Ecole Polytechnique Fédérale de Lausanne

Master Thesis:

Evaluation of the wind droughts at a global seasonal scale: variability, predictability and drivers

Author: Guth Basile

Under the supervision of Daniele Mari [1], Francisco J. Doblas-Reyes [2], Verónica Torralba [3], Jean-Marie Fürbringer [1]

Barcelona, 16th September 2022

With the support and resources of the Barcelona Supercomputing Center - Centro Nacional de Supercomputación

[1]Ecole Polytechnique Fédérale de Lausanne
[2]Barcelona Supercomputing Center - Centro Nacional de Supercomputación
[3]Euro-Mediterranean Center on Climate Change (CMCC)

Abstract

Facing the expansion of wind energy use and production, it becomes necessary to evaluate the current forecast systems ability to predict low wind speed events (or wind droughts). Moreover, identifying the drivers of these events might be used to improve their predictability. This thesis is the first to investigate which is the forecast system's ability to predict wind droughts. Two indices (Below Threshold and Mean Below Mean Threshold days) have been created to quantify the occurrences of the number of wind drought days in a season. These indices are based on the 10^{th} percentile mobile threshold on the daily near-surface (10m) wind speed values. Two datasets from the European Centre for Medium-Range Weather Forecasts have been employed: the ERA5 reanalysis is taken as an observational reference and the seasonal forecast system is SEAS5. The analysis focuses on the boreal winter (DJF) over the period from 1993 to 2016. The forecast quality assessment has been done in terms of Pearson's correlation coefficients and the Ranked Probability Skill Score. This work shows the limitations of SEAS5 to skilfully predict wind droughts. Furthermore, it has been found that there is not a linear relationship between seasonal mean wind speed and wind droughts. Finally, this work has revealed that the North Atlantic Oscillation is not playing a major role on the wind droughts over the North Atlantic domain.

Contents

Acronyms and abbreviations

CDAS := Climate Dynamics and Air-Sea interaction laboratory

CPC := Climate Prediction Center

DJF := December - January - February

GCM := Global Circulation Model

ITCZ := Inter-Tropical Convergence Zone

NAO := North Atlantic Oscillation

NOAA := National Oceanic and Atmospheric Administration

RPCA := Rotated Principal Component Analysis

RPS := Ranked Probability Score

RPSS := Ranked Probability Skill Score

WMO := World Meteorological Organisation

Definitions

Members: A forecast model runs several simulations in order to do a prediction. It gives a probability of results of the forecast. These simulations are called members and form an ensemble.

Climatology: Long-term average map of chosen variable computed over a baseline period. For example, the wind speed monthly climatology for January can be computed as the average wind speed for all the January months in the 1981-2016 period. The climatology maps are computed according to the following equation:

$$Climatology(lat, lon) = \frac{1}{N_{period}} \frac{1}{N_{member}} \Sigma_{i \in period} \Sigma_{j \in members} Variable(lat, lon, i, j)$$

where N_{period} is the number of years in the baseline period and $N_{members}$ is the number of members of the considered model. Remark, when the climatologies are computed for the variables obtained from the reanalyses, the average is only computed over the time dimension, as the observational datasets do not contain probabilistic information (i.e. ensemble members). Lat and lon represent respectively the latitude and the longitude of a grid point.

Composite: Average map of the chosen variable for all times that a given condition occurs [1]. For example, a pressure composite can be computed as the pressure average map on Switzerland over those months for which a storm is present over Lausanne.

$$Composite_R(lat, lon) = \frac{1}{N_{period_R}} \frac{1}{N_{member}} \Sigma_{i \in period_R} \Sigma_{j \in members} Quantity(lat, lon, i, j)$$

The difference between the climatology and the composite is the period used for the computation of the mean. For the composite, only those time steps satisfying a particular condition (days or months for which the NAO is positive) are averaged.

Starting date First day of a forecast system. The system is initialised with observed values and is in general launched over 6 months

Forecasts and Hindcasts As a forecast is a simulation into the future and tries to give new information to the user, a hindcast is a forecast done on previous years. In general hindcasts are performed to be compared with other systems or reanalysis.

1 Introduction

1.1 Motivation

Through the release of carbon in the atmosphere, Humanity acquired piece by piece some independence from the stochasticity of nature. However, a few thousands years after the first intentional fire, scientists, among whom Callendar [2] was the first to have mentioned the effect, discovered that a too fast and consequent release of greenhouse gases in the atmosphere was not a good idea. Consequently, modern society energy production turns back partly to the abundant solar and wind energy in order to limit global warming and its risks. According to Hannah Ritchie et al. [3], the solar and wind world energy shares went from respectively 0.01 and 0.17% to 1 and 2 % in fifteen years, with major increases in western countries. However, an increase in solar and wind also signifies an increase in climate dependence, hence, an increase in the intermittency of the energy production. Two subjects need to be studied: the storage and the predictability of the source. Both are complementary since the first offers an active solution to fill gaps where the energy production is not available and the latter gives information on the sizing, improves the decision-making and helps the optimisation on the electricity generation.[4]

One example of this situation is the summer and early autumn of 2021 when Europe was affected by a prolonged period with below-normal wind speeds (Bloomfield, 2021) [5]. In this context, the estimation of the frequency and duration of extreme winds and the anticipation of these persistent low wind speed episodes have become essential for the wind industry to mitigate the influence of wind variability in their activities." Climate predictions [...] have witnessed considerable improvements in the last decade, demonstrating that probabilistic forecasting can inform better decision-making for some forecast windows and regions." [4]. logistics. [6]

Indeed, seasonal climate forecasts can bring information about energy generation and consumption in specific locations, and therefore they support the management of energy supply/demand balance preventing the risks of blackouts. Wind variability information allows choosing the most adequate period to schedule maintenance of wind power plants and offshore

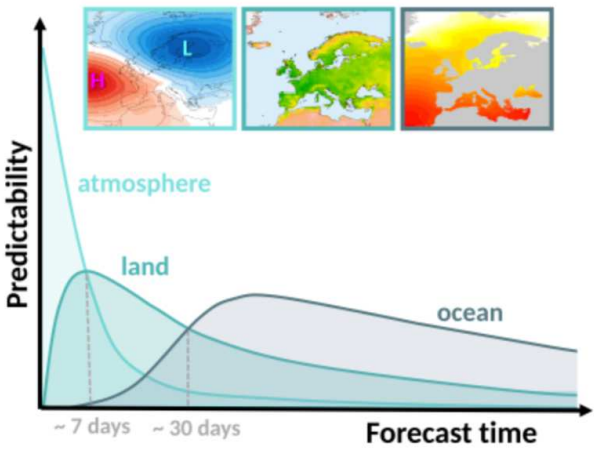


Figure 1: Illustration of the predictability change according to its nature and the forecast time. Figure from Verónica Torralba-Fernandez 2019 [1], adapted from Mariotti et al 2018, [8]

Although the climate system is known to be chaotic (Lorentz, [7]), climate predictability is supported by slower processes that need less precise initial conditions (illustrated by Figure 1). "Current climate prediction systems can provide accurate information [...] with lead times up to nine months" [9]. However, the ability of climate models to predict low wind speed events sufficiently in advance has not yet been examined. This work is intended to fill this gap of

knowledge, and it has two main aims: first, it will focus on the investigation of the recorded wind drought episodes in order to determine the physical phenomena driving these extreme events at the global scale. Secondly, it will evaluate the ability of the currently available seasonal forecast systems to predict wind droughts in advance.

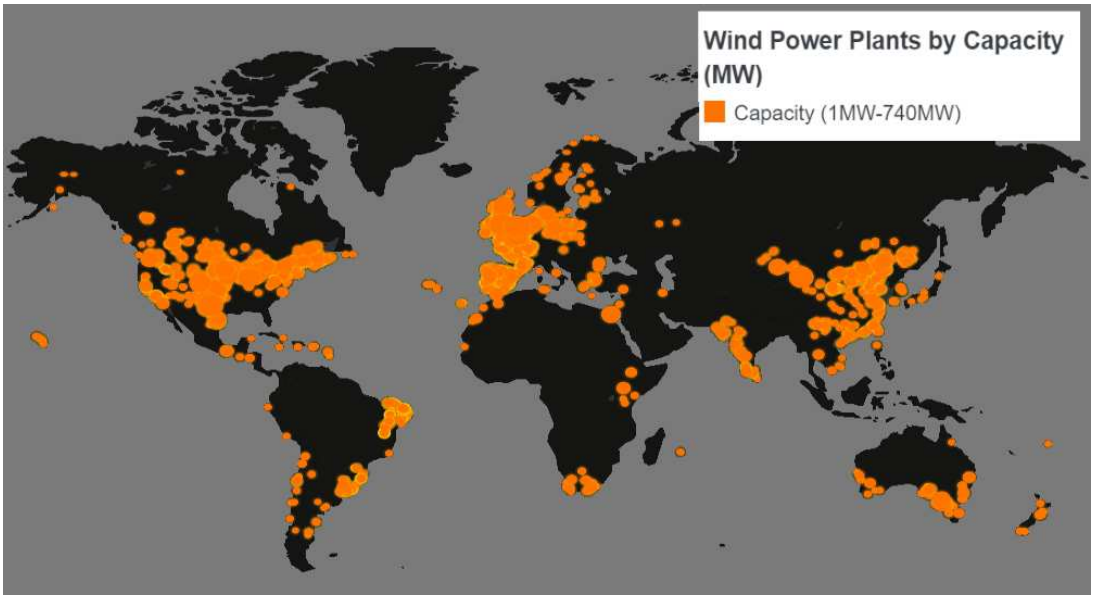


Figure 2: Wind power plant distribution around the world. They are mainly situated on land. Taken from RessourceWatch, source from the Global Energy Observatory of the KTH Royal Institute of Technology of Stockholm [10]

The first introduction sections purpose is to present the basis concepts necessary for the understanding of this work.
As an observational data-set, this work will use the ERA-5 reanalysis and the seasonal forecast system SEAS5 will be studied. In order to study low wind speed events, one has to give a definition on what is a wind drought. Two indices are then defined, the Below Threshold and the Mean Below Mean Threshold days. The results will then be analysed thanks two statistical tools, the Pearson's correlation coefficient and the Ranked Probability Skill Score (RPSS), a widely used tool to assess forecasts. All those metrics are explained in the section 2 (Data and Methods).
The results are then divided in three parts. The first will focus on the behaviour of the two created wind drought indices. Secondly, the evaluation of the SEAS5 forecast skill is done. The last part concerns the wind drought behaviour with the NAO index.
Finally, all the conclusion are summed up in the final section.

1.2 Surface wind in climate system

Because of Earth's shape, the solar irradiance is more intense on the equator than on the poles. The Earth also emits its black body radiation that cools down the atmosphere. As the atmosphere is a quasi-steady state, there is more radiation emitted than absorbed in the poles and the opposite occurs in the equator, where the limit between the two is approximately on the 40 and -40 latitudes.

The heat is then released on the poles and a gradient of temperature occurs on Earth. Consequently, if there were no rotation on our planet, there would be a constant wind from the equator to the North. [11]
The rotation brings the Coriolis' effect, which adds a rotation pattern to the wind that can't reach the poles. Instead, convective cells are observed.

The theory characterises the global atmospheric circulation with a division of each hemisphere in three cells: the Hadley, Ferrel and Polar cells.

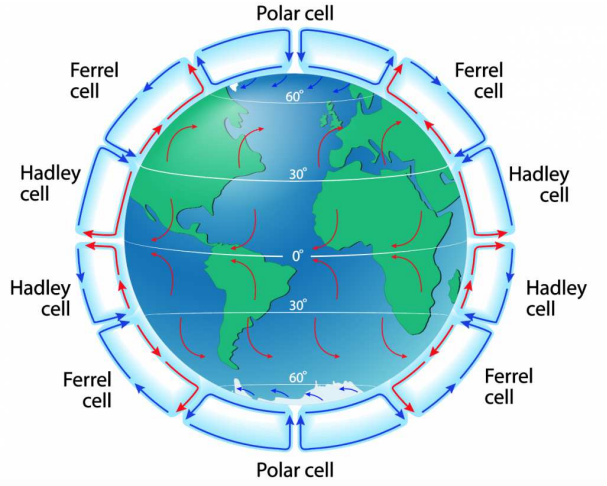


Figure 3: Illustration of the atmospheric convective cells [12].

The Hadley Cell: At the equator, or more precisely, in the Inter-Tropical Convergence Zone (ITCZ), the solar radiation warms the surface, and as a consequence, the air is ascending. By taking altitude the air cools which leads to the condensation of the water vapour and a release of its latent heat. The liberated heat of the water warms up the now dried air that keeps raising while the water rains down. This mechanism will take place until that the air is completely dry. The atmospheric level reached is the tropopause. While elevating, the air mass is already moving to the poles, then, the Coriolis effect implies a movement to the East. However, this movement is counterbalanced to the West by the fact, that while elevating, the air mass moves away from the Earth centre and by the conservation of the angular momentum, it slows the velocity. The result is that during the elevation the motion of the air follows the latitudes without consequent deviation.

At the tropopause, the air is cooled down by radiative cooling [13], slowly reducing its altitude. Coriolis' effect and the gain in movement speed create a movement eastward that reaches the latitude of $\pm 30°$. The depression created by the elevation of the air and the over-pressure created by its arrival engender a pressure gradient and thus, surface wind.

The Polar Cell: The atmosphere on the poles is so cold that all the air "falls" on the ground. The pressure increases consequently and then, pushes the air out of the poles. At the limit between the Ferrel and the Polar cells the air rises and a convective scheme takes place.

The Ferrel Cell: Ferrel Cells are intermediary cells doing the gear between the two above-mentioned cells. "The winds blowing towards the higher latitudes pick moisture from the oceans, and they meet cold air that has drifted from the poles at around 60° latitude. The convergence of these two air masses creates an area of unstable weather conditions associated with the mid-latitude depression" [14]. That is defined as baroclinic instability. Ferrel cells, unlike the two other cells, are thermally indirect. It comes from the fact that the temperature circulation inside the cell is reversed. Indeed, the cold air on the $\pm$ 60o is ascending while the air on the $\pm$ 30 ° latitude is

descending. This is so because the circulation is driven by the motions of the cells on either side i.e. the Hadley cells and the Polar cells. [15]

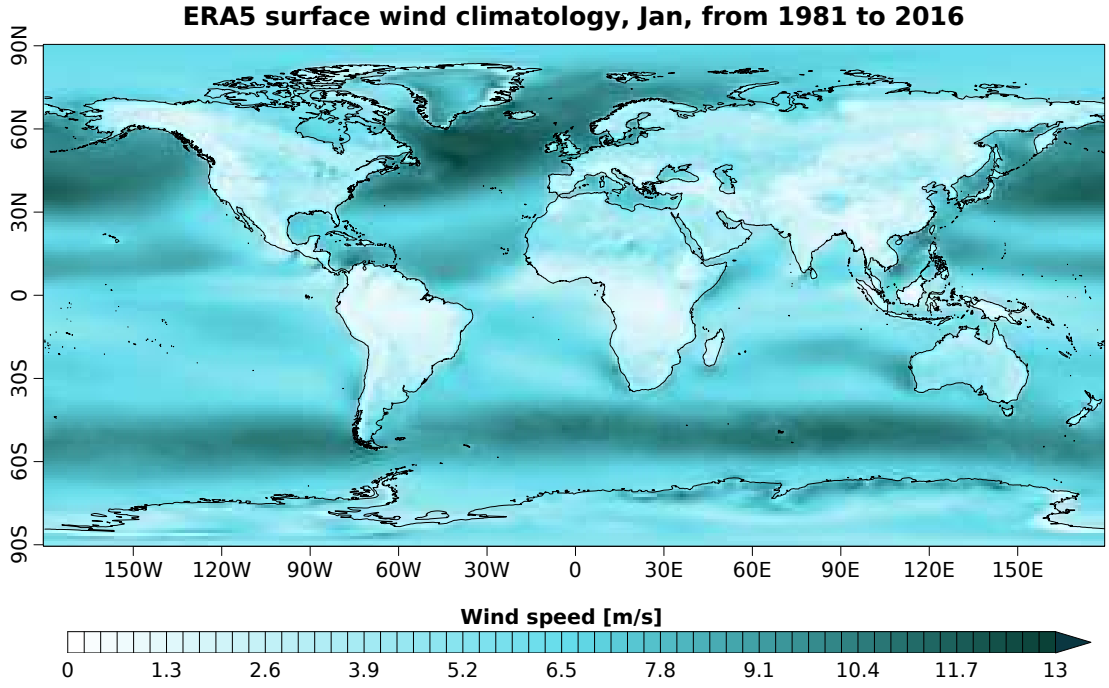


Figure 4: Wind speed climatology in January, from ERA5 monthly values.

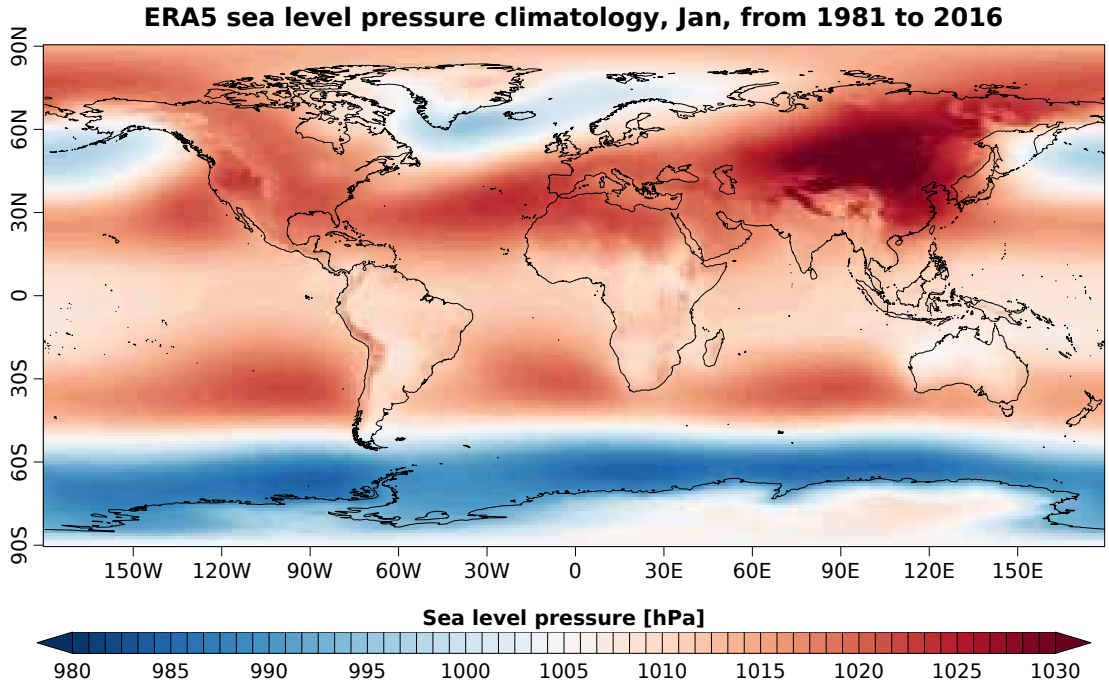


Figure 5: Pressure climatology in January, from ERA5 monthly values.

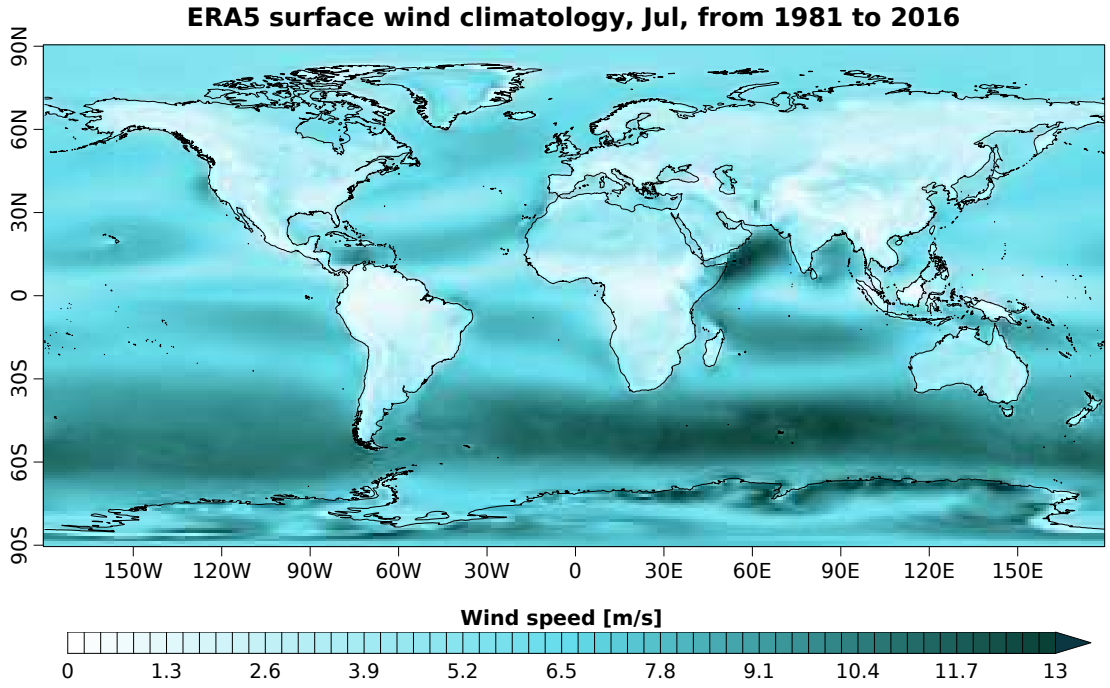


Figure 6: Wind speed climatology in July, from ERA5 monthly values.

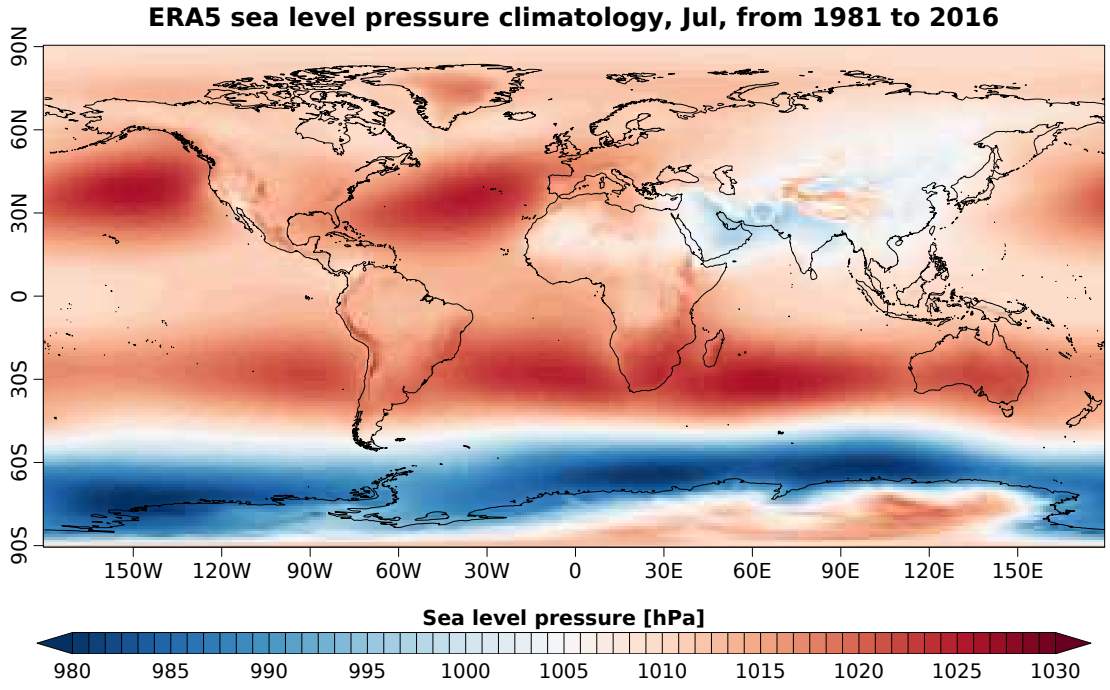


Figure 7: Pressure climatology in July, from ERA5 monthly values.

Figures 4, 5, 6 and 7 present the climatologies of the wind speed at 10m over- and the pressure on sea level for January and July. They are characterising the average seasonal behaviour of both atmospheric variables (i.e. wind speed and pressure) as January and August are good proxies for

the season. Indeed, these two months are good proxies for the seasonal behaviour, indeed even though the solstices are in late December and June, the atmospheric inertia brings the state at his extremes in mid-July and January [16].

By looking at Figure 4 and 6 one can see that the regions where a high climatological wind is observed, it corresponds in the climatological pressure to a high gradient zone (short distance of pressure colour change). Of course, surface roughness also plays a major role on the wind speed, which is shown by the speed values over the land being significantly lower than over the oceans. Figures 5 and 7 illustrate the spatial distribution of the sea level pressure. These patterns are associated with the convection cells, where a low pressure zone corresponds to the location where the air is elevating, and a high pressure zone correspond to the region where the air is descending. One can see that it is following a latitudinal distribution, which equatorial low pressure band host the above-mentioned ITCZ.

The comparison of these figures highlights the different climate conditions for Winter or Summer months. The ITCZ follows the "thermal equator" [17] , which is not completely centred on the equator during all the year. In January, it is shifted to the South of the equator, and in August it moves to the North. This is a consequence of the Earth obliquity, indeed, Damianos F. Mantsis shows that the Earth's obliquity, which is responsible for seasonal variability, imply a reduced meridional temperature gradient in the summer hemisphere than in winter and that by the thermal-wind relationship (Rind, 1998 [18]) respectively weaker/stronger westerlies [19]. Finally, one observation to made is the high climatological pressure on mountain locations.

1.3 Surface Wind in Northern Europe

The North Atlantic Oscillation (NAO)

Figure 8 shows the average pressure condition that occurs over North Atlantic in January. Because of the global atmospheric circulation, what we could identify as the beginning of the two Hadley cells (and the ITCZ) is bringing air from the equator down to the Azores, increasing the pressure locally. Moreover, between Iceland and Greenland air is ascending from the North Atlantic, causing a depression. A part of this air is also descending on the Azores. At the origin, this is the variation of this gradient that gives the first NAO index.

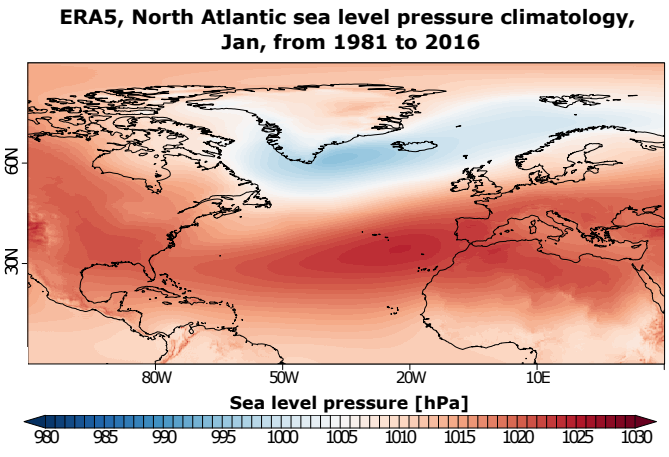


Figure 8: Sea level pressure climatology over the North Atlantic from ERA5 monthly values.

In Figures 5 and 7, one can observe that the January North Atlantic gradient is well stronger than in July. This aspect associated with the higher variability of NAO in winter is the reason why its influence is higher in this period and why the study will focus on it. It is responsible

for more than 35% of the total pressure variance in the North Atlantic sector, during Winter (DJF) [20].

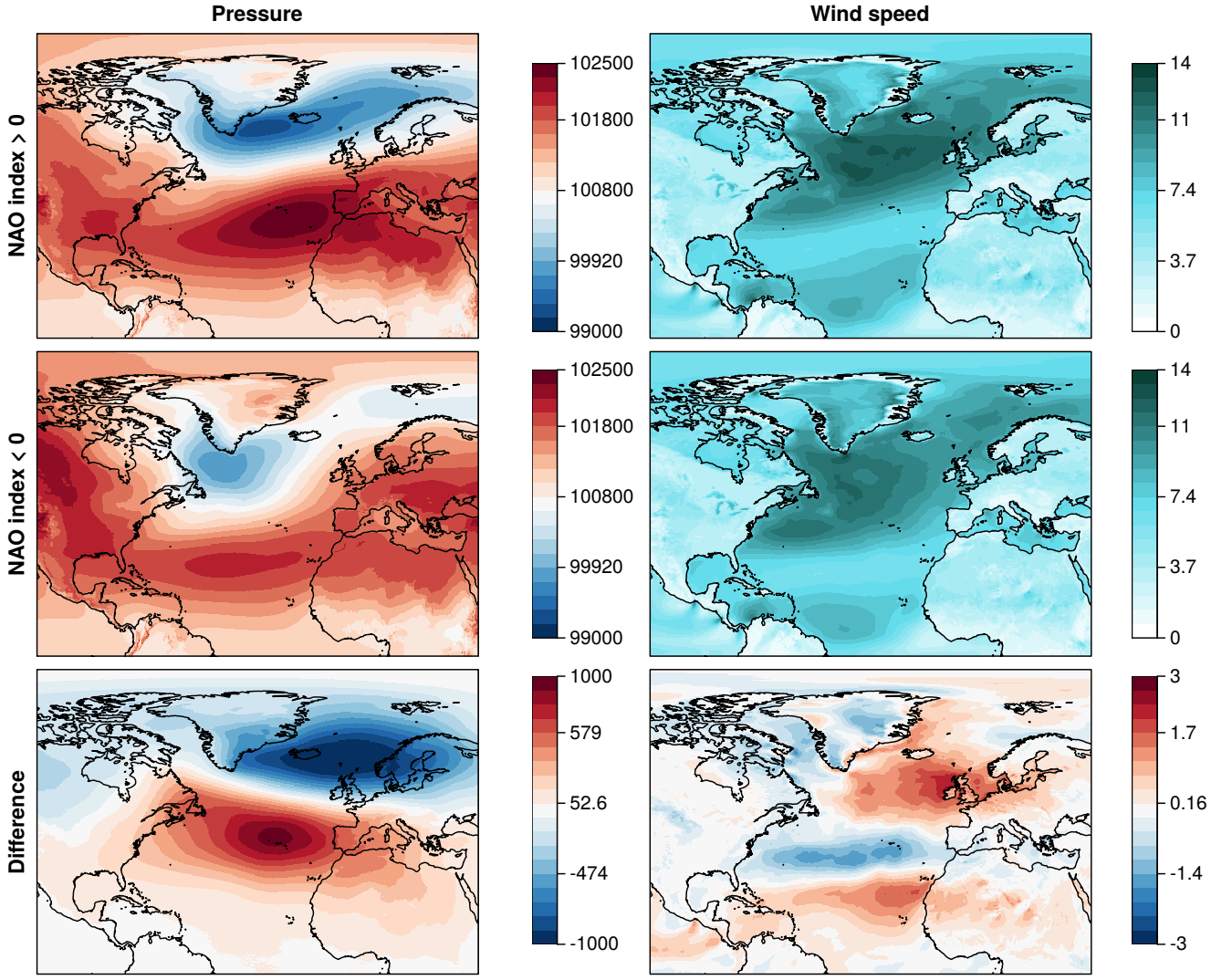


Figure 9: NAO composites from monthly mean values, ERA5. On the left are the pressure maps and in the on the right are the wind composites. The first row are the composites from a positive NAO and the second row from a negative NAO. Finally, the last row is obtained the difference between the positive NAO composite and the negative one.

More technically, Hurrell defines the NAO as : "The NAO is reflected in the spatial pattern of the two leading empirically determined orthogonal function (EOF) of the Northern Hemisphere boreal winter [...] but in order to see them clearly, it is necessary to rotate the EOFs in a manner that tends to simplify their spatial structure" [20], as it is done by Barnston and Livezey [21] and is called the Rotated Principal Component Analysis (RPCA). The RPCA is a Principal Component Analysis where a linear rotation transformation is applied to the components of the singular value decomposition.[22] The RPCA technique is applied to monthly standardized 500-mb height anomalies obtained from the CDAS (Climate Dynamics and Air-Sea interaction laboratory) in the analysis region (20° N-90° N) between January 1950 and December 2000. The standardised anomalies are calculated based on the 1950-2000 climatological daily mean and standard deviation. The result is then an index

that can theoretically go from $-\infty$ to $+\infty$. However, in practice, the index is most of the time fluctuating between -3 and +3.

The monthly values of the NAO Index from 1981 to 2016 used in this work are obtained from the Climate Prediction Center (CPC) of the North Oceanic and Atmospheric Administration (NOAA).

In other words, the North Atlantic Oscillation is a pressure oscillation pattern (first order mode in pressure variation) in the North Atlantic sector (20º-70ºN, 90ºW-40ºE). The composites of the Figure 9 shows the pressure and wind patterns associated with the sign of the NAO index. One can then observe the two patterns of the pressure oscillation. The pressure and wind conditions corresponding to the occurrences of the NAO are illustrated by the composites of Figure 9 for January. The increase of the NAO index intensifies the gap between the extremes: the highest pressure is higher, and the lowest pressure is lower when the NAO is positive. The positive NAO is also characterised by a shift of the high-pressure cell to the Iberian Peninsula. The low pressure zone is developing over Scandinavia. The changes in those pressure systems modify the

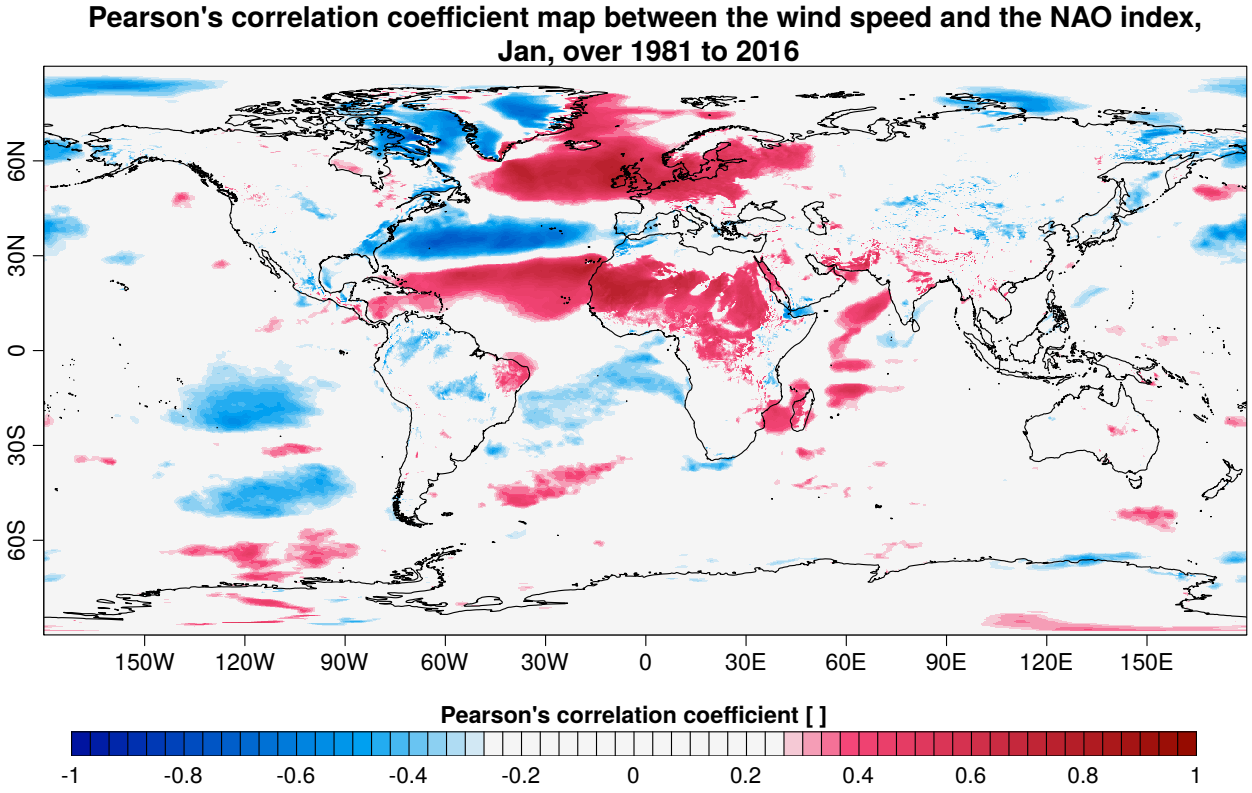


Figure 10: Correlations between the NAO index and the local wind speed monthly mean in January. As wind is almost linear with the pressure gradient, Pearson's correlations are considered as correct. In blue are represented negative correlations, and in red positive ones. Low absolute correlation values are set in white. The associated p-values to these correlations are given by the Figure 11

pressure gradients and the wind strength. For example, it is clear that the pressure gradient over the UK is increasing as it is situated between the reinforced high and low-pressure systems. The wind composite difference confirms this analysis [23]. The same conclusions can be

done over the Greenland Sea, Europa or middle Atlantic.

Figure 10 shows the correlations between the NAO index and the wind speed in January. A positive (negative) correlation means that when the NAO index is increasing, the wind speed increase (decrease) accordingly. Four highly correlated or anti-correlated zones are visible. The wind is a flow of air coming to compensate a depression from a high pressure zone, so thanks to the composite map (Figure 9) sorting the pressure according to the NAO index sign, and the climatology map pressure during January, it is easy to associate the negative correlations with a diminution of the local gradient pressure and the positive correlation with an increase of this gradient.

Finally, out of the North Atlantic sector some correlations are visible, symptomatic of atmospheric tele-connections.

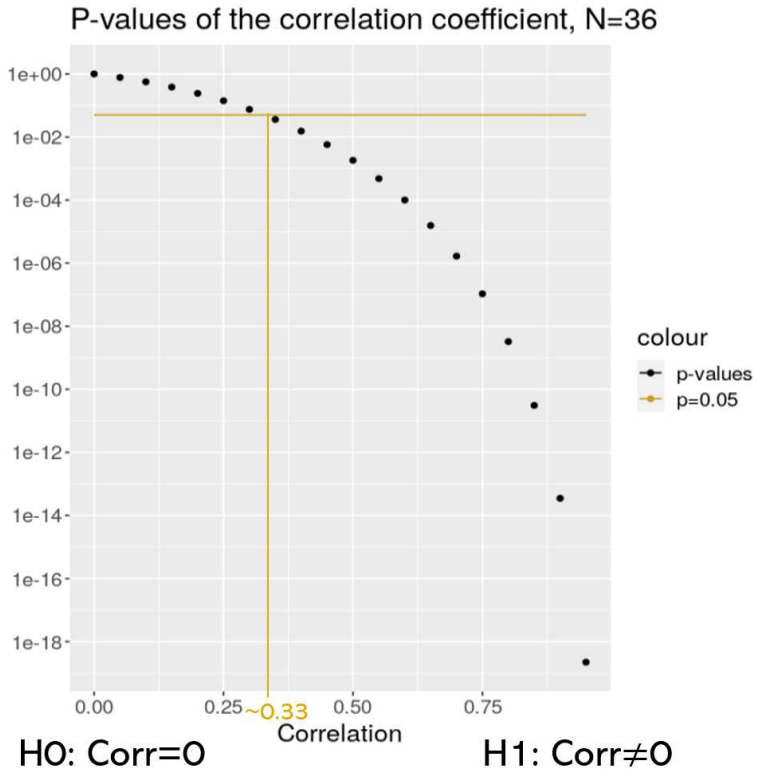


Figure 11: Two-tails p-values for a given correlation coefficient with 36 elements. In yellow is show the threshold of 95% significance.

North Atlantic Oscillation's predictability

NAO is responsible for more than 35% of the pressure variability (main driver of the wind) in its sector. The NAO index is extracted from an area, while the wind behaviour is computed point to point. There is then more hope that the general behaviour of the NAO index is better predicted than individual wind speed points.

To study the predictability of the NAO, it is convenient to have a look at its spectrum, showing the internal periods of oscillation.

Following the indication of Helen J. Wearing [24] the figures 12 and 13 have been produced. Those figures represent first the spectrum of the NAO index for various periods, smoothed with a modified Daniell kernel whose width is equal to 3. Indeed, this width has been chosen since the result is obtained from monthly values which are arbitrarily defined. For a given month, its monthly value could have been influenced by a phenomenon that took place a month before or could influence the month after.

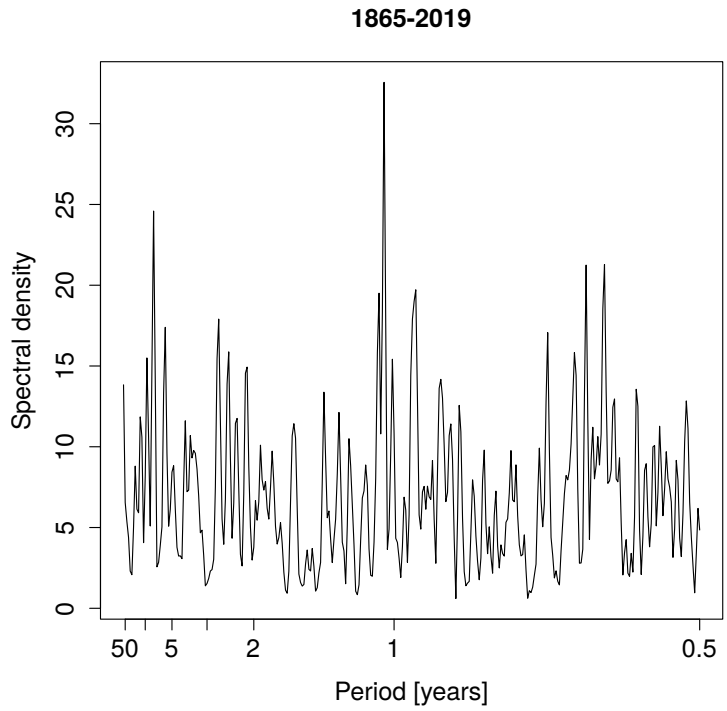


Figure 12: NAO index spectrum computed with monthly data over the 1865-2019 period from the national centre for atmospheric research (NCAR) [25], smoothed with modified Daniell kernel of months. On the logarithmic values is performed a linear fit that gives as an output

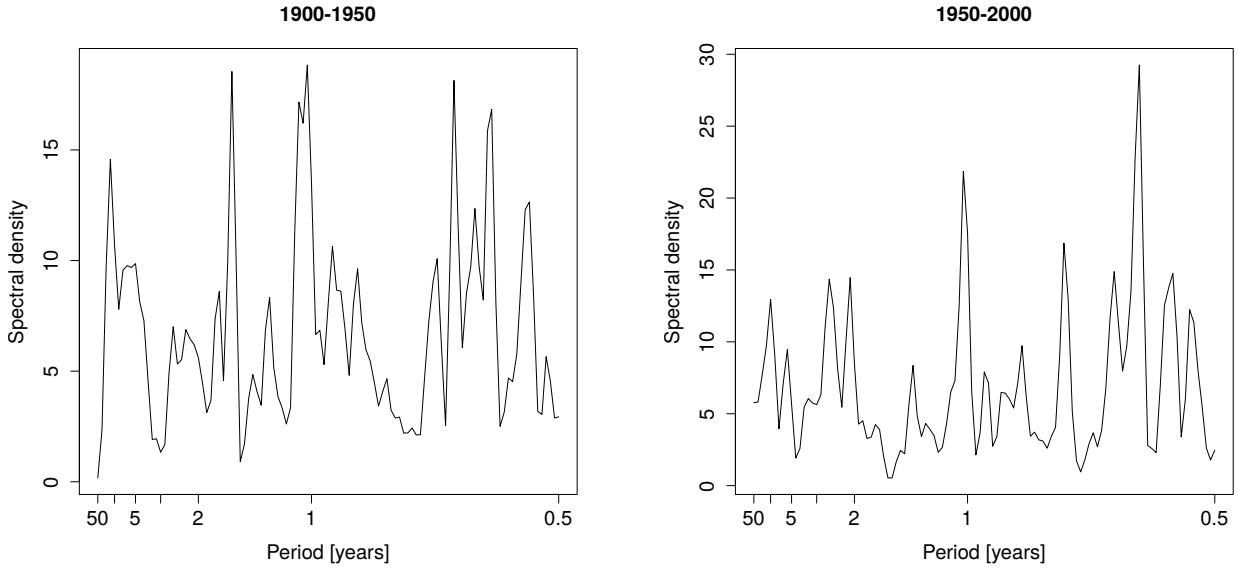


Figure 13: NAO index spectrum computed with monthly data on a fifty years old period, smoothed with 3 months width Daniell kernel.

Figure 12 shows the resulting NAO temporal index spectrum for a period from 1865 to 2019 which looks like a constant noise. There are no evident significance of the different peaks, and the geometry would be different if the selected period is different. Indeed, the figure 13 shows the

spectrum for two different periods, from 1900-1950 and 1950-2000. As expected, the two figures are rather different, confirming the precedent assumption. As a matter of fact, the NAO is influenced by a lot of different drivers that have a small impact [20]. Thus, it is unlikely that a simple model of the NAO could be developed, instead the most promising simulations are obtained by taking into account statistical and fluid physics in Global Circulation models (GCM), but are still not perfect. The predictability of the NAO by the GCMs which simulate the dynamics and state of the different layers of the Atmosphere and the Oceans, is limited [26].

However, from the Figure 13 some common characteristics appear. The most visible is the peak high spectral density on a period of one year, which highlights the seasonal sensibility of the NAO and which has already been mentioned in this work in section 1.3. Moreover, two other properties could be investigated, first the absence of any peak around a period of 1.3 years and the high spectral density around 0.6 year.

1.4 Wind records

1.4.1 Observations

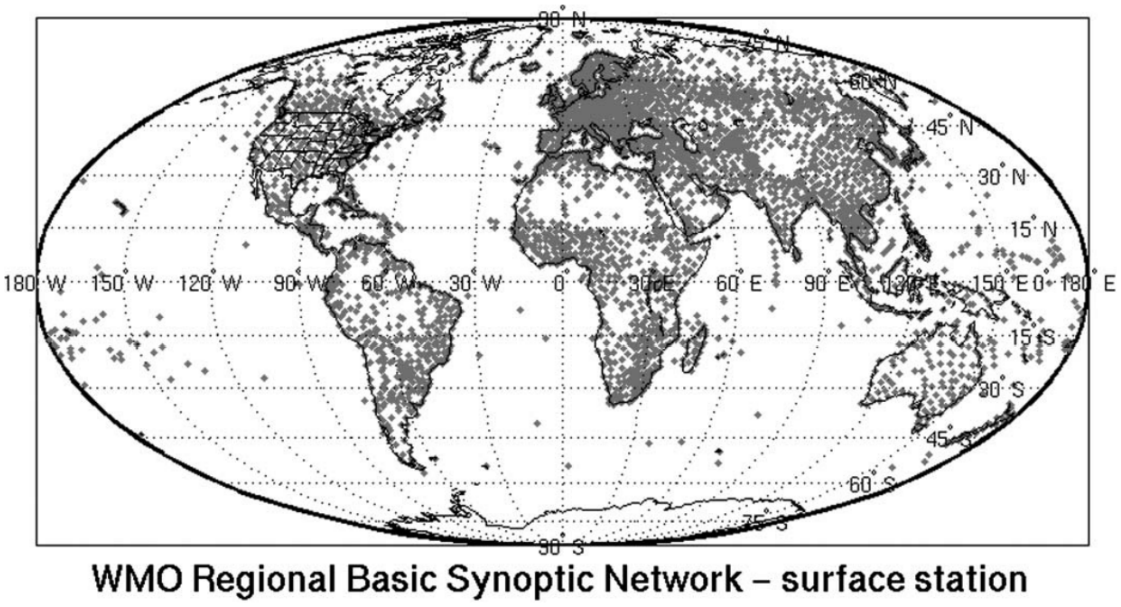


Figure 14: Global distribution of the World Meteorological Organisation (WMO) surface stations recording wind data in January 2012. Figure from Liu et Al., 2014 [27].

Wind is one of the most known and observed meteorological variables, however, the amount of observations is scarce. The main reason is the need of this variable to be recorded in-situ. There is no actual way to obtain global observations. Moreover, wind is almost only observable from surface, at the exceptions to some balloons observations. As an example, the global distribution of the World Meteorological Organisation (WMO) surface stations recording wind data in January 2012 is shown by Figure 14. It appears that the amount of observations is higher for developed

countries. These difficulties result in a very unequal and unsystematic distribution of the wind observation in the world, and particularly reduced on the ocean.

1.4.2 Reanalysis

Due to the lack of observational data in most parts of the world, the observational reference currently employed for research and also industrial activities are the **reanalyses**.

A reanalysis is the merging of observation and modelling, it provides wind speed values over the whole globe with variable spatial resolution. The data obtained through observation is used for their defined location and altitude, then the grid points for which no data is available is computed thanks to an internal model (illustration with Figure 15). These operations allow us to have a complete dataset of wind records over several decades (with ERA5, since 1955).

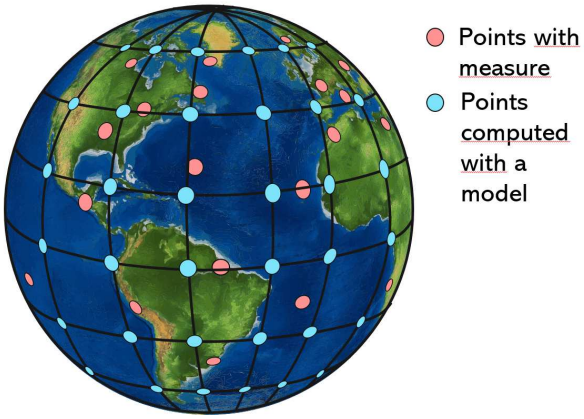


Figure 15: Illustration of how is working a reanalysis. Observations are illustrated in pink and are unequally distributed around the world, with a higher frequency on land than on the oceans. The blue points, which form a grid, are then computed with the help of the observations and a model.

1.4.3 Seasonal forecast systems

It has already been mentioned in the Introduction that seasonal climate can be predicted up to six months into the future. Therefore, most of meteorological and climate research centre are developing their own system and there are numerous of them (Meteofrance, der Deutscher Wetterdienst, the United Kingdom meteorological office...).

The forecast systems are composed of several quasi-independent models that focus on a specific part of the climate system. The models run independently, and the outputs are used as initialisation for the other models. For example, the ECMWF system is subdivided into different components such as high-atmosphere, land and oceanic surface. The grid cut covers the Earth's surface by a given number of smaller areas. In each subsurface, the dynamic is based on a set of equations describing the processes relative to the environment. In the case of the atmosphere, the grid is covering the Earth with almost identical isosceles triangles except at the equator, which is made of squares. The vertical axis (perpendicular to the surface) has a coarse layer density at the crust than at the top of the atmosphere.[28]

Forecast systems are going forward in time and therefore, they are very sensitive to the initialisation. Each simulation is independent, and even a small change in the initial state can lead to a rather different forecast [29]. This is why forecast systems are generating an ensemble of simula-

tions, called members, with slight differences in the initialisation in order to provide the forecasts with an estimation of the uncertainty. The analysis of the forecasts is always taking all results into account. In general, they are used to produce tercile probabilities. The three categories are "above normal", "normal" and "below normal" and usually represent 33% of the usual outcomes. [30] Therefore, the forecasts give an indication of how probable are specific climate conditions in the upcoming season. [31]

2 Data and Methods

2.1 The choice of the ERA5 reanalysis as the observational reference

Numerous reanalyses exist, and it is not possible to de fine the best from a general point of view, the different models having difference in their performance according to the variable. The choice has been done according to the work done by Ramon et Al. [32] which shows that ERA5 reanalysis provides the most accurate wind speed representation of the wind observations.

ERA5 (ECMWF ReAnalysis 5) is developed by the ECMWF (European Centre for Medium-Range Weather Forecast) and replaces the previous ERA-Interim reanalysis, which stopped being produced on 31 August 2019.
ERA5 provides hourly estimates numerous atmospheric, land and oceanic climate variables. The data covers the Earth on a 30km grid and resolves the atmosphere using 137 levels from the surface up to a height of 80km. ERA5 includes information about uncertainties for all variables at reduced spatial and temporal resolutions. ERA5 provides "quality assured" data from 1979 to the present (the data are delivered with a delay of three months) and a preliminary dataset from 1950 to the present (with a delay of five days). [33]

2.2 The SEAS5 as the seasonal forecast system

One would like to select the best forecast system available, however because of the lack of complete observations. It is needed to use reanalysis, which is of course not perfect and have bias. A complete work should include several reanalyses and a multimodel comparison. However, because of the limited time of this work only one of each is taken showing a first sketch of the methodology that can later be applied more widely. In order to have the best chance that the forecast system predicts the same as the reanalysis, it has been chosen to choose the system with has the biggest potential of having the same, i.e. from the same institution. Which is SEAS5.

SEAS5 is the fifth (and most recent) operational seasonal forecast system developed by the ECMWF. SEAS5 uses IFS Cycle 43r1 and the community ocean model NEMO (Nucleus for European Modelling of the Ocean). The ocean resolution is 0.25 degrees and 75 depth layers in SEAS5 (ocean model configuration: ORCA025z75). The vertical resolution is particularly high in

the uppermost part of the ocean, with an increase in the number of levels in the first 50 metres from 5 to 18. The sea-ice model is LIM2, part of the NEMO modelling framework. SEAS5 ocean and sea-ice initial conditions are provided by the new ocean analysis and reanalysis ensemble (ORAS5). ORAS5 is driven by ocean observations from floats, buoys, satellites and ships.

Regarding the atmosphere, the grid resolution is 36 km, the wave model resolution is 0.5° and the vertical resolution is L91.

Finally, SEAS5 hindcasts period covers 1981–2016 and has an ensemble size of 25 members. However, since the common period of all forecast systems is from 1933 to 2016 the latter period will be used.

2.3 A threshold to select low wind speed events

The goal of this thesis is the evaluation of the low wind speed episodes (i.e. wind droughts) which can affect wind energy production leading to economic losses. To take into account the day-to-day fluctuations of the wind speeds, wind drought indices have been defined on daily basis.

As it is aimed to represent anomalous events, the design of the threshold should be proper for each longitude and latitude. For each day of the year, the corresponding threshold is defined as the 10th percentile of all the days of the same date in the other years, which is defined as the 10th climatological percentile. Moreover, the threshold has been smoothed by taking three days before and three days after (i.e. 7 days window) to remove specific day-to-day fluctuations [34]. Hence, each day threshold is the result of the 10th percentile on a 24 years (1993-2016) · 7 (days) dataset.

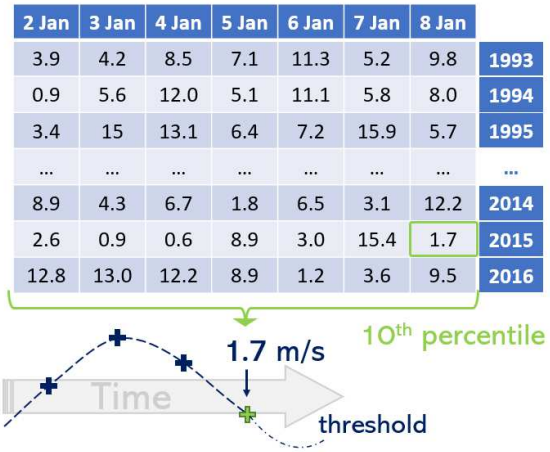


Figure 16: Schematisation of the daily threshold setting.

This thesis focuses on the evaluation of the wind drought days in the boreal winter defined by December, January and February (DJF). This season has been selected due to the high variability of wind speed compared with other seasons. This strong variability has a major impact on wind energy applications, therefore it is when the seasonal forecasts can be more valuable. The seasonal forecasts used will be those initialised the 1st of November. This choice is motivated by the fact that in real operations, the forecasts are available for the 15th of the month, and the first month is usually discarded for the forecast quality assessment, which is one of the main goals of this thesis.

2.4 Creation of two definitions to evaluate the wind drought days

The creation of an index to characterise the wind drought is stressed by the need for a common definition to evaluate the ability of the state-of-the-art seasonal forecast systems to predict the probability of wind droughts in a season. The wind drought indices have been defined to minimise the impact of the systematic errors inherent in the climate prediction systems. Models are inevitably subject to biases. However, by creating an index from internal variables of the systems, one obtains a comparable output. Hence, the next definitions characterise the local presence of anomalously low wind speed episodes.

The definitions of the different wind drought indices are greatly inspired by the work of Ohlendorf et al. (2020) [35]. They defined two indices for wind power generation and for data with a high temporal resolution. However, these indices have been adapted in this thesis to provide information on wind droughts at seasonal timescales. These indices are directly dependent on the threshold computation method, which is defined in the section 2.3. Therefore, this study defines:

- The Below Threshold (BT) days, are the days that the wind value is below a certain threshold. The threshold is the 10^{th} climatological daily percentile. This index is supposed to represent the occurrence of extremely low wind days.
- The Mean Below Mean Threshold (MBMT) definition is inspired by the Mean Below Threshold definition from Ohlendorf, but adapted to a mobile daily threshold. It concerns the days in which their conjoint mean wind speed is below the mean daily threshold of the corresponding period. The algorithm is privileging days occurring before a BT day since one expects that the physical conditions that may lead to wind droughts would take place before the event.

  This index is intended to be more flexible, smoothing the daily variability. Moreover, as the BT events could be seen as symptoms of a meteorological state, the presence of a favourable state does not impose a systematic occurrence of BT event and vice-versa. MBMT events, as wider, are supposed to be a better representation of this state. Nevertheless, the evaluation of the two indices might help to develop two different products that users can select depending on their specific needs.

In this work, the definitions have been used to output a number of days that are under the required conditions. There is no necessity of a minimum of consecutive days for an event to be considered as a wind drought. The metric is quantitative, and the evaluation of a more qualitative criterion on the length of the wind drought is unfortunately reserved for further analysis.

Figure 17 illustrates the main steps involved in the computation of the wind drought indices defined above. However, as the algorithm is maximising the period to the past, it could lead to some errors by missing intuitive MBMT periods. An example is given by the figure 18.

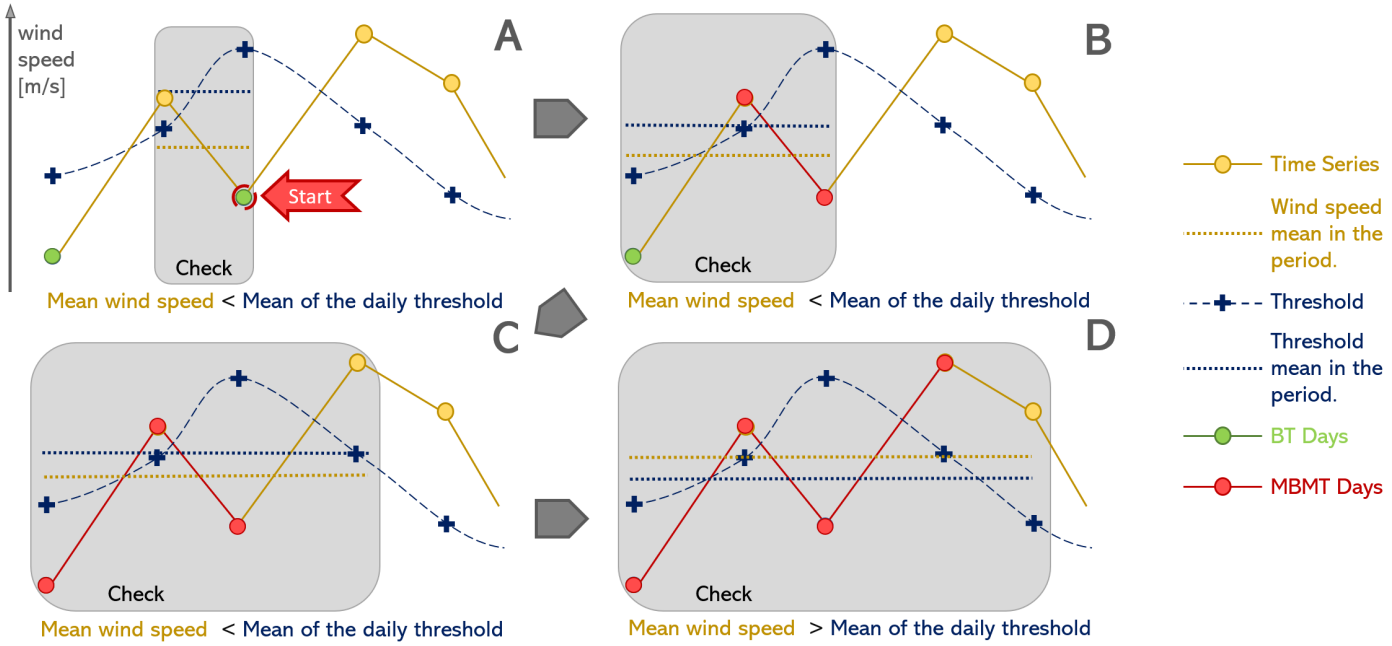


Figure 17: Illustration of the process leading to wind drought events selection. A. In the beginning, all days below the threshold are selected (below threshold (BT) day in green). Then for each BT period, the code tries to extend the period by checking the mean. The figure illustrates the MBMT selection process for the second BT point. First, the code compares the mean wind speed of the selected period to the corresponding threshold mean by including the previous (or left) point. B. If the condition is full-filled, the period will extend further to previous points until it can't. D-C. The process now tries to extend the MBMT period to the following days (the right) until the mean wind speed is not exceeding the mean threshold. The process will try again to extend to the previous period and then to the next days and so on until there are two consecutive failures.

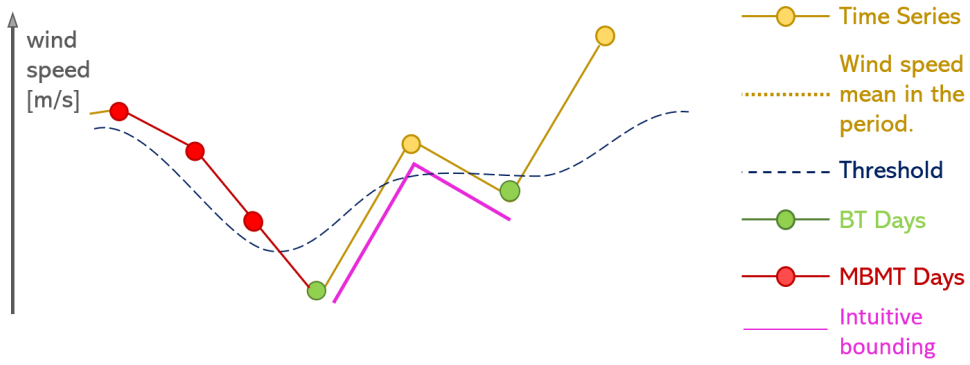


Figure 18: Illustration of the undesired behaviour, the algorithm is first expanding at maximum into the past, thus the mean value of the period is exceeding the mean threshold. The algorithm will not expand to the right. The second day below the threshold is too close to the threshold, then it also can't expand to the left. However, the desired behaviour would be to unify the different BT days in the MBMT definition.

Consequently, the correction consists in verifying all MBMT events resulting from the precedent algorithm. For a given event, the correction checks if it is not possible to extend the period to the right (the future) with different starting states which consist of a regression of the days before the

first BT event (see Figure 19).

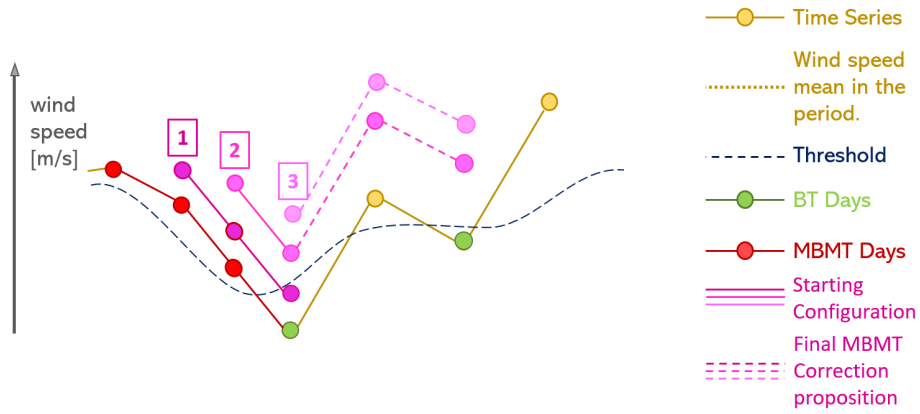


Figure 19: Illustration of the MBMT after process correction. First, it initiates by selecting all the days from the second day until the first BT day (included) of the MBMT period. Then days on the right are added to the selection until the MBMT definition is no longer true. Then the algorithm does the same but with a starting selection starting from the third day of the MBMT initial selection (in red). And so on until reaching the BT day. Finally, the MBMT period selected is the first period corrected that maximises the BT days number inside.

The results of the two indices are illustrated by Figure 20. One can see the BT wind drought days in green and the MBMT in red. The latter always need a BT day as a stem and extends the period, sometimes unifies independent BT periods as it can be seen on the end of the Time series. It implies that the number of MBMT days will always be higher than the number of BT days.

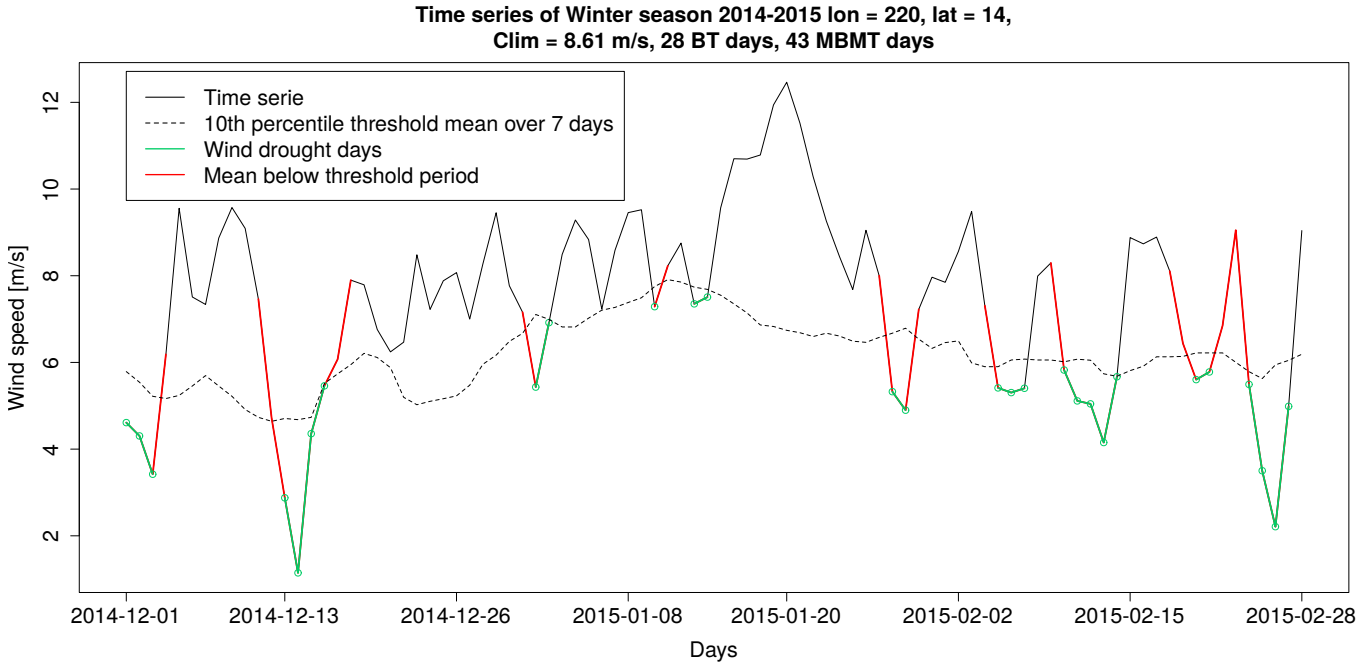


Figure 20: ERA5 Wind speed (m/s) from December 2014 – February 2015. Wind drought days as defined by the below threshold index are represented in green, and wind drought days based on the mean below the mean threshold are marked in red. The wind speed corresponds to a location in the middle of the Pacific [130° W, 12° N]. The algorithm find 28 BT days and 43 MBMT days.

2.5 Pearson's correlation coefficient and statistical significance

Pearson's correlation was used to compute one point correlation map of values of interest. For instance, it has been used to compare the monthly NAO index with the monthly wind speed and to compare the winterly output of the wind drought indices with the seasonal mean wind. Moreover, correlations are used to assess the similarities of the Forecast system with the reanalysis.

Pearson's correlation coefficient is an indicator of the linear correlation going from -1 (completely anti-correlated) to 1 (completely correlated). If the time series X and Y have a correlation equal to 1 (-1), it means that the variations of X are (inversely) proportional to the variations of Y. The correlation is given by:

$$cor_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y}$$

where $cov()$ is the result of the covariance and $\sigma_{X,Y}$ is the standard deviation of the X and Y time series. Assuming a normal distribution of the time series (even though this assumption might not be correct [36] for the wind speed at all timescales and for all the locations), the distribution of the correlation coefficient follows a t-distribution t(). Given by:

$$t(cor) = \frac{cor \cdot \sqrt{l-2}}{\sqrt{1-cor^2}}$$

where l is the length of the time series. The p-value giving the probability that the result is completely uncorrelated is then the probability that a 0 correlation would appear as the given value, multiplied by two because the sign of the correlation is not taken into account. Which means [37] :

$$p-value(cor) = 2 \cdot P(t >= T) = 2 \cdot \int_T^{\infty} \frac{cor \cdot \sqrt{l-2}}{\sqrt{1-cor^2}}$$

2.6 A tool to assess probabilistic forecasts: the Ranked Probability Skill Score

"The ranked probability score (RPS) measures the squared forecast probability error, and therefore indicates to what extent the forecasts lack success in discriminating among differing observed outcomes, and/or have systematic biases of location and level of confidence."

From the International Research Institute for Climate and Society of the British Columbian University. [38]

The Ranked Probability Skill Score (RPSS) is a widely used probabilistic metric for the forecast quality assessment. It gives a result from $-\infty$ (worst score) to 1 (perfect forecast) correcting the Ranked Probability Score (RPS).

As explained in the introduction, the output of a seasonal forecast over a variable consists in giving three probabilities of results. These probabilities can be given as the percentage of ensemble members in the below, near, or above normal categories. These categories are based on climatological terciles in which the number of occurrences in the reference period (1993-2016)

represents 33% the distribution. The RPS correspond to the mean square difference of cumulative probability between the forecast system output and the observational dataset. Of course, since the observation is not a stochastic variable, its probability is defined as 1 for the observed condition (for example, near average) and 0 for the other categories.

In this work, the RPS corresponds to:

$$RPS = \frac{1}{N_{Forecast}} \frac{1}{N_{cat} - 1} \Sigma_i^{M_{Member}} \Sigma_K^{N_{Cat}} \left(\frac{\Sigma_j^K M_{Occ,j}}{M_{Member}} - \Sigma_j^K P_{obs,j} \right)$$

Where $N_{Forevast}$ is the number of forecasts evaluated in a time series (for example in this work monthly forecasts cover 24 years which means that $N_{Forecast} = 24$), M_{Member} is the number of member in the forecast and N_{cat} is the number of category. In general, and as we do here, the outputs are divided in 3 classes (below/near/above)-normal, but other categories could be defined if the user is interested. $M_{Occ,j}$ is the number of members giving a result in the j^{th} category, $P_{obs,j}$ is the observational probability defined as above.
One should notice that as the definition uses cumulative probabilities, the last category will always have a contribution of 0 to the RPS.

The Ranked Probability Skill Score is the comparison of the RPS_F of a forecast system and the RPS_{ref} of a "reference forecast" [38] which in this case has been considered a climatological forecast giving always a 0.333 probability for each category. The equation is given by:

$$RPSS = 1 - \frac{RPS_F}{RPS_{ref}}$$

Positive RPSS is indicating that the forecasts provide more useful information than a climatological forecast. The RPSS=0 suggests that the forecast is not providing any added value with respect to a climatological forecast. RPSS below 0 suggests that the forecasts provide worse results than the climatology.

3 Results and Discussion

3.1 Behaviour of the wind drought indices

3.1.1 The computational weight of the MBMT

Dimension of data [Time · Orthogonal Dimension]	90 · 131 040	90 · 65 520	90 ·32 760	45 · 32 760	180 · 32 760
BT computation time [s]	7.2 ± 0.9	5.0 ± 0.5	3.1 ± 0.3	3.0 ± 0.2	3.3 ± 0.4
MBMT computation time [s]	1800 ± 200	1100 ± 200	580 ±40	300 ± 20	1300 ± 200
Ratio t_{MBMT}/t_{BT} [ ]	250 ± 70	220 ± 60	190 ± 30	100 ± 12	410 ± 90

Table 1: Table of the computation time respecting two the two indices. The threshold is set along the **Time** dimension. For a given input size, each result is obtained thanks to 3 different dataset run on three different computers.

The Table 1 shows the computation times for the two wind drought indices. These values have been obtained by running the function computing the indices for three equally sized dataset size and parallelised with 8 processors on a standard computer. The dimension of the dataset in input can be sorted into two categories. The dimension along which the threshold is computed, which is "Time", and the other on which the parallelisation is performed which can be named as the orthogonal dimension.

For a time dimension length of 90, which corresponds to the length of a daily seasonal time series, the ratio between the computation time of the MBMT and of the BT is about 220. The ratio seems to decrease with the global dataset size reduction, but the uncertainty is too wide to assert. The ratio seems also to be highly variable with the length of the time dimension. It's not possible to exclude a linear relationship, but it is more probable that this time increases in a non-linear way.
While for the MBMT, the computation time appears to be linear in any dimension, the BT index seems to be only linear in the orthogonal dimensions. When changing the threshold computed dimension size, it appears that the computational time is not changed. It signifies that the computation of the BT is negligible compared to the resources needed for the parallelisation, which is still a surprising result.

3.1.2 A relationship between the BT and the MBMT in ERA5

Figure 21 presents the distribution of the MBMT days according to the BT days computed thanks to ERA5 during the boreal winter from 1981 to 2016. One can see a saturation appearing for the seasonal number of MBMT days equal to 90, which corresponds to the length of the season (i.e. the maximum wind drought days).

A linear fit is performed on the whole dataset by excluding the saturated data. It gives a linear relationship of $N_{MBMT} = 1.57 \cdot N_{BT} - 0.98$. The linear regression result is mainly driven by the

highest occurrences that are of low seasonal numbers of wind drought days. There are at least 3 orders of difference in the occurrence between the low and the high number of seasonal wind drought days. The mean shows an edge of around 25 BT days, but the reasons behind this need to be further explored. The mean collapse appearing after 40 BT days is certainly an already existing consequence of the saturation. Indeed, no MBMT days can be above 90, so it means that the distribution is shifted down. The robustness of this analysis can be improved by using longer-time series for two main reasons, first, it will avoid the saturation on 90 days, and second, will show if the edge is symptomatic of a non-linear behaviour or if it is just a consequence of the asymptote $N_{BT} = N_{MBMT}$.

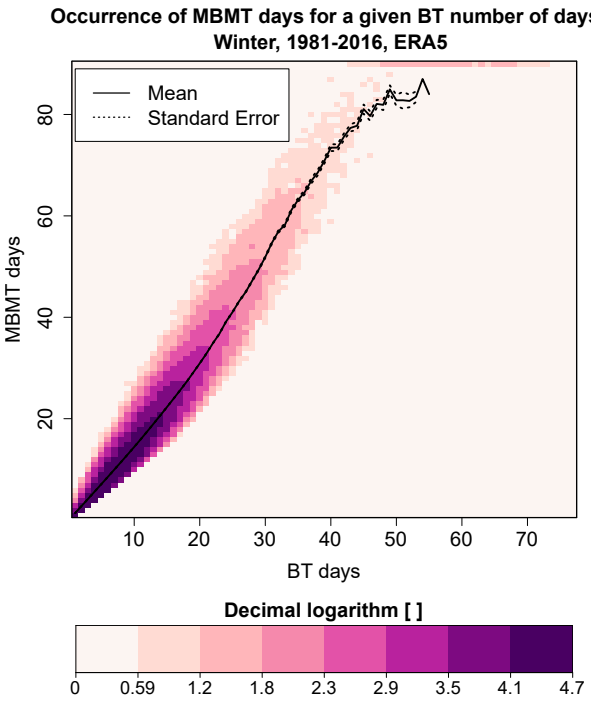


Figure 21: Occurrence of MBMT days for a given BT number of days, Winter (DJF), 1981-2016, ERA5. The distribution is displayed on a logarithmic scale. In the graphic, all the values that were 0 before the appliance of the logarithm are set again to 0. The linear regression gives $N_{MBMT} = 1.57 \cdot N_{BT} - 0.98$ with an uncertainty on the coefficient of respectively 10^{-4}, 10^{-3} and p-values of 10^{-16}

Figure 22 shows the seasonal climatologies of each index. Since the length of the DJF seasonal time series is 90 days, the climatological value corresponds to 9 days (by construction). The noisy results are due to the running window applied to smooth the 10% percentile (see Section 2 for the details).

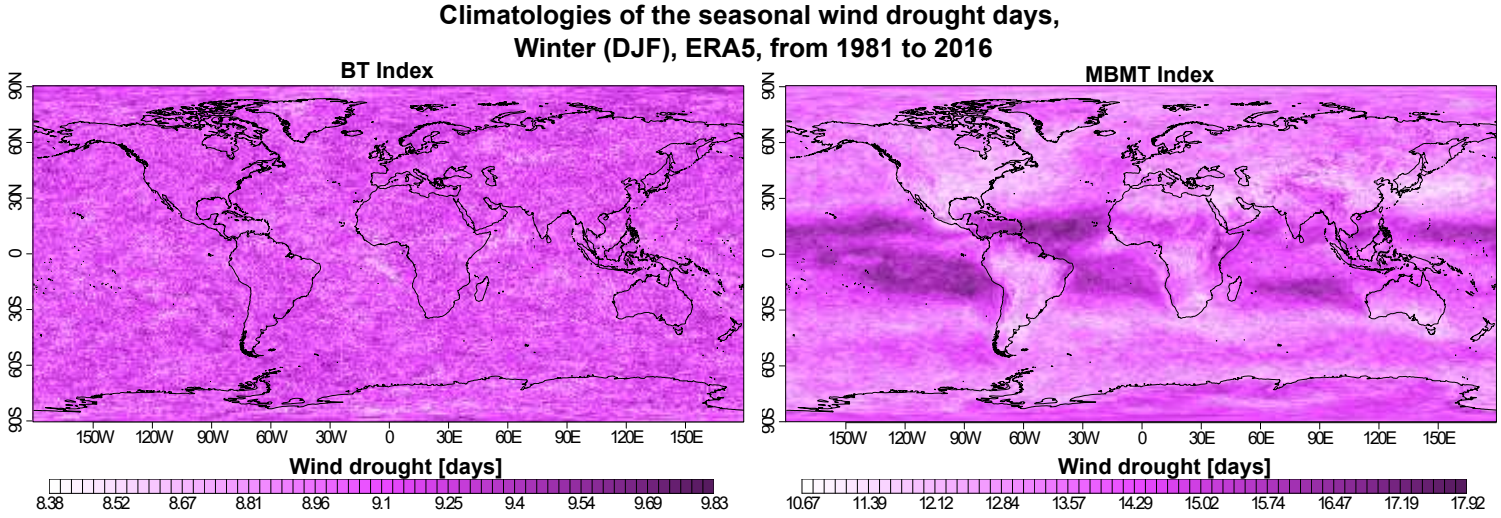


Figure 22: Climatologies of the seasonal wind drought days of the two drought indices, computed over Winter (DJF), from ERA5, from 1981 to 2016. Surprisingly the MBMT index shows patterns.

The MBMT index climatology is giving a more specific pattern. Actually, they are similar to what is obtained in Figure 23 showing the mean normalised one-day difference of the ERA5 wind from 1981 to 2016. This similarity is intuitive since the MBMT is extending from stems, which means that if from one day to another the difference in wind speed is high (i.e., if the running difference is really high), it is less likely that the neighbour day will be a Mean Below Mean Threshold day.

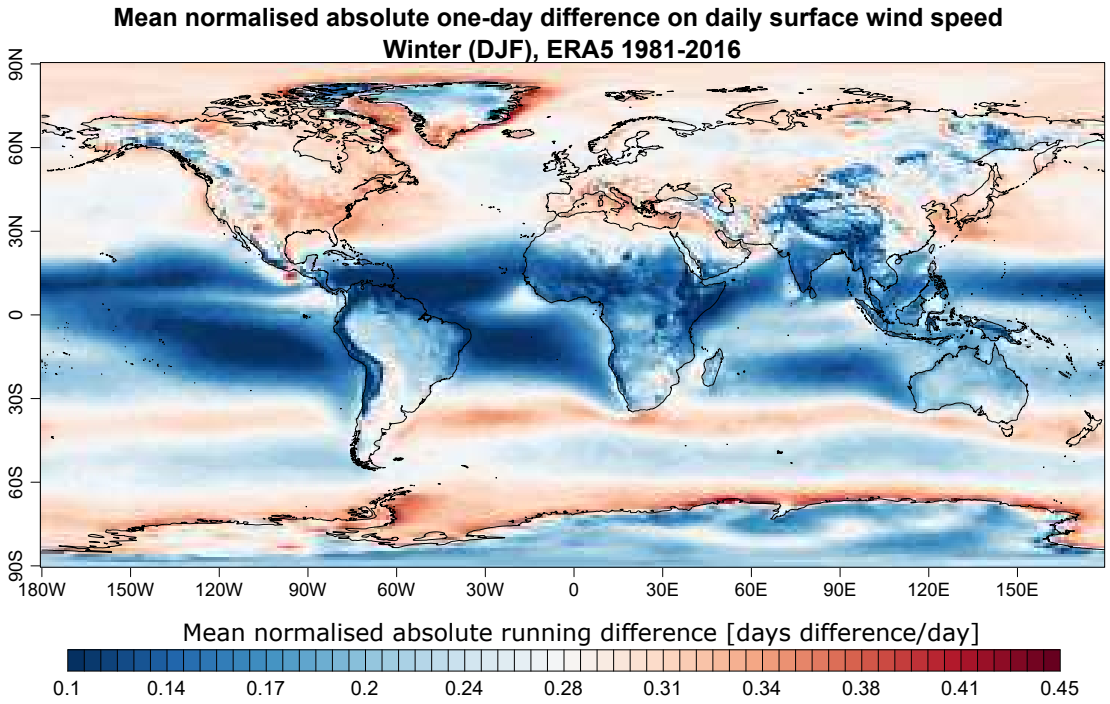


Figure 23: Mean absolute running difference on the surface wind of ERA5 during Winter season from 1981 to 2016

Finally, Figure 22 underlies a spatial variability in the BT to MBMT relationship. The MBMT climatology ranges from 11 to 17, meaning a linear relationship coefficient varying from 1.2 to 1.9 with the BT climatology.

3.1.3 Decadal evolution of the wind drought days

Figure 24 shows the linear trends in the winterly number of wind drought days according to the MBMT index from 1993 to 2016. This trend is the steering coefficient of a linear fit obtained from the winter number of wind droughts days time series over the years. Only trends with p-values lower than 0.1 are shown. The BT trends, visible in the appendix (Figure 38) exhibit the same behaviour as figure 24 but with smaller coefficients. Indeed, by definition the MBMT occurrence can only be higher than those for the BT. This is because a BT day is needed as a stem for MBMT periods which implies that the occurrence of BT days will result in more MBMT days. But, the absence of BT occurrence also results in the absence of MBMT days. Hence,

the MBMT linear trends slope can only be bigger. Nevertheless, the different indices have a similar significance. 12620 points have valid p-values in the BT definition and 12721 for the MBMT.

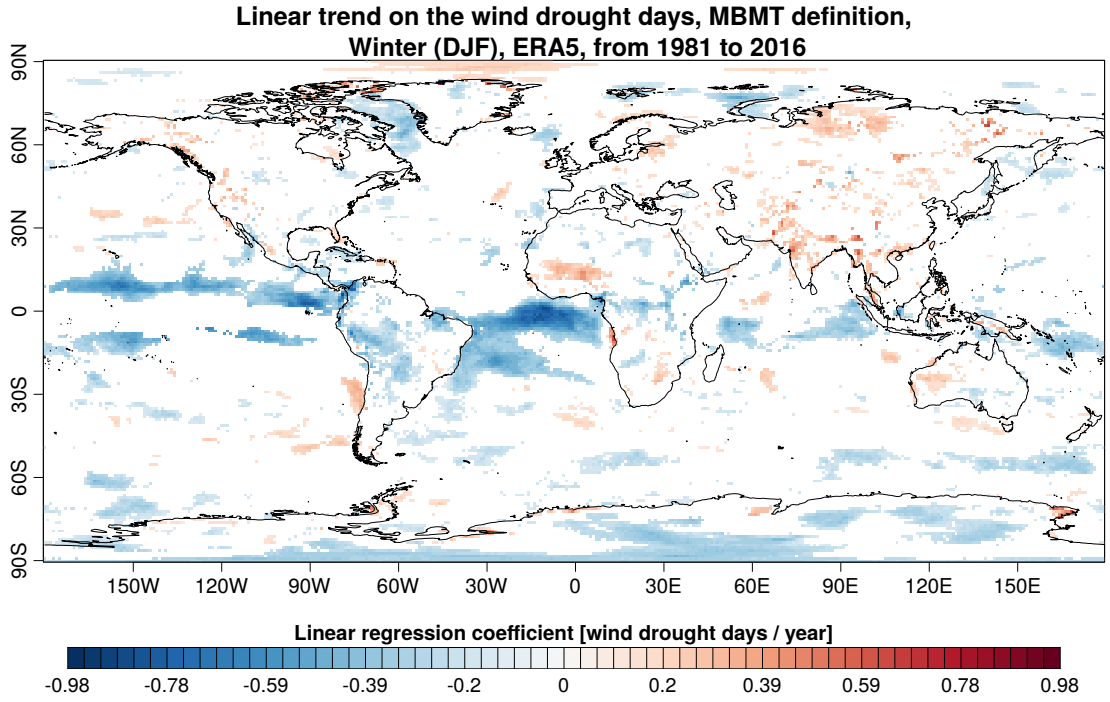


Figure 24: Linear trend on the MBMT days in the boreal winter from 1981 to 2016 provided by ERA5. Only the trends with their corresponding p-values under 0.1 are shown. The figure shows a clear negative trend along the equator and some on the South Pole. The map is globally anti-correlated as one could have expected, when the wind is higher there are less drought days.

As the prominence of white underlies, the trends are mostly insignificant. However, one can see negative trends on the South Pole and along the equator. The MBMT wind drought trend reaches -0.86 with a p-value of order 10^{-5} (-0.49 BT days/year with a p-value of order 10^{-5}). This might suggest that the MBMT is more variable than the BT index. The values of the linear trends represent a difference of 31 days (18 days) of wind drought days in winter between the 1981 and the 2016. Extreme tendencies occur around the equator and especially the ITCZ. Indeed, along this line, the air is ascending, which means that there is in general no strong wind at these latitudes (Figure 4). A change in this location is then relatively more important and reinforce the tendencies. In fact, there is a global increase in wind speed over the oceans noted by different authors (Young et al. 2011 [39], Zheng et al. 2016 [40]). The cause is attributed to climate change (L'Heureux et al. 2013 [41], England et al. 2014 [42]), that is the strengthening of the Walker circulation (which is in a few words, the model of closed air circulation in the tropics, strongly linked with the ocean circulation) [43]. CITE

Figure 25 shows the linear trend of the winter seasonal mean wind speed. The wind speed trends are opposite to the BT trends. The threshold the 10^{th} over the years (see section 2.3). If the

mean wind speed decreases along the years, one expects more values below this threshold. Even though there is a relatively important increase in wind speed on the Antarctic Sea, there are not significant trends in the wind drought days.

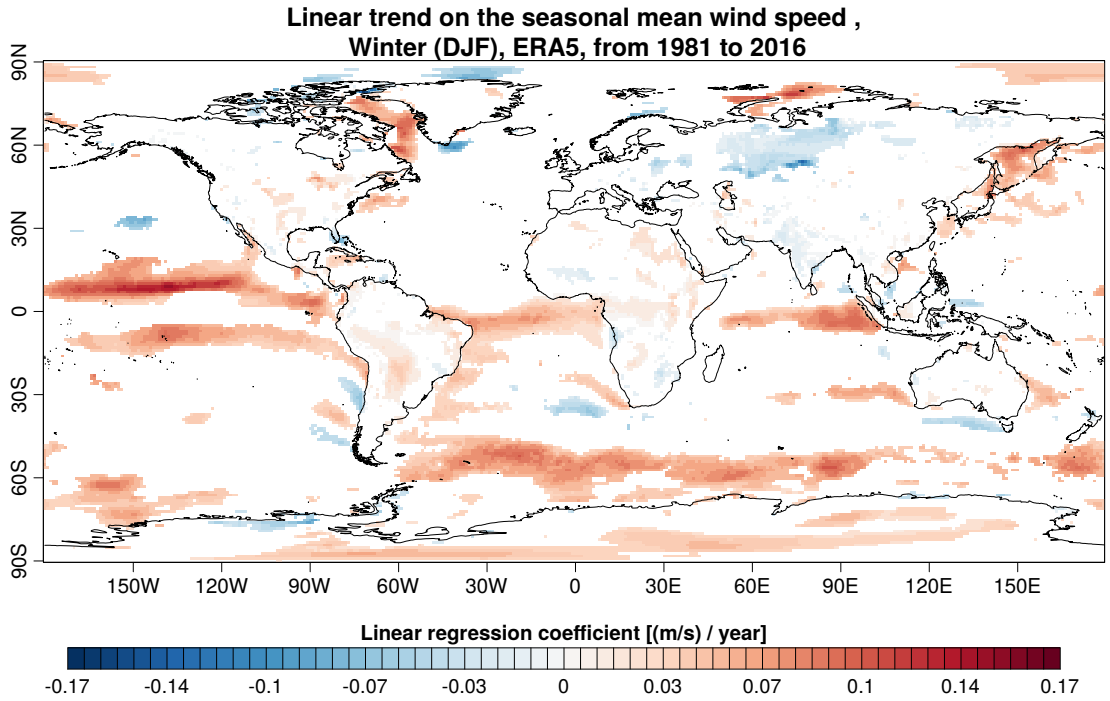


Figure 25: Linear trend of the winterly seasonal mean wind speed trend from 1981 to 2016 provided by ERA5 . Only the trends with p-values under 0.1 are shown. The figure shows similar trends as the Figure 24.

Figure 26 shows the evolution of the winter number of wind drought and the mean seasonal speed from 1981 to 2016, in two different locations. The top figure (lat 0°S, lon 9°W) corresponds to the location where the number of drought trends reaches the local maximum intensity in the South Atlantic (close to the African coast). Secondly, the bottom figure illustrates the wind speed and the wind drought indices for a grid point in Denmark (lat 55°N, lon 9°E).

Once again, the two indices have a similar behaviour. There are no years where the BT index is giving a minimum that is not followed by the MBMT. However, the "response" of the MBMT index is not always proportional to the BT.

Figure 26 (top) shows the negative trend of the number of wind drought days in this area responding to the increasing trend of the mean wind speed. Even a seasonal variation of the wind drought indices of about 10% implies a reduction of 4 wind drought days.
Generally, the local minimum in the seasonal mean speed results in the occurrence of wind drought. However, the time series of Denmark (lat 55°N, lon 9° E) shows disagreements with the mean wind speed. For example in 2012, a high number of wind droughts were identified even though that the seasonal average speed for this year is not a local maximum.

Figure 27 shows a map of the correlation between the winter number of wind droughts and the seasonal mean speed. Even if the map shows a prominent blue and has a mean value of -0.56 with the MBMT index (-0.55 MBMT), there is an important spatial variability. One can see from Figure 14 that the wind stations are most of the time situated in un- or low correlated areas. There is no clear pattern about where these (anti-)correlations are higher or not. It seems that it is in general high along the equator.The normalised standard deviation shown in the Figure 28 seems to have a similar distribution than the Figure 27.

Indeed, the locations that have a low normalised standard deviation also have high anti-correlations. It is expected that a lower variability will give the seasonal wind mean speed a better indication of the shift of the wind speed distribution. Therefore, it means that the white area in the Figure 27 corresponds to points where the seasonal wind mean speed indicator does not catch the intra-seasonal variability of the wind. Finally, the mean correlation from the BT index in both hemispheres is -0.56 (-0.55 MBMT) which shows that there is no evidence of a difference between summer and winter in the link between the indices and the mean wind speed values .

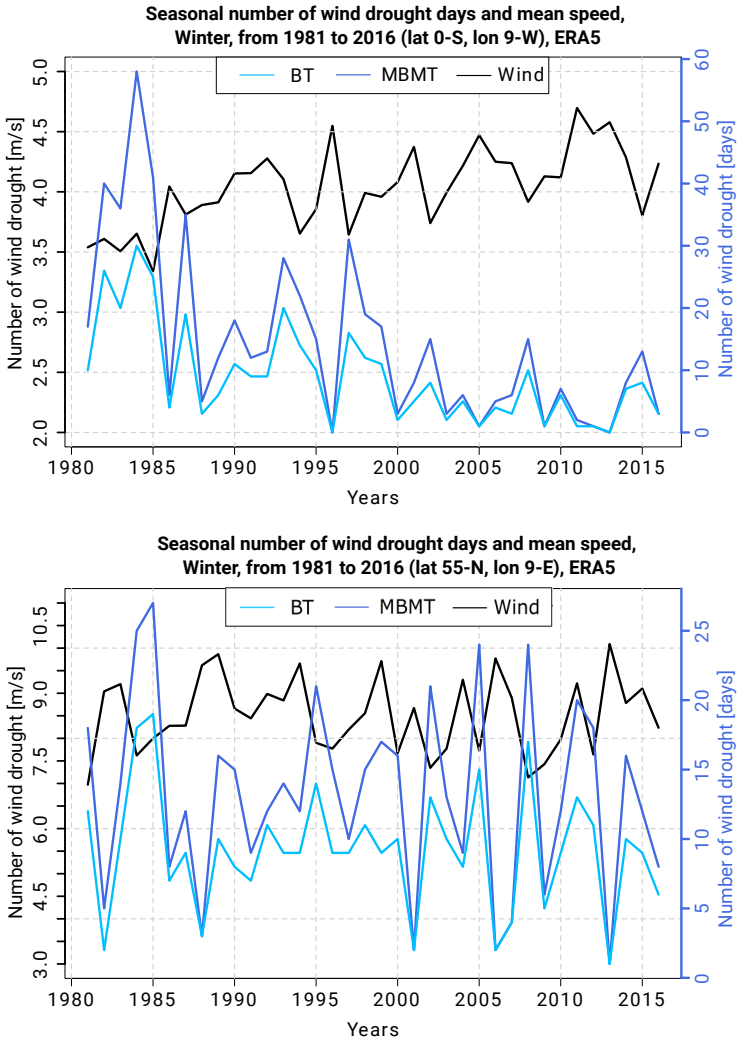


Figure 26: Time series of the number of wind drought days during winter and of the mean seasonal speed along the years. They correspond to two different latitudes and longitudes (**Upper figure:** South Atlantic, **Lower figure:** Denmark) The two indices are represented. The **black** is the mean seasonal wind per year and the **blue** lines are the number of low wind events.

From Figure 28 a route along the equator of local maximums of the coefficient of variation. It corresponds to the position of the ITCZ. As it has already been said in the beginning of this section, the ITCZ shifts spatially from one month to the next and thus leads to high differences in the wind in a short times.

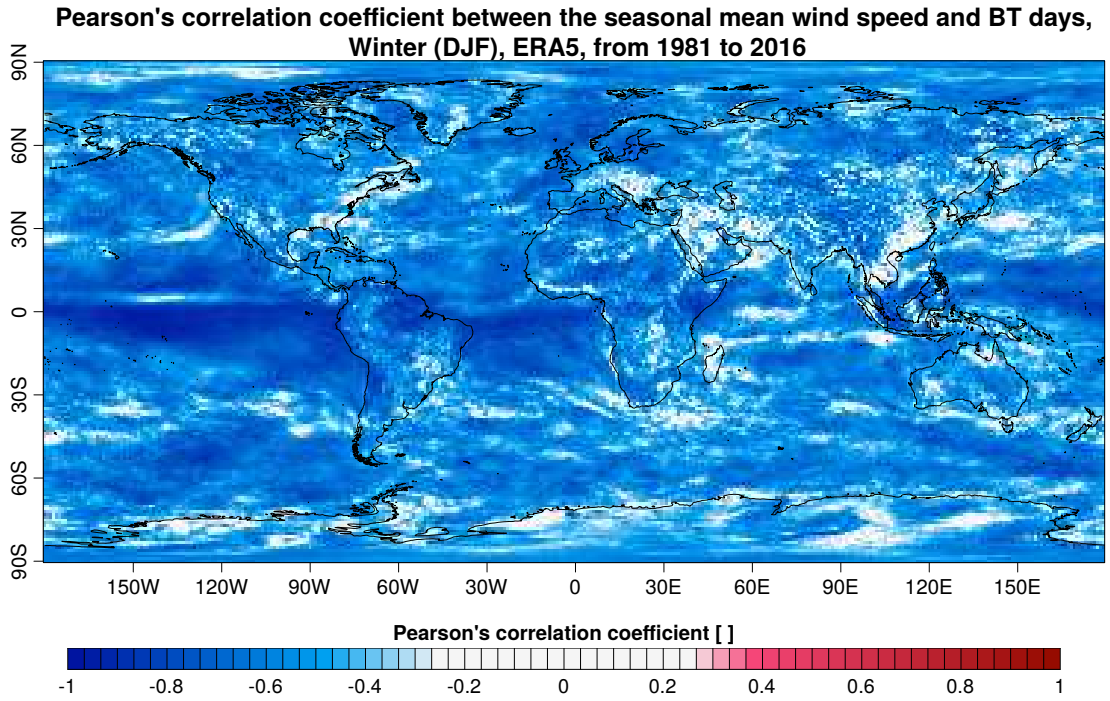


Figure 27: Correlation maps of the seasonal mean wind speed with the winter number of wind drought according to the MBMT indices. Data from ERA5. The corresponding figure for the BT is in the appendix (Figure 36).

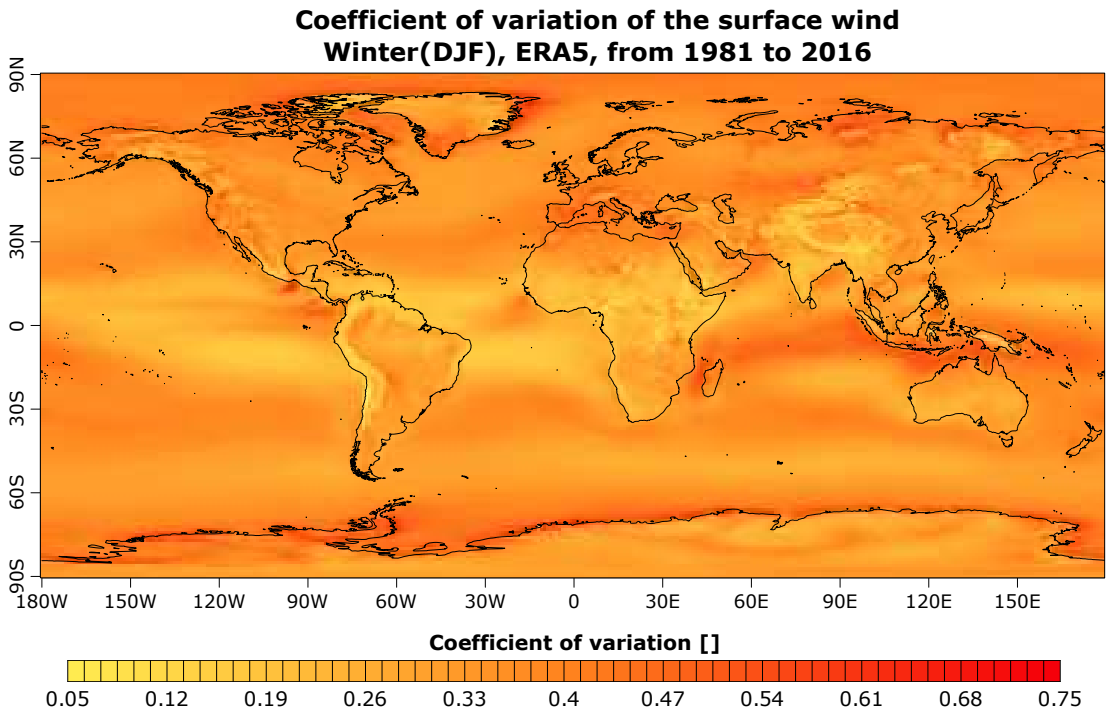


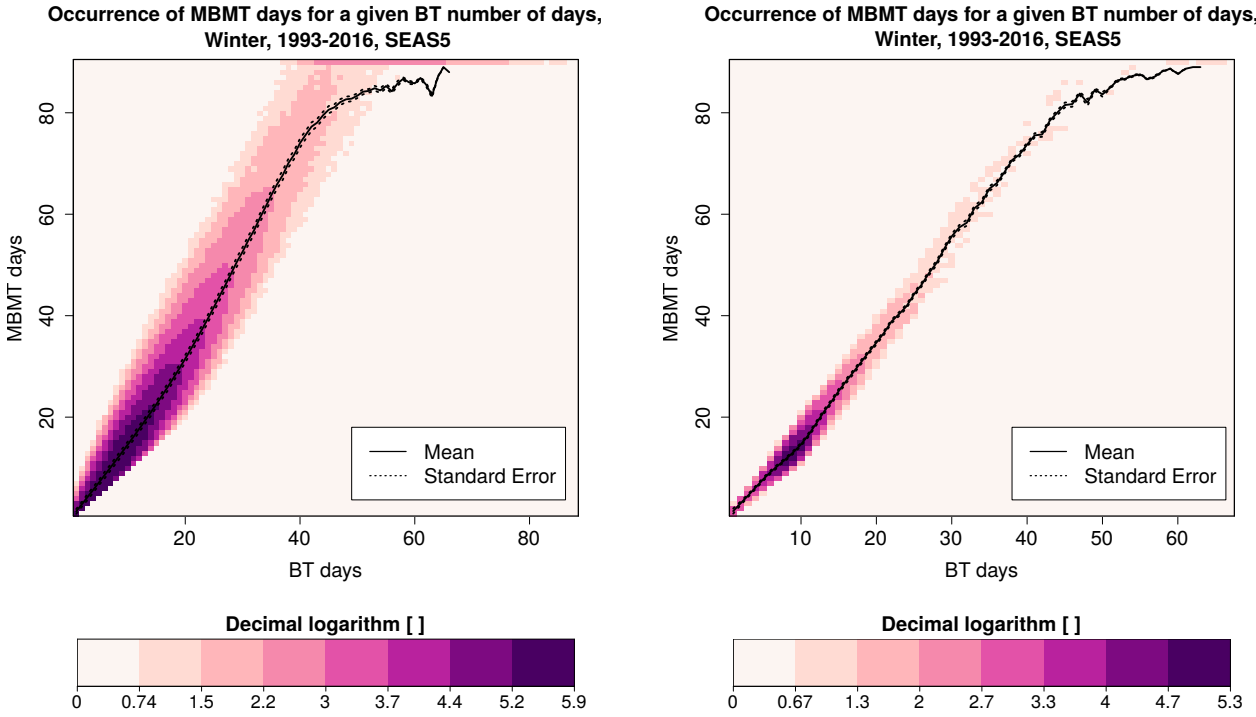
Figure 28: Coefficient of variation obtained from ERA5's winter daily surface wind from 1981-2016. The coefficient of variation is given by the standard deviation normalised by the average. The Tropics are generally characterised by a low coefficient of variation.

3.1.4 Summary

Technically, the computation time of the MBMT index takes about 220 times longer than for the BT for a seasonal time series length (90 points). This means that the MBMT computation requires the use of more computational resources. Moreover, the strong seasonal relationship encountered between the two indices suggests that it could be judicious to save this time for other applications. However, climatologies suggest that this relationship is spatially dependent and that it varies according to the normalised running difference of the daily wind speed.
The significant trends of the wind drought days over the years are following the trends of the seasonal mean wind speed. Nevertheless, the seasonal mean wind speed has a low global correlation (-0.55) with the wind droughts which implies that this indicator is unable to catch the inter-annual variability of the wind droughts. This idea is supported by the similar pattern observed with high correlations and the coefficient of variation of the daily wind. The seasonal mean wind speed is therefore not reliable as a predictor of the seasonal number of wind drought days.

3.2 Seasonal forecast quality assessment of the wind drought days predictions

3.2.1 Relationship between the BT and MBMT index in SEAS5



SEAS5 Members. The linear regression given by this ensemble is $N_{MBMT} = 1.58 \cdot N_{BT} - 0.93$ with a p-value of 10^{-16} and respective uncertainties of orders 10^{-4} and 10^{-3}.

SEAS5 ensemble mean. The linear regression given by this ensemble is $N_{MBMT} = 1.69 \cdot N_{BT} - 1.92$ with a p-value of 10^{-16} and respective uncertainties of orders 10^{-4} and 10^{-3}.

Figure 29: Occurrence of seasonal MBMT days for given BT days from 1993-2016 with the SEAS5 ensemble mean. In the graphic, all the values that were 0 before the appliance of the logarithm are set again to 0 and occurrences < 1 are not shown. N_X is the number of wind drought days of the variable X. Decimal numbers of wind drought days obtained by the mean are rounded to the next integer.

The relation between the SEAS5 seasonal predictions of the BT and the MBMT indices is shown in Figure 29. It has been evaluated for the boreal winter season in the 1993-2016 period. On the left, is the distribution for all members predictions, and on the right, for the ensemble mean. By a quick comparison with Figure 21, one can observe that the distribution of all the members together is very close to the one obtained by ERA5 as evidenced by the closeness of the linear fits. From this result one would expect the predictions of SEAS5 to be close to the ones of ERA5.

However, the ensemble mean is showing some differences. Even by taking into account the reduction of the sample size ($36 \cdot 90$ to $24 \cdot 90$ which means a $\log_{10}(3) \approx 0.5$ difference order), the SEAS5 seasonal predictions exhibit a narrower distribution than the one of ERA5.ERA5 have $\approx 10^{2.3}$ seasonal occurrences of 25 BT days and SEAS5 $\approx 10^{1.3}$. Therefore, the mean applied on the members that centres the distribution on 9 BT days and 12 MBMT days (Figure 29) and the high occurrences are less probable. Finally, the linear regression gives a coefficient of ≈ 1.7 between the two indices, with a lower intercept than the one obtained for ERA5.

3.2.2 Detailed SEAS5 predictions for January 1998

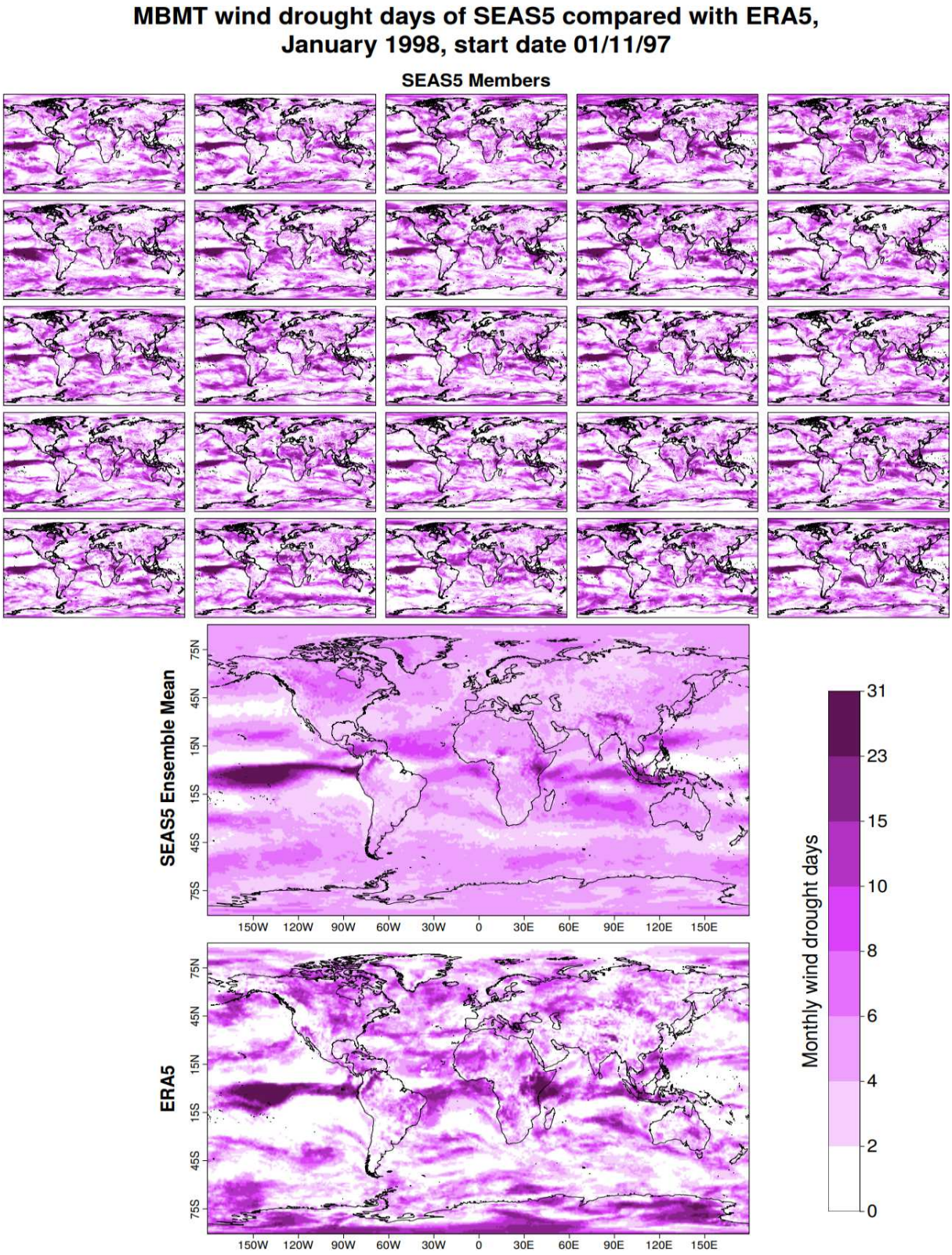


Figure 30: MBMT number of wind drought days predicted by the seasonal forecast system SEAS5 and the corresponding wind droughts obtained from ERA5 for the month of January 1998. First are represented all the different members (simulations) given by SEAS5. In general, the same behaviour is observed but the small changes in the initialisation change the results. The BT index gives a similar figure and is in the appendix (Figure 40).

Figure 30 shows the number of wind drought days predicted by SEAS5 and given by ERA5 for January in 1998. First, the 25 ensemble members are displayed. In general, similar behaviour is observed for all the ensemble members, but the small changes in the initialisation lead to non-negligible variations. The ensemble mean underlies this variation, firstly by the reduced contrast and then by the smoothing of the patterns. As an example, one can look at the Pacific and see that it is characterised by an intense line of wind drought number along the ICTZ. In the ensemble mean, the pattern of this high wind drought zone is well thinner. Another example is the land surface where the wind drought days range between 0 and 1, which is a lower number of days compared to those obtained for the reanalysis. This is the result of the smoothing corresponding to the ensemble mean computation. Nevertheless, the members have similar behaviour to the ERA5 observational reference.

3.2.3 Mean bias of the wind drought days seasonal forecasts

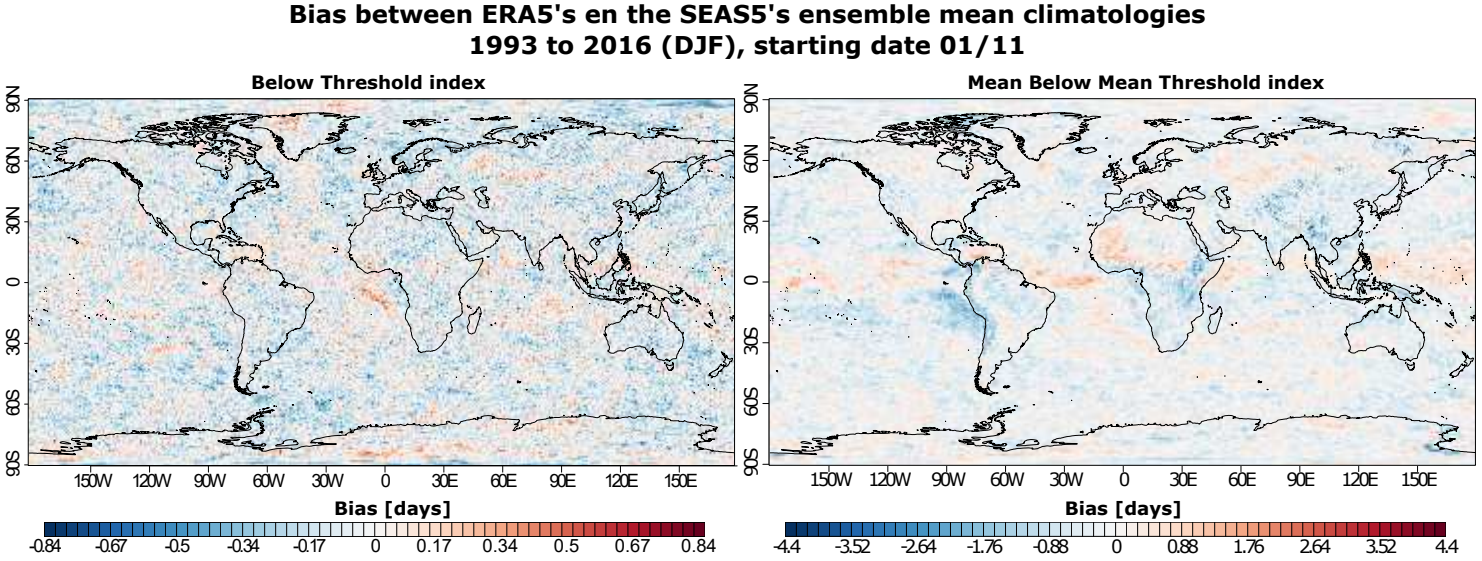


Figure 31: Bias between ERA5 and SEAS5 ensemble mean for Winter. The data is from 1993 to 2016 and the forecast system is launched on the 1^{st} November. The bias is the climatology difference between the reanalysis and the ensemble mean of the seasonal forecasts. Blue colours indicate that SEAS5 is predicting more wind droughts than those from the ERA5 reanalysis and orange indicate the opposite.

	Global mean	Global absolute mean	Equatorial mean	Equatorial absolute mean	Extra-tropics mean	Extra-tropics absolute mean
BT [days] (Uncertainty $\approx 10^{-3}$)	-0.07	0.14	-0.05	0.14	-0.07	0.14
MBMT [days] (Uncertainty $\approx 10^{-3}$)	-0.19	0.42	-0.22	0.53	-0.19	0.39

Table 2: Global, equatorial and extra-tropics mean and absolute mean bias of the Figure 31. The absolute mean is the mean of the absolute values.

Figure 31 shows the bias, which is the difference between the seasonal climatologies of ERA5 and of SEAS5 ensemble mean for both wind drought indices. The BT days difference between the climatologies of the seasonal forecasts and the reanalysis is really low and never exceeds a day, which means that the climatological behaviours of the BT index is consistent for the seasonal predictions and the reanalysis. This is due to the definition of the BT index as in terms of a relative threshold (i.e. the 10^{th} percentile) which is defined separately in the seasonal predictions and in the reanalysis. Even though there is a slight overestimation of the wind drought days by the forecast system since the mean bias is negative, this definition leads to an unbiased BT index.

The results of the MBMT index are different to those obtained for the BT. First, the mean bias reaches 4.4 days in specific locations (Figure 31) with a mean bias of -0.19 days with an absolute mean of 0.42 days (Table 2). Both MBMT means have a ratio with BT means well away from the 1.57 coefficient of the linear relationship. It may indicate a difference in the seasonal behaviour of the two indices that has not been emphasised.

The equatorial and tropical regions show stronger differences between the two climatologies which is consistent with the results obtained in previous sections. Except for Asia, the biases over the extra-tropics is noisy in a similar way to the BT index. Indeed, the absolute mean between the 25° North and South is equal to 0.53 days while the extra-tropics have an absolute mean of 0.39 days. As a comparison, the BT index shows respectively absolute means of 1.40 and 1.34 days. This result was expected since the BT index is directly the representation of the 10^{th} percentile. Hence, the climatology of this definition should give 10% of the period (with variations since the threshold is set over a week and the algorithm is seeking the one that is approximating the best, which depends on the distribution). By contrast, the MBMT index is well more dependent on the smoothness of the wind speed daily values and their distribution.

3.2.4 Deterministic evaluation: Pearson's correlation coefficient

Figure 32 shows the correlations maps between the wind drought days from ERA5 and the SEAS5 ensemble mean from 1993 to 2016. These correlations allow assessing if the seasonal predictions have the potential to detect year-to-year fluctuations in the occurrences of wind droughts. The Pearson's correlation coefficient ranges between -1 and 1. If $cor = 1$ ERA5 and SEAS5 fluctuations are perfectly corresponding. When $cor = 0$, it indicates that there is no association.The ensemble mean prediction does not provide added value relative to the retrospective observational climatology (which would be a constant time series). Finally, if $cor < 0$ the ensemble mean predictions are worse than the past climatology, Positive correlation is the minimum for predictions to have potentially useful additional information. (Jolliffe and Stephenson, 2012, [44]).

Since negative correlations signify unskilful predictions, the mean is computed instead of the absolute mean, which shows that the global correlations are low.

As the mean correlations are showing, the general quality of the forecast is decreasing with the forecast time. This is due to the reduction of the influence from the initial conditions (observations at the time the forecasts are issued) with the forecast time. However, the decrease of the correlation between the two last months is really marginal. Still looking at the means, the forecast over the season is giving a better result since the information requested from the system is less precise.

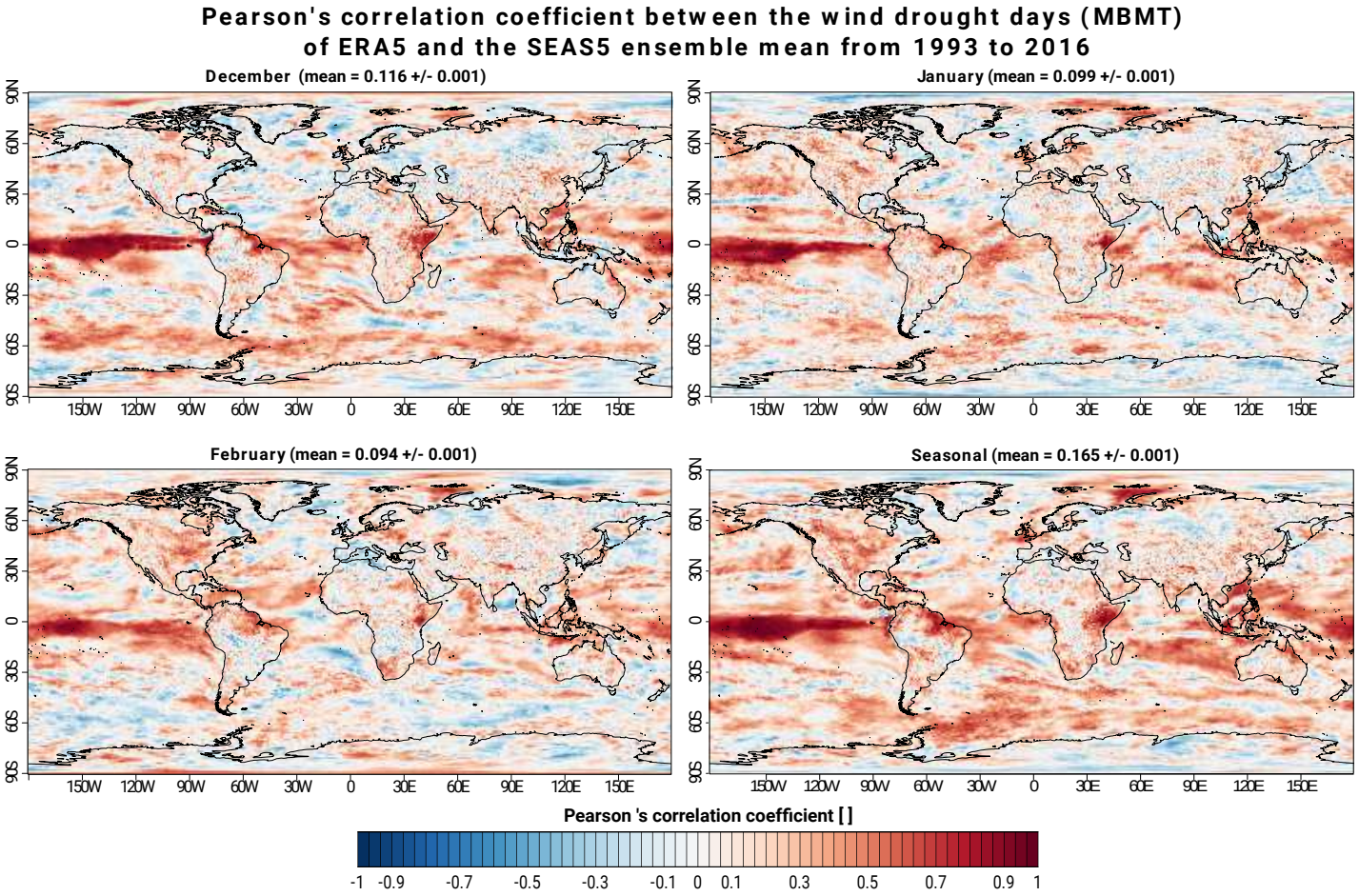


Figure 32: Pearson's correlation coefficients between the MBMT wind drought days of ERA5 and the SEAS5 ensemble mean from 1993 to 2016.The predictions are initialised the 1st of November. Each respective map mean is shown. The seasonal forecast is the one showing the highest correlation. The BT index shows a similar behaviour and is shown in Figure 39. The monthly maps show the correlation over time of the monthly number of wind droughts whereas the seasonal map is the correlation over time of the winter seasonal number of wind droughts (sum over three months).

By looking more into the spatial distribution of the maps shown by Figure 32, it is clear that better skill on the yearly fluctuations is demonstrated in the equatorial Pacific, North Brazil, and North Eastern Africa (all along the equator). All three locations seem to be well influenced by the El Niño Southern Oscillation phenomena in winter, and this phenomenon is currently well predicted by the different forecast systems [9]. Furthermore, these regions are characterised by a small coefficient of variation (Figure 28) which might suggest a more predictable wind behaviour.

3.2.5 Probabilistic evaluation: the Ranked Probability Skill Score

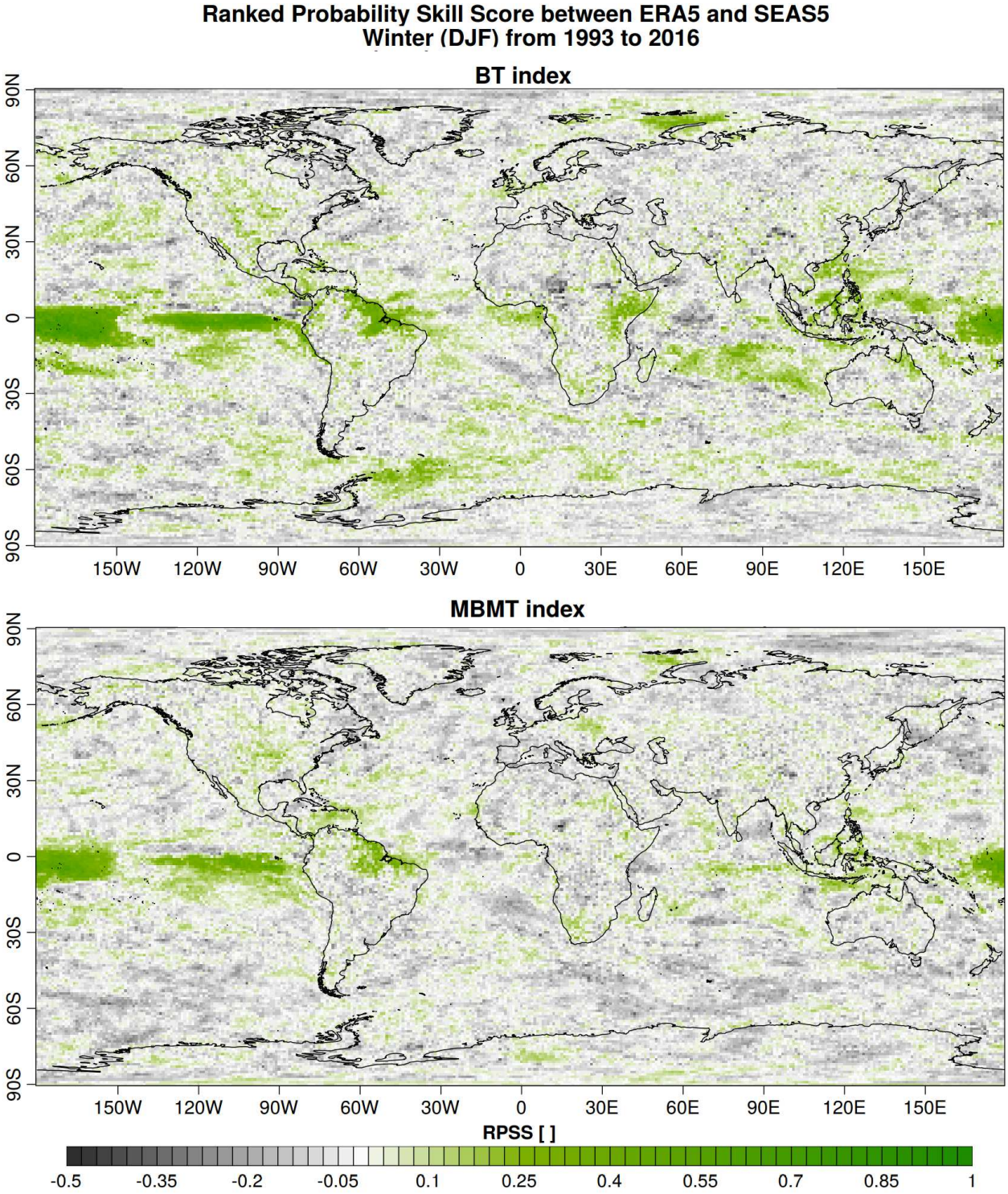


Figure 33: Ranked Probability Skill Score between ERA5 and SEAS5 in winter from 1993 to 2016 with the MBMT index. The best skill is obtained along the equatorial pacific. The predictions are initialised the 1st of November. BT RPSS average $= -(13.1 \pm 0.5) \cdot 10^{-3}$, MBMT RPSS average $= -(3.21 \pm 0.4) \cdot 10^{-3}$. A score of 1 is considered as a perfect forecast and a score below 0 corresponds to unskilful forecasts (represented in grey).

Figure 33 presents the result of the RPSS score of the SEAS5 winter predictions (starting date 1*st* of November) compared to the ERA5 reanalysis. The two indices show better performance along the equatorial and tropical latitudes, and especially in the pacific, compared to the extra-tropics. The BT index seems to have in general a better skill than the MBMT. The difference is particularly marked in the equatorial and tropical Pacific. North America and Europe, which are relevant regions for the wind industry show positive RPSS values. This implies the seasonal predictions of the wind droughts could benefit the wind industry in that areas. However since the RPSS values are quite low, more careful analysis should be done on these areas.
However, BT better performance is not systematic. For example, the SEAS5 seasonal prediction system is more skilful for the MBMT than the BT over Belarus and Ukraine.

While it might be expected that in general, a better skill would be obtained for the BT index that is based on the 10^{th} percentile, and consequently, less dependent on the wind distribution Than the MBMT, the MBMT shows better skill on some regions as Belarus and Ukraine. The presence of numerous points suggest a significant result. It may indicate that the wind behaviour is different on these points and that difference in skill needs to be further explored.

3.2.6 Summary

The two indices show a global unbiased behaviour between the forecast system and the observational dataset, which is a requirement of a good index. However, some small biases (less than two days) are emerging with the MBMT wind drought days and would need a more specific analysis.

The mean performed on the forecasts ensemble is responsible for the difference in the MBMT and BT days distribution and thus, the difference in the wind behaviour between the observation and the ensemble mean. The example output of Figure 30 underlies indeed that the ensemble mean of the forecast system is fading the contrast of the output at a monthly scale, only showing a similar behaviour on the equatorial latitudes and particularly over the pacific. Indeed, the Ensemble mean correlation between the forecast system and the reanalysis and RPSS show that the equatorial region is well predicted. This region is dominated by the ENSO phenomenon which is well understood and skilfully predicted at seasonal time scales [9].
Nevertheless, the results of those metrics suggest that the seasonal prediction systems need to be improved to provide more skilful predictions of wind droughts on a global scale. However, there are some exceptions such as north Brazil, and some locations in North America and Europe that can benefit from the climate information provided by the seasonal prediction systems.

3.3 The North Atlantic oscillation impact

3.3.1 Relationship between the wind drought days and the NAO

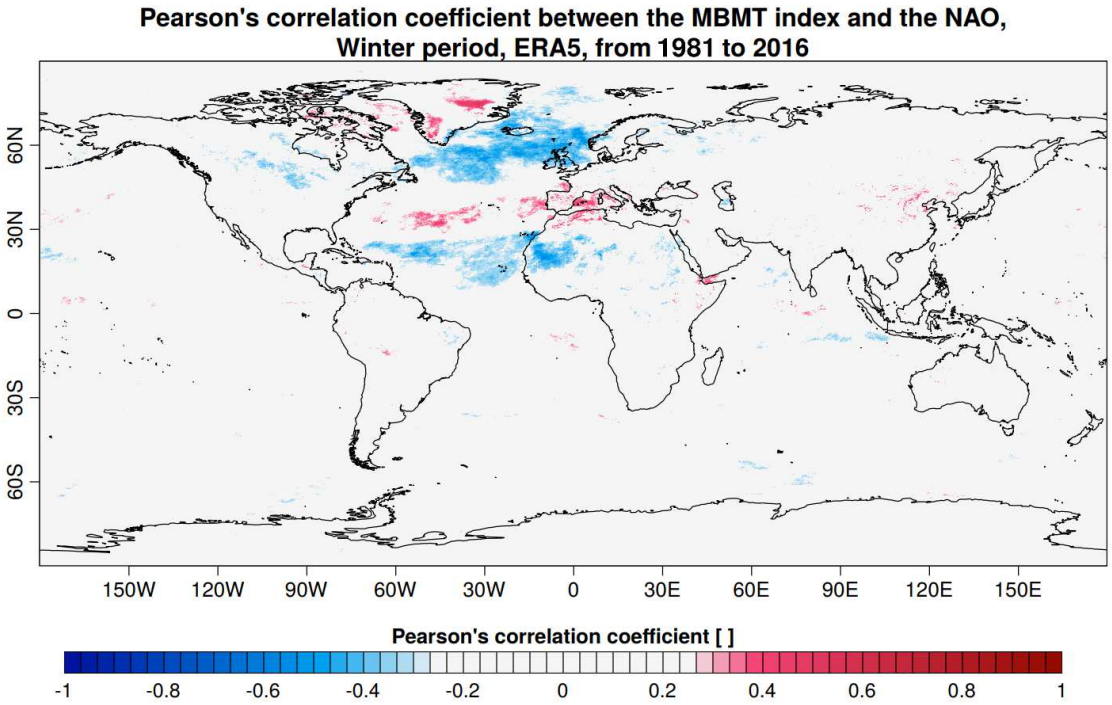


Figure 34: Pearson's correlation coefficient between the monthly wind drought obtained with the MBMT index and the NAO monthly value. The BT figure is similar and available in the appendix (Figure 37).

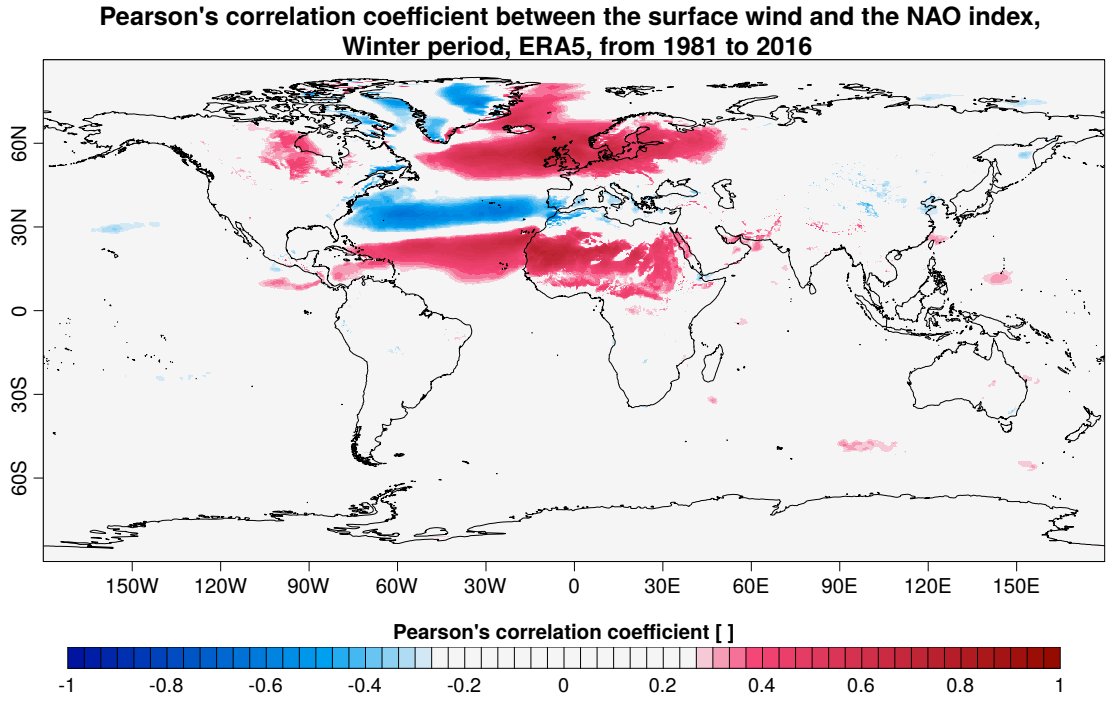


Figure 35: Pearson's correlation coefficient between the monthly value of the 10m surface wind speed (ERA5) and the NAO monthly value.

From January's monthly correlations between the wind and the NAO index in winter (Figure 10 and later, Figure 35), one has seen that the NAO is a good descriptor of the monthly Euro-Atlantic wind behaviour. One could then expect a similar behaviour between the NAO and the wind drought indices.

Figure 34 shows the Pearson's correlation coefficient between the monthly wind drought obtained with the MBMT index and the monthly NAO index. When the correlations are negative, it means that if the monthly value of the NAO index is positive the number of wind drought days during this month would be lower than usual. When the correlation is positive, the opposite happens.

The patterns observed correspond to the ones from Figure 35 showing the correlations between the surface wind speed and the NAO index (similar to Figure 10 but over the whole winter). Indeed, the sign of the significant correlations are opposed: where the NAO induce a higher monthly wind speed, the number of wind drought days during the year decrease. But, these correlations are weak, the absolute value is around 0.4. It can be understood as the consequence of the low correlation of seasonal wind speed with the wind droughts shown in Figure 34. Moreover, the correlations are mainly present over the sea and only the United Kingdom and North West Africa are well covered. The monthly value of the NAO index cannot be used as the only predictor of the monthly wind drought days occurrence.

The patterns formed by Figure 34 are not included in the patterns from Figure 35, there are small zones, such as the Mediterranean that is significant with the wind drought indices but not the wind speed. It highlights, even more, the difference in the behaviour of the low-speed events and the monthly wind speed. It justifies the need for the BT and MBMT index as a metric for wind droughts.

3.3.2 Summary

Driven by the fact that the monthly NAO index is a good indicator of the seasonal wind speed, this work has been interested in the link between the NAO and the wind drought number of day. Unfortunately, this relationship has proved to be weak, revealing the difference in behaviour between the wind and the low-speed events. The NAO is not a predictor of the wind drought days and one needs the BT and the MBMT index to characterise the wind droughts.

4 Conclusion

In this work, two indices were developed in order to characterise the occurrences of low wind speed events in a season, the Below Threshold index and its extension the Mean Below Mean Threshold definition. A solid linear relationship of 1.6 has been found between the two indices at the seasonal scale. Consequently, it may be sufficient for a seasonal forecast to only compute the BT index and save the consequent computation time that would have been used for the MBMT. However, the systematic errors affecting the seasonal predictions of the two indices show different

patterns. While the BT index shows to be fairly unbiased, the MBMT shows stronger patterns between the climatology of the predictions and the reanalysis that are clued by their climatologies. It demonstrates that the MBMT completes the BT information given on the wind nature during a month.

Generally, significant trends of wind droughts days are following significant trends of seasonal wind speed. Over the tropical regions, there is an increase in the mean wind speed over the last decades that results in the reduction of wind drought days. The inter-annual fluctuations of the mean seasonal wind speed are associated with the opposite changes in the wind drought days, but the correlation is not 1. Indeed, this thesis exhibits a correlation of -0.55 between the wind drought days (according to the BT index) and seasonal wind speed in winter. It also has been found that this weak correlation is locally dependent, underlying the different performance of the wind drought between the tropical and extra-tropical regions. The seasonal mean wind speed is then not an indicator for the occurrence of seasonal low wind speed events, and the specific indices created, are more useful to better describe the inter-annual variability of the wind droughts.

NAO which is the main phenomenon driving the wind speed over Europe shows low correlations with the wind drought indices. This suggests that the wind droughts in winter cannot only be explained in terms of this pattern and that the link with other teleconnections should be investigated. Finally, the forecast quality assessment shows better performance of the forecasts for seasonal means than for monthly values. In addition, SEAS5 quality of forecast is decreasing from 15% after the first usable month.

The very good representation of the El Niño Southern Oscillation (ENSO) phenomenon by the seasonal prediction systems [9] might lead to the high RPSS score obtained for the wind drought indices over the equatorial Pacific. Moreover, locations that are strongly influenced by ENSO through teleconnections such as Northern Brazil and Eastern Africa show a fair skill. Also visible, relevant regions for the wind industry as North America and Europe, show positive RPSS values. The seasonal predictions of the wind droughts could then benefit in that areas.

However, the SEAS5 seasonal forecasts of wind droughts are unskilful in most of the regions. Of course, the robustness of this result needs to be confirmed by the assessment of other seasonal prediction systems and other observational references. It has been shown that a multimodel combination of seasonal ensemble forecasts can lead to better forecast quality than the best single forecast system. Besides error compensation, the multimodel combination also improves consistency and reliability [45].

Further development

A sketch of the computation time behaviour of the indices has been done and has underlaid the difference in results according to the specific dimension size (along threshold or orthogonal) of both indices. However, this results lack of precision and need a more careful study, with more points and less uncertainty. In addition, the study has shown a saturation of the MBMT wind drought days equal to the seasonal time length. It therefore suggests that some wind droughts are wider than

the seasonal scale to be studied. The saturation also hide the full distribution of the MBMT days according to the BT ones and the de-saturation with a longer threshold dimension scale would be a benefit for exhibiting the behaviour.

With the limited time, some choices have been made to have a coherent analysis despite the robustness of some results. It has been preferred to have a global overview on the prediction assessment of the low wind speed events but with only one reanalysis and Forecast system. However, this work shows limits at some points. For instance, the low positive skill shown on regions where the wind industry is well implemented, draws attention and further analysis should examine the predictability of the wind drought days over these areas.

The behaviour of the two indices is not completely understood. Next studies should look more precisely at the dependence of the MBMT on the distribution shape (Gaussian, Weibull, etc...) and the 1-day difference of the time series concerned. Also, the behaviour of the indices is likely to be depending on the daily threshold definition. As a first shot this work used a percentile threshold, but more complex and thoughtful threshold can certainly be better scoped on the interest of wind industries.

The SEAS5 seasonal predictions of wind droughts have been assessed. However, the performance relative to the duration and intensity of these droughts, i.e. the number of consecutive days with low wind speed values should be also explored. Indeed, it is reasonable to think that small consecutive-days events are more frequent and stochastic. In addition, the link between the wind drought length and the resilience of the energy storage and wind industries should be studied in order to extract what types of drought is the most problematic for the users. The difference in wind drought's prediction handling according to the strength event has to be reviewed.

Even though ERA5 is considered the best reanalysis representing the surface wind [32] and that SEAS5 has been chosen to be the most critical system, the very first step to go forward of this work would be to generalise the forecasts assessment for various reanalysis and forecasts systems, thus performing multimodel analysis. As a reanalysis MERRA-2 also performs well in the representation of surface winds [32] while the seasonal forecast systems from the Copernicus Climate Change initiative can be exploited for the construction of an effective multimodel.

The NAO has little influence on the wind droughts, however, it would be relevant to study other important large-scale patterns that act as predictability sources at seasonal timescales (e.g. ENSO, Euro-Atlantic teleconnections, Pacific-North American teleconnection, Indian Ocean Dipole, etc).

References

[1] V. Torralba-Fernández, "Seasonal climate prediction for the wind energy sector: methods and tools for the development of a climate service," p. 297,

[2] G. S. Callendar, "The artificial production of carbon dioxide and its influence on temperature," *Quarterly Journal of the Royal Meteorological Society*, vol. 64, no. 275, pp. 223–240, 1938, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.49706427503, ISSN: 1477-

870X. DOI: 10.1002/qj.49706427503. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/qj.49706427503 (visited on 09/12/2022).

[3] H. Ritchie, M. Roser, and P. Rosado, "Energy," *Our World in Data*, Nov. 28, 2020. [Online]. Available: https://ourworldindata.org/electricity-mix (visited on 08/20/2022).

[4] A. Soret *et al.*, "Sub-seasonal to seasonal climate predictions for wind energy forecasting," *Journal of Physics: Conference Series*, vol. 1222, no. 1, p. 012 009, May 2019, Number: 1 Publisher: IOP Publishing, ISSN: 1742-6596. DOI: 10.1088/1742-6596/1222/1/012009. [Online]. Available: https://doi.org/10.1088/1742-6596/1222/1/012009 (visited on 03/16/2022).

[5] H. Bloomfield. "Climate change, "wind droughts" and the implications for wind energy," Energy Post. (Nov. 15, 2021), [Online]. Available: https://energypost.eu/climate-change-wind-droughts-and-the-implications-for-wind-energy/ (visited on 09/11/2022).

[6] "Factsheets - earth system services." (), [Online]. Available: https://ess.bsc.es/portfolio/factsheets (visited on 07/07/2022).

[7] W. M. Organization (WMO) and W. M. O. (WMO), *The nature and theory of the general circulation of the atmosphere*, ser. WMO. Geneva: WMO, 1967, 161 p.

[8] A. Mariotti, P. M. Ruti, and M. Rixen, "Progress in subseasonal to seasonal prediction through a joint weather and climate community effort," *npj Climate and Atmospheric Science*, vol. 1, no. 1, pp. 1–4, Mar. 26, 2018, Number: 1 Publisher: Nature Publishing Group, ISSN: 2397-3722. DOI: 10.1038/s41612-018-0014-z. [Online]. Available: https://www.nature.com/articles/s41612-018-0014-z (visited on 07/07/2022).

[9] F. J. Doblas-Reyes, J. García-Serrano, F. Lienert, A. P. Biescas, and L. R. L. Rodrigues, "Seasonal climate predictability and forecasting: Status and prospects," *WIREs Climate Change*, vol. 4, no. 4, pp. 245–268, 2013, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcc.217, ISSN: 1757-7799. DOI: 10.1002/wcc.217. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/wcc.217 (visited on 09/14/2022).

[10] Vizzuality. "Global power plant database." (), [Online]. Available: https://resourcewatch.org/data/explore/Powerwatch (visited on 08/28/2022).

[11] "Global circulation patterns - met office." (), [Online]. Available: https://www.metoffice.gov.uk/weather/learn-about/weather/atmosphere/global-circulation-patterns (visited on 09/01/2022).

[12] "What is global atmospheric circulation?" Internet Geography. (), [Online]. Available: https://www.internetgeography.net/topics/what-is-global-atmospheric-circulation/ (visited on 09/10/2022).

[13] H. Su *et al.*, "Tightening of tropical ascent and high clouds key to precipitation change in a warmer climate," *Nature Communications*, vol. 8, Jun. 7, 2017. DOI: 10.1038/ncomms15771.

[14] Dashamlav. "Atmospheric circulation: Hadley, ferrel and polar cells," Dashamlav. (Jan. 29, 2021), [Online]. Available: https://dashamlav.com/atmospheric-circulation-hadley-ferrel-polar-cells/ (visited on 09/11/2022).

[15] W. Anderson, "Overturning circulations (or: Why the hadley and ferrel cells exist)," p. 14,

[16] "Pourquoi les saisons météorologiques diffèrent-elles des saisons astronomiques ? | météo-france." (), [Online]. Available: https://meteofrance.com/magazine/meteo-questions/pourquoi-les-saisons-meteorologiques-different-elles-des-saisons-astronomiques (visited on 09/11/2022).

[17] *Intertropical convergence zone*, in *Wikipedia*, Page Version ID: 1102271189, Aug. 4, 2022. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Intertropical_Convergence_Zone&oldid=1102271189 (visited on 09/12/2022).

[18] D. Rind, "Latitudinal temperature gradients and climate change," *Journal of Geophysical Research: Atmospheres*, vol. 103, pp. 5943–5971, D6 1998, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/97JD03649, ISSN: 2156-2202. DOI: 10.1029/97JD03649. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1029/97JD03649 (visited on 09/12/2022).

[19] D. F. Mantsis, A. C. Clement, A. J. Broccoli, and M. P. Erb, "Climate feedbacks in response to changes in obliquity," *Journal of Climate*, vol. 24, no. 11, pp. 2830–2845, Jun. 1, 2011, Publisher: American Meteorological Society Section: Journal of Climate, ISSN: 0894-8755, 1520-0442. DOI: 10.1175/2010JCLI3986.1. [Online]. Available: https://journals.ametsoc.org/view/journals/clim/24/11/2010jcli3986.1.xml (visited on 09/12/2022).

[20] J. W. Hurrell, Y. Kushnir, G. Ottersen, and M. Visbeck, "An overview of the north atlantic oscillation," in *The North Atlantic Oscillation: Climatic Significance and Environmental Impact*, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/134GM01, American Geophysical Union (AGU), 2003, pp. 1–35, ISBN: 978-1-118-66903-7. DOI: 10.1029/134GM01. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1029/134GM01 (visited on 07/19/2022).

[21] A. G. Barnston and R. E. Livezey, "Classification, seasonality and persistence of low-frequency atmospheric circulation patterns," *Monthly Weather Review*, vol. 115, no. 6, pp. 1083–1126, Jun. 1, 1987, Publisher: American Meteorological Society Section: Monthly Weather Review, ISSN: 1520-0493, 0027-0644. DOI: 10.1175/1520-0493(1987)115<1083:CSAPOL>2.0.CO;2. [Online]. Available: https://journals.ametsoc.org/view/journals/mwre/115/6/1520-0493_1987_115_1083_csapol_2_0_co_2.xml (visited on 07/20/2022).

[22] "Climate prediction center - teleconnections: North atlantic oscillation." (), [Online]. Available: https://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/nao.shtml (visited on 07/19/2022).

[23] R. Clark, P. Bett, H. Thornton, and A. Scaife, "Skilful seasonal predictions for the european energy industry," *Environmental Research Letters*, vol. 12, p. 024 002, Feb. 1, 2017. DOI: 10.1088/1748-9326/aa57ab.

[24] W. Helen J., *Spectral analysis in r*, Jun. 8, 2010.

[25] Hurrell, "NAO index data provided by the climate analysis section (2003, updated regulary), boulder, USA,"

[26] H. Lin. "Predictability and prediction of the north atlantic oscillation." (2012), [Online]. Available: https://www.semanticscholar.org/paper/Predictability-and-Prediction-of-the-North-Atlantic-Lin/ed1cebf0928971c6096d8a10e54065f86f43abcf (visited on 08/21/2022).

[27] C.-Y. Liu, G.-R. Liu, T.-H. Lin, C.-C. Liu, H. Ren, and C.-C. Young, "Using surface stations to improve sounding retrievals from hyperspectral infrared instruments," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 11, pp. 6957–6963, Nov. 2014, Conference Name: IEEE Transactions on Geoscience and Remote Sensing, ISSN: 1558-0644. DOI: 10.1109/TGRS.2014.2305992.

[28] "Met office seasonal prediction system: GloSea5," Met Office. (Mar. 11, 2022), [Online]. Available: https://www.metoffice.gov.uk/research/climate/seasonal-to-decadal/gpc-outlooks/user-guide/technical-glosea5 (visited on 03/11/2022).

[29] P. Grace, *Analysis, reanalysis, forecast—what's the difference?* Dec. 4, 2014.

[30] M. Dave and K. Sarah, *A practical guide to seasonal forecasts*, Jun. 2019.

[31] "Introduction to seasonal forecasting," Met Office. (), [Online]. Available: https://www.metoffice.gov.uk/research/climate/seasonal-to-decadal/gpc-outlooks/user-guide/background (visited on 08/20/2022).

[32] J. Ramon, L. Lledó, V. Torralba, A. Soret, and F. J. Doblas-Reyes, "What global reanalysis best represents near-surface winds?" *Quarterly Journal of the Royal Meteorological Society*, vol. 145, no. 724, pp. 3236–3251, 2019, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.3616, ISSN: 1477-870X. DOI: 10.1002/qj.3616. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/qj.3616 (visited on 07/07/2022).

[33] A. Guillory. "ERA5," ECMWF. (Nov. 3, 2017), [Online]. Available: https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5 (visited on 07/07/2022).

[34] I. Mahlstein, C. Spirig, M. A. Liniger, and C. Appenzeller, "Estimating daily climatologies for climate indices derived from climate model data and observations," *Journal of Geophysical Research: Atmospheres*, vol. 120, no. 7, pp. 2808–2818, 2015, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2014JD022327, ISSN: 2169-8996. DOI: 10.1002/2014JD022327. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/2014JD022327 (visited on 09/11/2022).

[35] N. Ohlendorf and W.-P. Schill, "Frequency and duration of low-wind-power events in germany," *Environmental Research Letters*, vol. 15, no. 8, p. 084 045, Aug. 2020, Number: 8 Publisher: IOP Publishing, ISSN: 1748-9326. DOI: 10.1088/1748-9326/ab91e9. [Online]. Available: https://doi.org/10.1088/1748-9326/ab91e9 (visited on 03/10/2022).

[36] R. Marcos, N. González-Reviriego, V. Torralba, A. Soret, and F. Doblas-Reyes, "Characterization of the near surface wind speed distribution at global scale: ERA-interim reanalysis and ECMWF seasonal forecasting system 4," *Climate Dynamics*, vol. 52, Mar. 1, 2019. DOI: 10.1007/s00382-018-4338-5.

[37] *Pearson correlation coefficient*, in *Wikipedia*, Page Version ID: 1106396722, Aug. 24, 2022. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Pearson_correlation_coefficient&oldid=1106396722 (visited on 08/27/2022).

[38] *Descriptions of the IRI climate forecast verification scores*, Jul. 2013. [Online]. Available: https://iri.columbia.edu/wp-content/uploads/2013/07/scoredescriptions.pdf.

[39] I. R. Young, S. Zieger, and A. V. Babanin, "Global trends in wind speed and wave height," *Science*, vol. 332, no. 6028, pp. 451–455, Apr. 22, 2011, Publisher: American Association for the Advancement of Science. DOI: 10.1126/science.1197219. [Online]. Available: https://www.science.org/doi/10.1126/science.1197219 (visited on 09/14/2022).

[40] C. W. Zheng, C. Y. Li, J. Pan, M. Y. Liu, and L. L. Xia, "An overview of global ocean wind energy resource evaluations," *Renewable and Sustainable Energy Reviews*, vol. 53, pp. 1240–1251, Jan. 1, 2016, ISSN: 1364-0321. DOI: 10.1016/j.rser.2015.09.063. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032115010333 (visited on 09/15/2022).

[41] M. L. L'Heureux, S. Lee, and B. Lyon, "Recent multidecadal strengthening of the walker circulation across the tropical pacific," *Nature Climate Change*, vol. 3, no. 6, pp. 571–576, Jun. 2013, Number: 6 Publisher: Nature Publishing Group, ISSN: 1758-6798. DOI: 10.1038/nclimate1840. [Online]. Available: https://www.nature.com/articles/nclimate1840 (visited on 09/15/2022).

[42] M. H. England *et al.*, "Recent intensification of wind-driven circulation in the pacific and the ongoing warming hiatus," *Nature Climate Change*, vol. 4, no. 3, pp. 222–227, Mar. 2014, Number: 3 Publisher: Nature Publishing Group, ISSN: 1758-6798. DOI: 10.1038/nclimate2106. [Online]. Available: https://www.nature.com/articles/nclimate2106 (visited on 09/14/2022).

[43] *Walker circulation*, in *Wikipedia*, Page Version ID: 1083460103, Apr. 18, 2022. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Walker_circulation&oldid=1083460103 (visited on 09/16/2022).

[44] I. T. Jolliffe and D. B. Stephenson, *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. John Wiley & Sons, Jan. 25, 2012, 304 pp., Google-Books-ID: DCxsKQeaBH8C, ISBN: 978-1-119-96107-9.

[45] R. Hagedorn, F. J. Doblas-Reyes, and T. Palmer, "The rationale behind the success of multi-model ensembles in seasonal forecasting — i. basic concept," *Tellus A: Dynamic Meteorology and Oceanography*, vol. 57, no. 3, pp. 219–233, Jan. 1, 2005, Publisher: Taylor & Francis _eprint: https://doi.org/10.3402/tellusa.v57i3.14657, ISSN: null. DOI: 10.3402/tellusa.v57i3.14657. [Online]. Available: https://doi.org/10.3402/tellusa.v57i3.14657 (visited on 09/14/2022).

Appendix A: Correlation

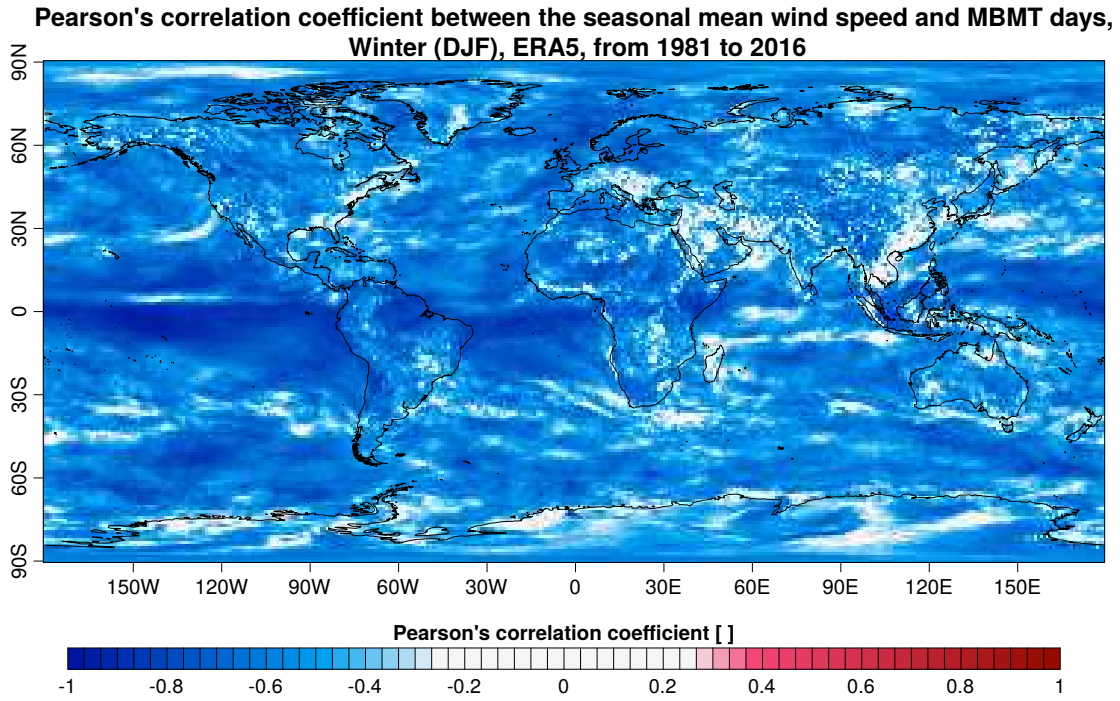


Figure 36: Correlation map of the seasonal mean wind speed with the number of wind drought days (BT index)

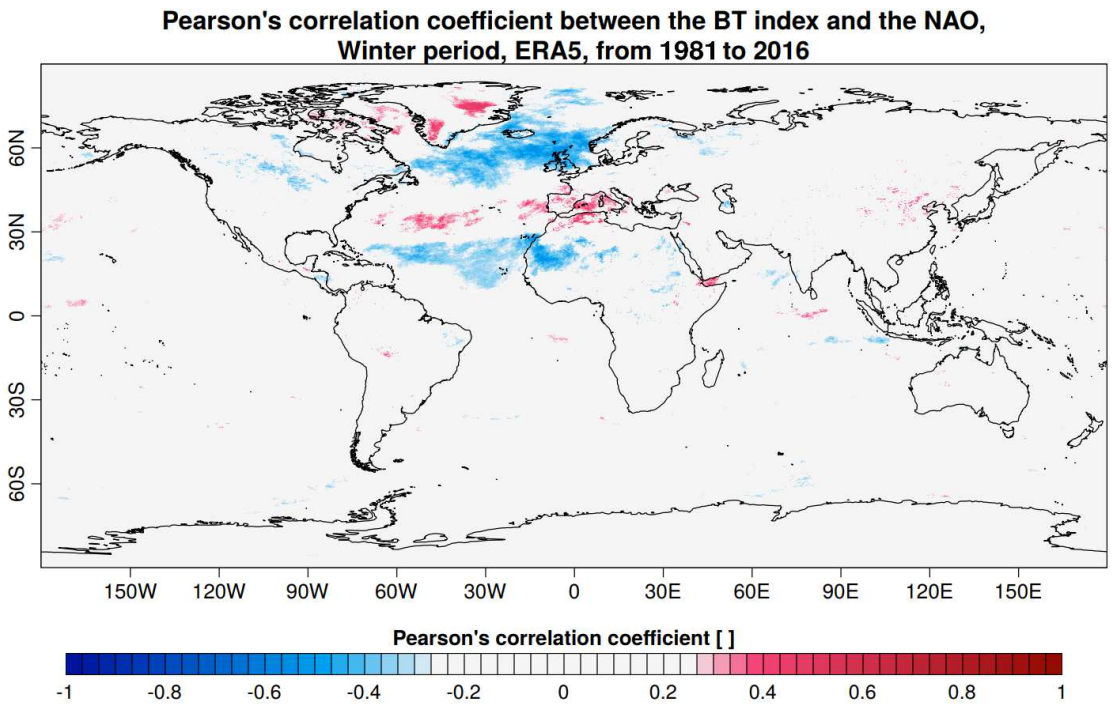


Figure 37: Pearson's correlation coefficient between the BT index and the NAO, Winter period, ERA5, from 1981 to 2016

Appendix B: Trends

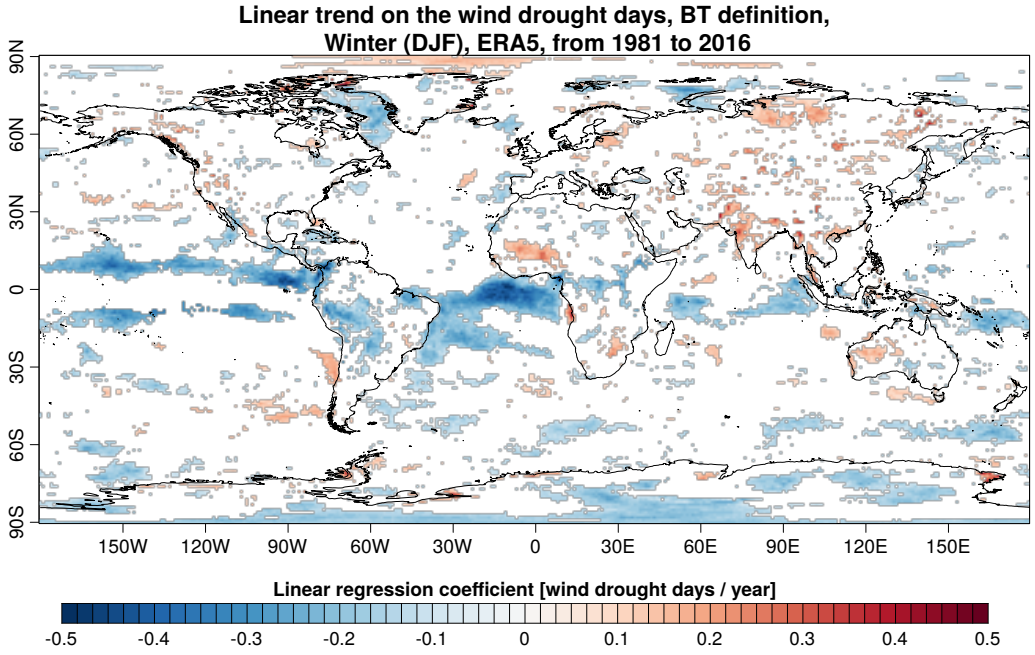


Figure 38: Occurence trend on the wind drought days, BT definition ,ERA5, from 1981 to 2016

Appendix C: Statistical Tools

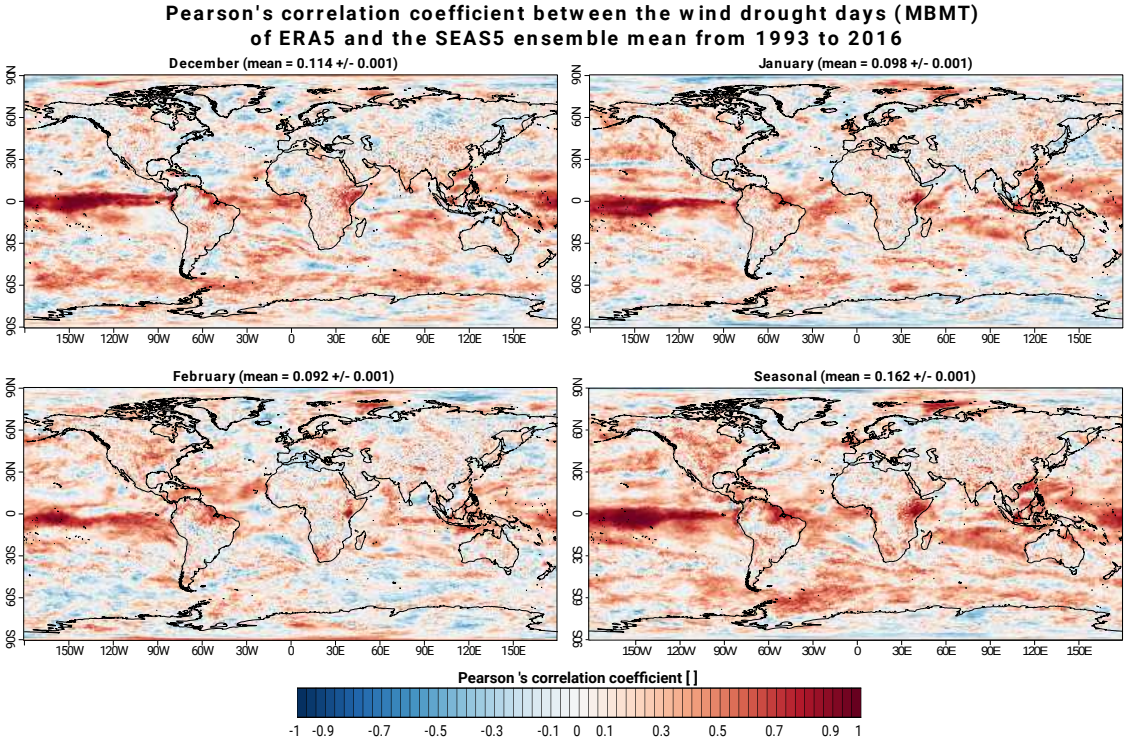


Figure 39: Pearson's correlation coefficient between the wind drought days (MBMT) of ERA5 and the SEAS 5 ensemble mean from 1993 to 2016.

Appendix D: Comparison between SEAS5 predictions and ERA5

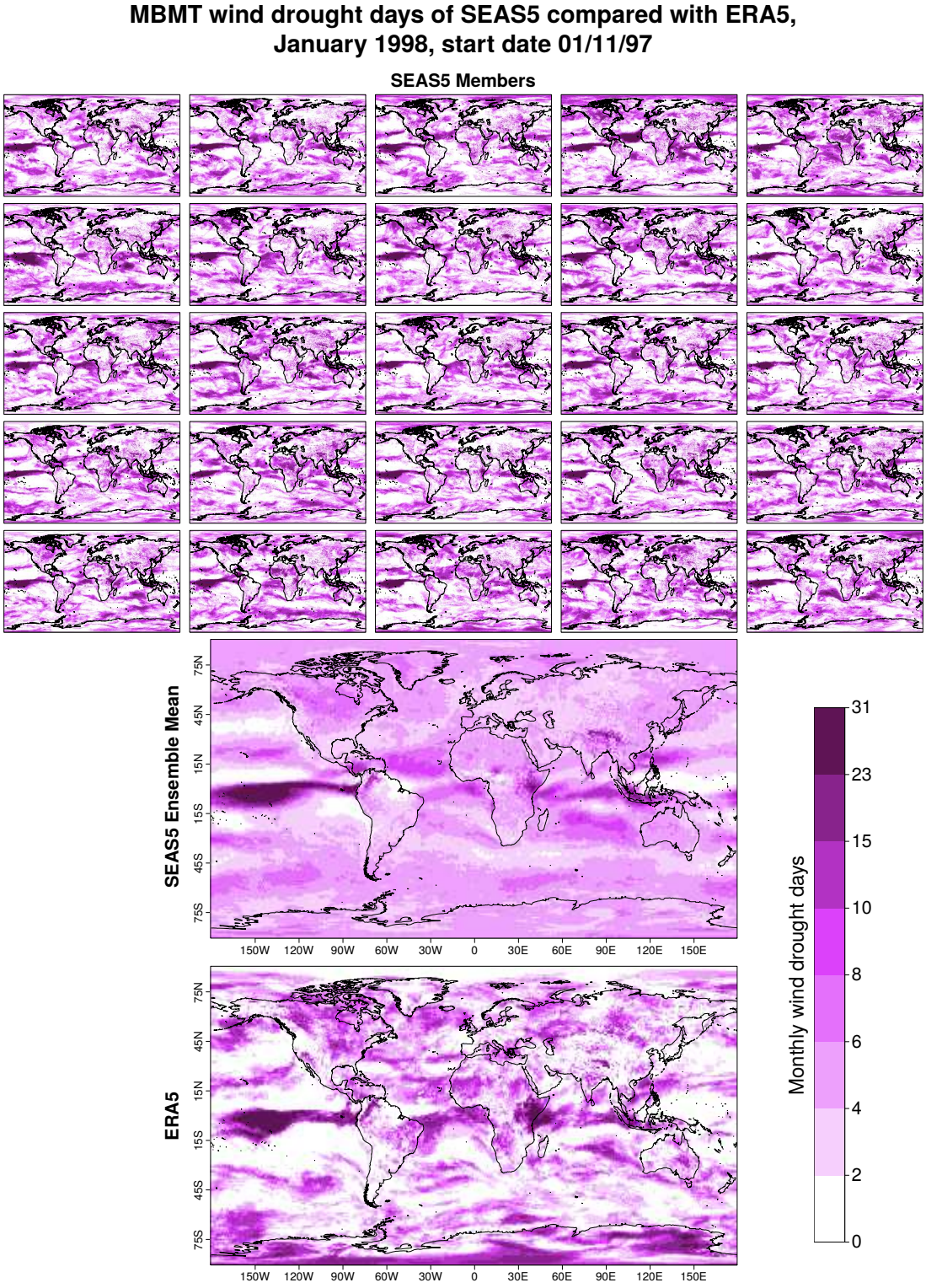


Figure 40: BT number of wind drought days predicted by the seasonal forecast system SEAS5 and the corresponding wind droughts obtained from ERA5 for the month of January 1998. First are represented all the different members (simulations) given by SEAS5. In general, the same behaviour is observed but the small changes in the initialisation change the results.

Appendix E: R scripts and Packages

This work has been done on Rstudio 99.9.9 with R 3.6.1, and has used the following libraries:

- s2dverification
- s2dv
- multiApply
- ClimProjDiags
- lubridate
- CSTools
- pointr
- ggplot2
- coe