the months considered including the ones in which the model did not behaved precisely.

4.4. Sensitivity Analysis

4.4.1. Sensitivity Analysis on SWH

The present study pursued has took some forced assumptions particularly regarding the criteria of maintenance operability and the limit to apply the FCM model in the scheduling strategy. For these reasons, in this section are shown two sensitivities analysis on the parameters mentioned, to evaluate the impact that such assumptions have brought to the results and to consider different strategies methods.

Firstly, is expanded the knowledge of the influence that the SWH limit has on the evaluation of a safe maintenance workable day. In fact, an approach that seeks to minimize the decisional error on work with unsecure weather conditions derived by forecast uncertainty is pursues.

The approach presented in [55] used a normalized alpha-factor (α) in the SWH operational limit factor. Five different values of the coefficient for wave heights were applied within this work (*0.9, 0.95, 0.98, 0.99, 1*). Correcting the value of wave height reducing the value limit permits to operate with a more conservative strategy. The time frame under consideration is limited to 2022-2023 due to the incomplete weather information provided in the 2021 and the results are normalized on the case with a coefficient factor equal to 1.

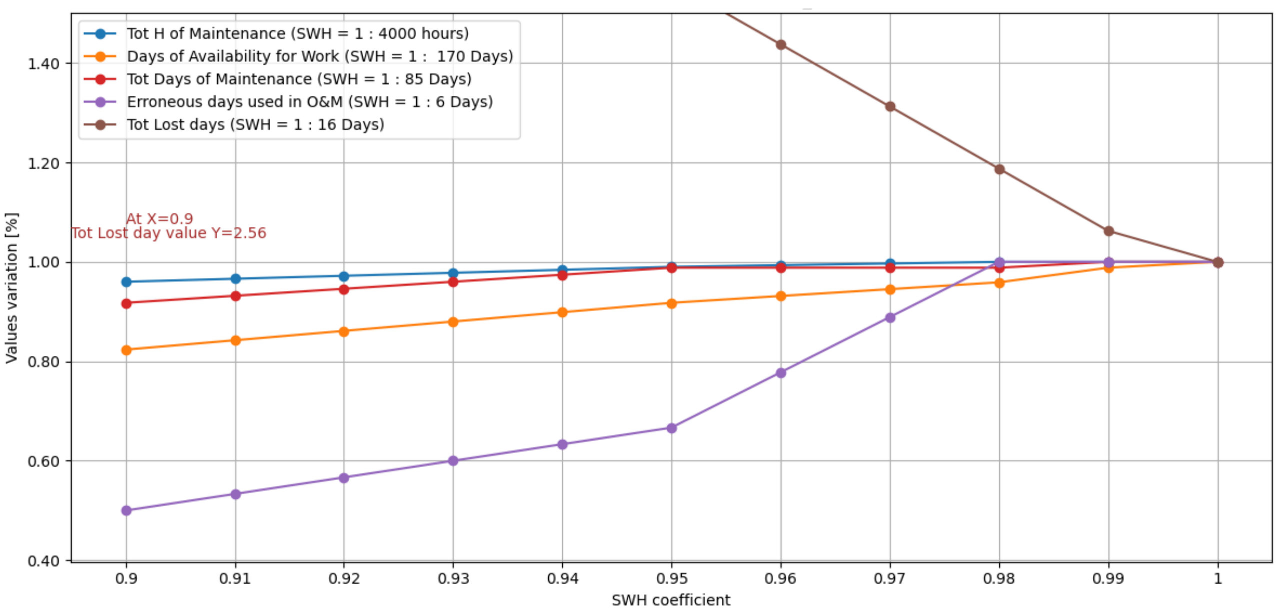

Figure 40: SWH sensitivity analysis 2022-2023 for 1 Day lead time forecast

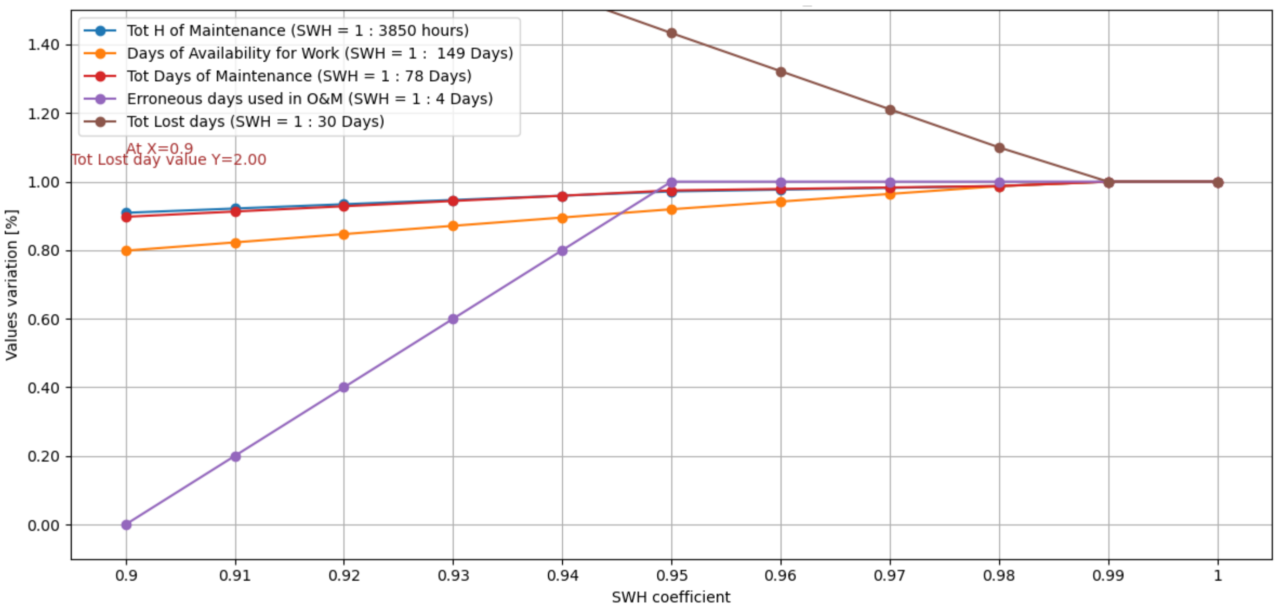

Figure 41: SWH sensitivity analysis 2022-2023 for 3 Day lead time forecast

4.4.2. Sensitivity Analysis on FCM membership limit

By choosing a more cautious approach, it is observable that the number of working days available for the conduction of O&M activities decreases for each year under review. So could result in a drop of hours reached for maintenances operations that could be carried out. Decreasing the SWH limit will also reduce the erroneous days registered, with the consequence of the safety enhance for the crew during the activities, cutting down the likelihood of any financial losses as a result of an abortion of the operation under execution.

As second deeper study on the parameters used, the influences that the membership limits have on the energetics results have been carried out. A sensitivity test changing this variable could enhances the degree of accuracy in the analysis when allocating days to certain similar clusters affecting the final energies results related. The variations occurred to the outcomes resulted from such investigation are normalized on the case with 0.5 as membership limit and are shown in the Figure 42 and Figure 43.

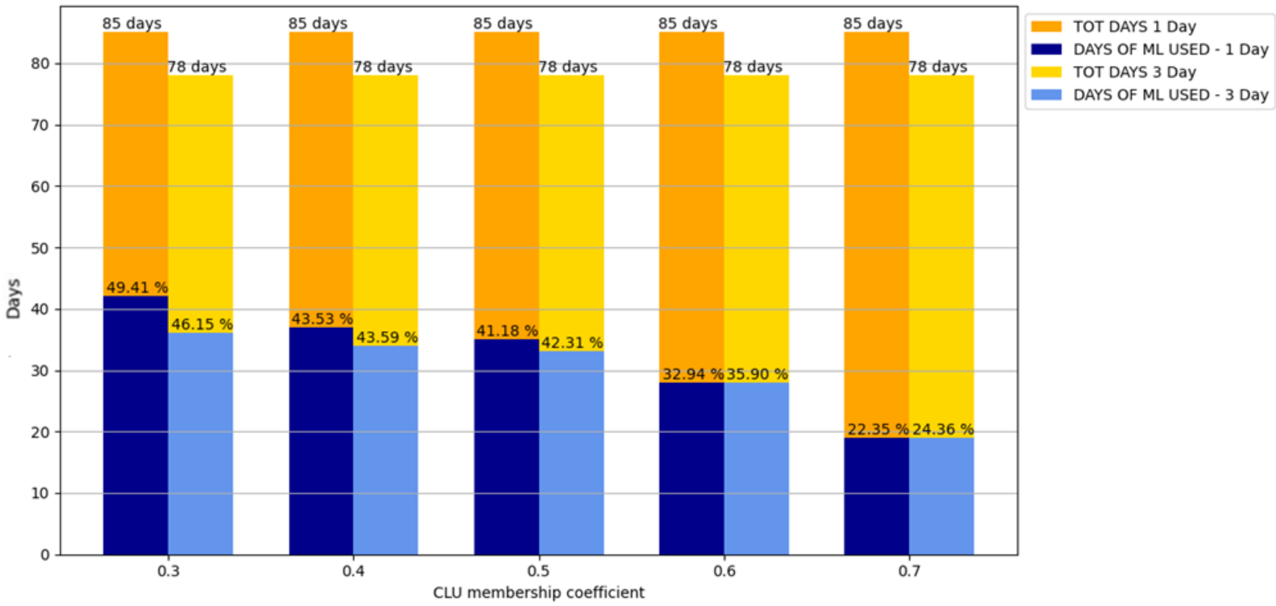

Figure 42: Membership limit sensitivity analysis 2022-2023 for 1- and 3-days lead time weather forecast

A reduction in the number of days where FCM integration occurs is shown. This is similar to the findings in the study discussed in the section 4.1.2 where the clustering of daily wind with low representative wind speed shows lower maximum membership values identified by data-driven method.

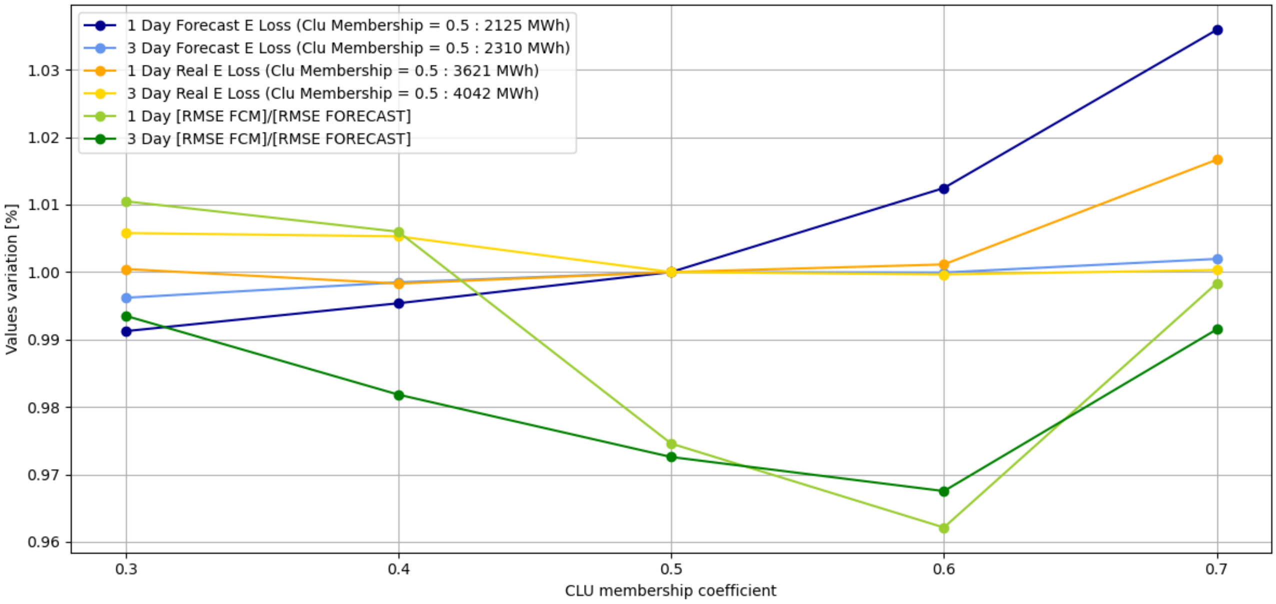

Figure 43: Membership limit sensitivity analysis on energetic results for 2022-2023 for 1- and 3-days lead time weather forecast

This phenomenon is attributed to the higher directional variability of the wind at lower velocities. In the context of this study, where preventive maintenance operations are scheduled during periods of modest wind speeds, it is observed that raising the membership limit does not necessarily lead to improved accuracy of the FCM compared to forecast data for these wind speeds. This occurrence can be attributed to the forecast's challenge in predicting such days. Utilizing FCM algorithm with a lower maximum membership value could enhance the results without overly constraining the model, reconducting to historical patterns days with high wind direction variation.

5. Conclusions and future work

The purpose of this thesis was to continue the previous study [1] and [2] on the optimization of scheduling strategy for preventive O&M activities, considering the wake effect. In this investigation it has been worked with assumed real data obtained from the ERA5 hindcast to understand the influence of imprecise weather forecasts when calculating the energetic productions by Floris tool. A further optimization it has been introduced integrating the FCM data-driven method to reduce the prevision uncertainties on weather conditions.

Following so the mentioned precedent studies, it has been applied a similar strategy for optimize the selection of the turbines to maintain considering the sea and wind conditions. More precisely, the SWH parameters it has proved itself crucial for the evaluation if a day is available or not to sustain a maintenance activity ensuring the safety of the crew members and the complete of the mission. In addition, the availability of novel vessels shown the increment of feasible days to operate on the wind turbines.
The wind direction instead, it has shown to be the most influent parameter in the selection of the turbines combination on which will be conducted the maintenance. In fact, considering the direction of the wind on the wind farm, it is possible to evaluate which set of aerogenerators has the lowest impact on downtime losses occurred due to their shutdown needed to carry out the activity planned.
The use of a classification algorithm based on previous weather conditions could help the knowledge of weather characteristic on a specific area and, if trained properly, could lead to an improvement of the quality of the data available. In this research, such strategy it has been applied only on the wind direction estimation, as considered the most important feature for the scheduling strategy chosen.

The results obtained confirmed the previous founds regarding the possibility of energy losses reduction by a scheduling strategy for O&M activities. Specially, the use of a specific implemented non-binary method to plan operations, it has demonstrated itself as a smart strategy to optimize energy production while conducting preventive maintenance operations. Although, the imprecision of weather forecast data could lead to errors in the calendar's activity generating economic losses. Moreover, suboptimal decisions could be taken due to forecast uncertainties decreasing the potential of such approach. The improvement of energy generation did not achieve remarkable results due to the low range of optimization reachable and due to mandatory assumptions, that needed to be taken as the daily representative wind values.

The integration of the FCM method to assess the forecasts of wind directions it proved itself a powerful tool that require a very low computational time. With the simple knowledge of past weather data, it is possible to improve the quality of the data obtained. The decision taken along the training of the model influenced the results achieved and the model did not behave equally in all the months that it has been applied.
Higher performances of the machine learning model have been encountered in longer lead time estimation of weather condition, where the uncertainties of the forecast rise.

The integration of the data-driven model for the AEP investigation shown more similar energetic outcomes to the assumed real climate conditions reaching an improvement of more than 1% for some cases. Such results are reached even if the validations of the parameter estimated did not demonstrated an improvement of estimation of the wind direction.

The validations of the data achieved with the use of the FCM in the O&M instead, shown an increase of the precisions that reached more than the 3% of RMSE in the 3-day lead time forecast case. Such improvement is reduced when evaluated the energetic losses due to the influence of the uncertainty related to the wind speed. In fact, the calculations are highly affected by this parameter due to the cubic relation that has with the power computation.

The sensitivities analysis conducted give a wider understand of the strategy applied with the possibility to feature the schedule with different approach methods, increasing the safety of the decisions.

The use of an hindcast as referenced data introduce a bias error due to the imprecision of the model used for the presumption of the past climate as shown in the investigation. Such assumption it has been taken due to lack of real observation for the location studied. Anyway, this aspect does not compromise the validity of the work, as the model has been trained on the same data used for the validation it shows the potentials that a data-driven method could have in assess the uncertainties relying on past data.

A possible future development can be achieved testing the strategy with real weather observations authenticating the results obtained. Furthermore, more complex data-driven method could be applied on the model focusing on regressive algorithm that are able to give specific estimations based also on various other parameters. Such integrations can be used for the SWH and wind speed estimations that, even if do not influence the optimal combination of turbines to operate, greatly affect the energy losses and the decision maker. In final, wider consideration can be taken in the feature of the non-binary algorithm, deciding to operate each turbine only one time a month or modifying the wind speed limits across the months analyzed. This strategy modification will influence the final results obtained from the investigation variating the approach aiming to other objectives.

This thesis can be considered as a first step in the O&M scheduling planning using the contributions of data-driven method. The availability of the most precise data possible is a necessity to plan and manage interventions in an OWF. The integration of methods that take into considerations past data and improve the inputs available rise the security of interventions and the ensure economic investments. Moreover, the understanding of the variations between the expected outcomes and the occurred obtained in consequence of some decisions, helps in increasing the knowledge in OWF field.

Bibliography

[1] A. M. R. Lourenço, *Advanced Operation and Maintenance Strategies Considering Turbine Interaction - Impact on Wind Farm Economics,* Lisbon,: Instituto Superior Técnico, 2023.

[2] C. Bussolati, *Advanced Wind Farm Operation and Maintenance Strategies Considering Turbine Interaction,* Lisbon: Instituto Superior Técnico, 2023.

[3] IEA, "World Energy Outlook 2023," 2023.

[4] IPCC, "Climate Change 2022, Mitigation of Climate Change," 2022.

[5] I. Tiseo, "Statista," 2023. [Online]. Available: https://www.statista.com/statistics/500524/worldwide-annual-carbon-dioxide-emissions-by-select-country/.

[6] IEA, "Energy Technology Prospective 2023," 2023.

[7] E. Institute, "Statistical Review of World Energy (2023)," 2023.

[8] E. Commission, *The European Green Deal,* 2019.

[9] Ember, "European electricuty review 2024," 2024.

[10] O. W. i. Data, "Global Carbon Budget," 2023.

[11] REN, "REN Data Hub," [Online]. Available: https://datahub.ren.pt/pt/eletricidade/balanco-mensal/?date=2023-12-31.

[12] P. M. Council, *Plano Nacional Energia e Clima 2030,* 2020.

[13] G. o. Portugal, "NATIONAL ENERGY PLAN AND CLIMATE 2021-2030 Update/Revision," 2023.

[14] Mingyang, "Mingyang Smart Energy," [Online]. Available: https://www.myse.com.cn/en/wind-turbine/index.aspx. [Accessed 03/2024].

[15] Iren, "Iren," 07 2022. [Online]. Available: https://www.gruppoiren.it/it/everyday/abcircular/2022/o-come-offshore-cosa-sono-gli-impianti-eolici-offshore-e-come-funzionano.html#:~:text=I%20vantaggi%20dell'offshore%20(e,cosa%20c'%C3%A8%20da%20migliorare)&text=Inoltre%2C%20grazie%20alla%20maggiore%20s. [Accessed 03 2024].

[16] DNV, "DNV," DNV, 2023. [Online]. Available: https://www.dnv.com/focus-areas/floating-offshore-wind/index. [Accessed 2024 03].

[17] T. Group, "DeepWater Buoyancy," 07 2023. [Online]. Available: https://deepwaterbuoyancy.com/comparing-fixed-and-floating-offshore-wind-turbines/#:~:text=Floating%20turbines%20are%20particularly%20suitable,locations%20inaccessible%20for%20fixed%20turbines.. [Accessed 03 2024].

[18] T. S. a. P. D. -. N. R. E. Laboratory, "Cost of Wind Energy Review," 2021.

[19] A. H. S. B. E. R. M. H. D. G. Emma C. Edwards, "Evolution of floating offshore wind platforms: A review of at-sea devices," *Elsevier,* vol. 183, 2023.

[20] O. Bagherieh, *Gain-scheduling control of floating offshore wind turbines on large platforms,* 2011.

[21] Y. T. J. W. Warit Werapuna, "Wind Shear Coefficients and their Effect on Energy Production," *Energu Procedia,* vol. 8, pp. 1061-1066, 2017.

[22] L. Anjum, "Wind Resource Estimation Techniques-An Overview," *International Journal of wind and Renewable Energy,* vol. 3, pp. 26-38, 2016.

[23] B. A, "The Maximum of the Theoretically Possible Exploitation of Wind by Means of a Wind Motor," *Wind Engineering,* vol. 37, no. 4, pp. 441-446, 2013.

[24] D. S. N. J. a. E. B. T. Burton, Wind Energy Handbook, John Wiley & Sons, 2001.

[25] J. M. A. R. J.F. Manwell, Wind Energy Explained, Wiley, 2009.

[26] S. V. P. V. B. P. Parikshit G. Jamdade, "Assessment of Power Coefficient of an Offline Wind Turbine Generator System," *International Journal of Engineering Research & Technology (IJERT),* vol. 2, 2013.

[27] F. B. a. M. Cathelain, "An alternative form of the super-Gaussian wind turbine wake," *Wind Energy Science,* vol. 5, p. 1225–1236, 2020.

[28] S. R. K. &. Z. A. Voutsinas, "On the Analysis of Wake Effects in Wind Parks. Wind Engineering," *Wind Engineering,* vol. 14, no. 4, p. 204–219, 1990.

[29] S. H. A. H. A. Zervos, "Three-dimensional free wake calculation of wind turbine wakes," *Wind Engineering and Industrial Aerodynamics,* vol. 27, pp. 65-76, 1988.

[30] L. P. C. a. F. P. ´e-Agel, "A Wind-Tunnel Investigation of Wind-Turbine Wakes: Boundary-Layer Turbulence Effects," *Boundary-Layer Meteorology,* vol. 132, no. 1, pp. 129-149, 2009.

[31] H. Y. Haiying Sun, "Study on three wake models' effect on wind energy estimation in Hong Kong," *Energy Procedia,* vol. 145, pp. 271-276, 2018.

[32] M. B. P. F. G. N. a. J. R. Maria Krutova, "Evaluation of Gaussian wake models under different atmospheric stability conditions: Comparison with large eddy simulation results," *Journal of Physics: Conference Series ,* vol. 1669, 2020.

[33] "Analytical Modeling ofWind Farms: A New Approach for Power Prediction," *Energies,* vol. 9, 2016.

[34] F. P.-A. Majid Bastankhah, "A new analytical model for wind-turbine wakes," *Renewable Energy,* vol. 70, p. 116–123, 2014.

[35] J. H. A. Crespo, "Turbulence characteristics in wind-turbine wakes," *Wind Engineering and Industrial Aerodynamics,* vol. 61, p. 71 85, 1996.

[36] P. F. R. K. L. A. M.-T. C. J. B. R. M. a. E. S. J. King, "Controls-Oriented Model for Secondary Effects of Wake Steering," 2020.

[37] J. K. K. D. E. S. J. R. A. S. P. M. J. K. L. P. Fleming, "Initial results from a field campaign of wake steering applied at a," *Wind Energy Science,* p. 273–285, 2019.

[38] J. K. C. J. B. E. S. R. M. N. H. A. F. a. P. Fleming, "Overview of FLORIS updates," *Journal of Physics: Conference Series,* vol. 1618, no. 2, 2020.

[39] J. K. E. S. J. R. A. S. P. M. J. K. L. P. M. P. Fleming, "Continued results from a field campaign of wake steering applied at a commercial," *Wind Energy Science,* vol. 5, no. 2, p. 945–958, 2020.

[40] I. Katic, J. Højstrup and N. Jensen, "A Simple Model for Cluster Efficiency," *EWEC'86,* vol. 1, pp. 407-410, 1987.

[41] L. I. R. B. C. L. B. Filippo Campagnolo, "Further calibration and validation of FLORIS with wind tunnel data," *Journal of Physics Conference Series,* vol. 2265, no. 022019, 2022.

[42] A. S. V. Y. L. d. J. J. T. Z. J. Zhengru Ren, "Offshore wind turbine operations and maintenance: A state-of-the-art review," *Renewable and Sustainable Energy Reviews,* vol. 144, 2021.

[43] R. Buizza, *Sources of uncertainties,* 2014.

[44] R. G. H. T. D. Owens, "ECMWF Forecast User Guide," 2018.

[45] D. Richardson, *Ensemble forecasting,* 2015.

[46] D. M. I. D. M. R. J. D. J. Q. a. K. W. V. M. Catterson, "An economic impact metric for evaluating wave height forecasters for offshore wind maintenance access," *Wind Energy,* vol. 9, pp. 199-212, 2016.

[47] A. K. D. I. Ravi Kumar Pandit, "Data-driven weather forecasting models performance comparison for improving offshore wind turbine availability," *IET Renewable Power Generation,* vol. 14, pp. 2386-2394, 2020.

[48] H.-q. T. X.-f. L. Y.-f. L. Hui Liu, "Wind speed forecasting approach using secondary decomposition algorithm and Elman neural networks," *Applied Energy,* vol. 157, pp. 183-194, 2015.

[49] R. Z. F. L. L. Z. Q. L. Yun Wang, "A review of wind speed and wind power forecasting with deep neural networks," *Applied Energy,* vol. 304, p. 117766, 2021.

[50] L. Zadeh, "Fuzzy sets," *Information and Control,* vol. Volume 8, pp. 338-353, 1965,.

[51] R. E. W. F. James C. Bezdek, "FCM: The fuzzy c-means clustering algorithm," *Computers & Geosciences,* vol. 10, pp. 191-203, 1984.

[52] C. C. Mitsubishi Corporation, "Mitsubishi Corporation," Mitsubishi Corporation, Chiyoda Corporation, 2015. [Online]. Available: https://www.mitsubishicorp.com/jp/en/pr/archive/2015/html/0000028944.html. [Accessed April 2024].

[53] H. B. B. B. P. B. G. H. A. M. S. J. N. J. P. C. R. R. R. I. S. D. S. A. S. C. D. D. T. J.-N. Hersbach, "ERA5 hourly data on single levels from 1940 to present," Copernicus Climate Change Service (C3S) Climate Data Store (CDS), 2023. [Online]. Available: 10.24381/cds.adbb2d47. [Accessed April 2024].

[54] Anker, "Anker," Anker, June 2023. [Online]. Available: https://www.anker.com/blogs/home-power-backup/electricity-usage-how-much-energy-does-an-average-house-use. [Accessed March 2024].

[55] Z. G. Mengning Wu, "Methodology for developing a response-based correction factor (alpha-factor) for allowable sea state assessment of marine operations considering weather forecast uncertainty," *Marine Structures,* vol. 79, no. 103050, 2021.

[56] N. R. E. Laboratory, *Turbine Library Interface*.

[57] B. Sanderse, "Aerodynamics of wind turbine wakes, Literature review," *ECN,* vol. 9, 2009.

[58] P. M. Sforza, "Chapter 10 - Propellers," in *Theory of Aerospace Propulsion (Second Edition)*, Pasquale M. Sforza, 2017, pp. 487-524.