EVALUATING THE QUALITY OF RECENT REANALYSES FOR WIND SPEED ALONG THE FRENCH COASTLINE

FRANCE ENERGIES MARINES

Rte Le réseau de transport d'électricité

EDF

EVALUATING THE QUALITY OF RECENT REANALYSES FOR WIND SPEED ALONG THE FRENCH COASTLINE

AUTHORS:
Anindita Patra, Tessa Chevallier, and Youen Kervella, France Energies Marines
Boutheina Oueslati, EDF
Laurent Dubus, RTE

Please cite this document as follows:
Patra A., Chevallier T., Kervella Y., Oueslati B. and Dubus L. Evaluation of recent regional climate reanalyses considering wind speed along the French coastline. 2024, 38 pages.
Published: September 2024
Graphic design: halynea.com
Cover page photo credit: France Energies Marines / Abode Firefly

Table of contents

List of figures

List of tables

EXECUTIVE SUMMARY

In this study, the accuracy of 10 m and 100 m wind speed estimates from several climate reanalyses—ERA5, ERA5 corrected by COSMO-REA6, COSMO-REA6, and CERRA—is evaluated along the various coastlines of France. The reference datasets consist of SYNOP wind measurements from Météo-France at 10 m height and LIDAR measurements at 100 m height above sea level. The interannual variability, seasonal and diurnal cycles, wind speed distribution, and extremes are assessed using a range of statistical metrics. These include bias, correlation, normalized standard deviation, diurnal cycle representation, frequency distribution, interannual variability, and extreme value analysis. Results show that CERRA is the most relevant reanalysis for most stations along the French coastline.

GLOSSARY

CC Climate Change
CDS Climate Data Store
CDF Cumulative Distribution Function
CDF-t Cumulative Distribution Function - transform
CERRA Copernicus Regional Reanalysis for Europe
CMIP Coupled Model Intercomparison Project
CORDEX Coordinated Regional Climate Downscaling Experiment
ECMWF European Centre for Medium-Range Weather Forecasts
ESGF Earth System Grid Federation
GCM Global Climate Model
GEV Generalized Extreme Value distribution
NWP Numerical Weather Prediction
EDA Ensemble Data Assimilation
DGEC Direction Générale de l'Énergie et du Climat
LIDAR Light Detection and Ranging
SYNOP Surface Synoptic Observations
IEA International Energy Agency
IPCC Intergovernmental Panel on Climate Change
PBIAS percentage bias
NSTD Normalized standard deviation
RMSE Root Mean Square Error
NOAA National Oceanic and Atmospheric Administration
PDF Probability Density Function
RCM Regional Climate Model
RCP Representative Concentration Pathway
WCRP World Climate Research Programme
WRA Wind Resource Assessment
WRF Weather Research and Forecasting

I. INTRODUCTION

France has a unique electricity mix, with nuclear power accounting for 71% of production in 2020, followed by renewables (wind, solar, hydro, bioenergy) at 21%, and fossil fuels at 8% (RTE, 2020). As the share of wind power in France's electricity mix continues to grow (RTE, 2021), the sensitivity of the electrical system to climate variability also increases (Bloomfield et al., 2016). France possesses the second-largest offshore wind resource potential in Europe, behind the UK, making it a promising candidate for the development of offshore wind farms. It is therefore essential to investigate France's offshore wind resource and its long-term evolution over the lifespan of wind farm installations.

Due to the sparse availability of wind measurements over coastal and offshore regions—both spatially and temporally—assessments must rely on high-resolution, gap-free wind data from numerical model simulations. Reanalysis datasets offer the advantage of homogeneous spatiotemporal coverage, enabling the characterization of wind resource variability across large domains and the investigation of long-term trends. However, a thorough evaluation of these datasets is necessary to identify potential limitations, such as systemic model errors (Bloomfield et al., 2016), coarse spatial resolution (Dawkins, 2019), and misrepresentation of extreme events (Cannon et al., 2015).

Several studies have assessed reanalyses for wind resource evaluation (Carvalho et al., 2014a; Miao et al., 2020; Jourdier, 2020; Potisomporn et al., 2023). These studies generally report decent correlations (mean Pearson R > 0.75) between reanalyses and observations, but also highlight significant mean biases (often > 0.3 m/s) and RMSE values exceeding 3 m/s (Potisomporn et al., 2023).

Global reanalyses such as ERA5 have shown better performance in terms of mean wind speed bias and RMSE compared to other global products (Gualtieri, 2022; Fan et al., 2020). However, regional reanalyses should be preferred—where available—to capture fine-scale features critical for wind energy applications. Fan et al. (2020) recommend selecting different reanalysis products depending on the region to minimize discrepancies with observations. Regional datasets are particularly valuable in complex coastal and mountainous areas, where they better resolve terrain features, land-use discontinuities, and stability variations (Carvalho et al., 2014b). Moreover, improvements in global datasets can positively impact regional products, whose accuracy depends on the quality of the global reanalyses used as boundary conditions (Clarke et al., 2021).

Jourdier (2020) evaluated several reanalyses using wind mast data (55–100 m above ground) over continental France. In contrast, this study focuses on surface winds along the French coastline, incorporates newly available regional reanalyses, and extends the evaluation period to 1995–2018 to investigate long-term trends. The reference measurements (see Data section 1) differ from those used in previous studies. In addition to mean wind speeds and percentiles, seasonal and diurnal cycles, as well as seasonal variations, are also assessed.

Evaluating model performance is challenging due to the non-homogeneous temporal coverage of observational data, which varies by station. While many metrics can be computed, their relevance depends on the intended application. Metrics calculated over the entire period may not fully capture seasonal or diurnal patterns and must therefore be evaluated individually at each station. Jourdier (2020) emphasized that diurnal cycle evaluation is particularly insightful, as it reveals intra-day biases and reflects how well boundary layer processes are parameterized in the models. Accordingly, this study includes a detailed comparison of seasonal and diurnal cycles between reanalyses and observations.

The global reanalyses ERA5 and ERA5 corrected by COSMO-REA6, as well as the regional reanalyses COSMO-REA6 and the Copernicus Regional Reanalysis for Europe (CERRA), are considered. The regional simulations benefit from higher spatial resolution compared to global datasets. All reanalyses are evaluated against in-situ measurements: surface wind data from automated weather stations and 100 m wind speeds from floating LIDAR stations. The quality of each

reanalysis is assessed along three major coastlines of mainland France: the English Channel, the Atlantic Ocean, and the Mediterranean Sea.

Figure 1. The planned steps to deliver 2CNOW project

The 2CNOW project follows a multi-step methodology, as illustrated in Figure 1. This document presents the analyses completed in Steps 1 to 3, focusing on wind data. The main objective is to identify the most suitable reanalysis dataset, which will subsequently be used to compare and correct wind speed distributions derived from Coupled Model Intercomparison Project Phase 6 (CMIP6) simulations at selected offshore locations.

The structure of the manuscript is as follows:

- **Section II** introduces the datasets, including both observational measurements and reanalyses, and describes the quality control procedures and statistical methods used for evaluation.
- **Section III** presents the results for 10 m and 100 m wind speeds, detailing error metrics at each station, seasonal and diurnal cycles, interannual variability, wind speed distributions, and extreme values (low and high percentiles).
- **Section IV** summarizes the main conclusions and outlines perspectives for future work.

II. DATA AND METHODS

1 Data

This section describes the measurements and reanalysis data (Table 1) evaluated in this study.

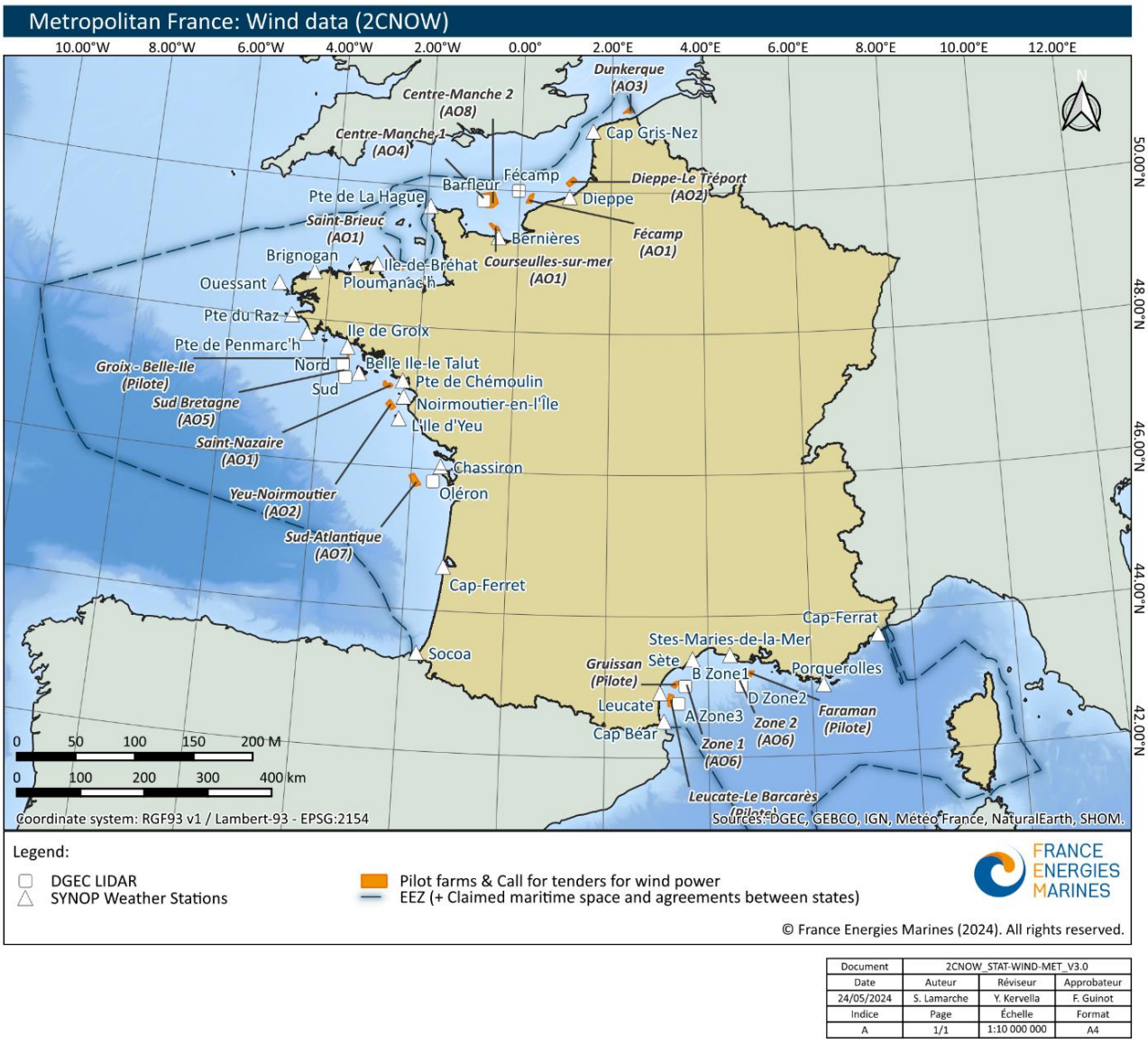


Figure 2. SYNOP stations and LIDAR stations around France used in this study

1.1 SYNOP

The SYNOP wind measurements are made in accordance with World Meteorological Organization (WMO) standards at 10 meters above terrain using verified and calibrated anemometers. As per WMO standards, the automated weather stations (AWS) are representative of the surrounding area, (mostly, not always) away from buildings and trees, which would otherwise influence the measurements. The description of the French stations and the available parameters can be found here: https://meteo.data.gouv.fr/datasets/6569b4473bedf2e7abad3b72.

The stations are maintained regularly and are calibrated to ensure correct readings. All stations send data every 10 minutes, while they report back every hour with the highest measurements in the latest 1-hour period (https://meteodata.dk/windsynop.php). We have selected all the coastal stations (closest to the coast, a few km maximum from the coastline) around France and then we have applied a quality control (section 2.1) to keep only the best stations for our comparisons (Figure 2).

1.2 ERA5

ERA5 is the fifth generation ECMWF atmospheric reanalysis that is based on the Integrated Forecasting System (IFS) Cy41r2 (Hersbach et al., 2020). The horizontal resolution is 31 km (0.28°) and the temporal resolution is hourly during the period 1940 to present. On the vertical, there are 137 levels from the surface to the model top located at 0.01 hPa or 80 km altitude. 4DVar is used for the assimilation of a variety of conventional and satellite based observational data. Quality-assured monthly updates of ERA5 are published within 3 months of real time. Access to the data is provided by the Meteorological Archival and Retrieval System (MARS) of ECMWF and the Copernicus Climate Data Store (CDS).

The ERA5 dataset contains one (hourly, 31 km) high resolution realisation (referred to as "reanalysis" or "HRES") and a reduced resolution ten-member ensemble (referred to as "ensemble" or "EDA"). Generally, the data consists of “analyses” and “forecasts”, initialized twice daily from analyses at 06 and 18 UTC. Using data assimilation, each previous forecast is combined with the newly available observations in an optimal way to produce a new best estimate of the state of the atmosphere, called the analysis, from which an updated, improved forecast is issued.

An issue with the diurnal cycles of wind speeds in ERA5, associated with the assimilation process, (affects mostly low latitude oceanic regions but also Europe and North America), was reported by Jourdier (2020). To overcome this issue, ERA5 forecasts (“fc”) are used here instead of analysis (“an”) data.

The global models like ERA5 have rather low spatial resolution. A resolution of a few dozens of kilometers is not sufficient to capture small-scale features which may be important for accurately simulating wind energy production at local scales. Thus, higher resolution, limited spatial extent models are also investigated.

1.3 COSMO-REA6

The regional European reanalysis COSMO-REA6 was developed within the Hans-Ertel-Centre for Weather Research and uses operational NWP model COSMO version 4.25 of the German Meteorological Service (DWD) with initial and boundary conditions from the ERA-Interim reanalysis (Dee et al., 2011) at 6-hourly interval. The model domain is adjusted to match the EURO-CORDEX region. The configuration used incorporates a horizontal resolution of 6 km with a non-hydrostatic model formulation. In the vertical, the terrain following hybrid coordinate system consists of 40 main levels.

For data assimilation, COSMO-REA6 employs a Newtonian relaxation scheme (nudging) to combine prognostic model variables with observations. Observations are assimilated stemming from radiosondes, SYNOP stations, ships, buoys or aircraft. Satellite data is not assimilated by COSMO-REA6 (Bollmeyer et al., 2015). The COSMO-REA6 data is publicly available as part of DWD’s open data (https://opendata.dwd.de/climate_environment/REA). The COSMO-REA6 reanalysis is chosen because of its high spatial resolution and availability from 1995 to 2018 at hourly frequency.

1.4 CERRA

The CERRA system is based on the HARMONIE-ALADIN data assimilation system. The model runs with a 5.5 km horizontal grid spacing and with 106 vertical levels with the domain larger than the Euro-CORDEX domain. The system uses lateral boundary conditions obtained from the ERA5 global reanalysis. It runs with a 3 h cycle producing 6 h forecasts at all analysis times except at 00:00 and 12:00 UTC where 30 h forecasts are produced. The CERRA system employs the 3D variational analysis (3D-VAR) method. At fixed points in time the model state is adjusted based on the observed state, considering the error statistics of both model and observations. The CERRA high-resolution system has been running

with eight assimilation cycles per day performing analyses ("an") at 00, 03, 06, 09, 12, 15, 18 and 21 UTC. The forecast lengths vary between 6 and 30 hours depending on the starting hour. The forecast model is then started from the analysis, and the output is saved hourly for the first six hours. Due to the forecast lengths, the forecasts are overlapping and for every hour of the day data can be chosen from the forecasts ("fc"). More information can be found in (Schimanke et al., 2021b, a) and via the link:
https://cds.climate.copernicus.eu/cdsapp#!/dataset/10.24381/cds.622a565a?tab=overview

After the assimilation of the observations (the analysis), the model state is not completely balanced dynamically, meaning that the model atmosphere might contain high-frequency waves. That might affect the quality of the forecasts closest to the analysis time if they are not rapidly damped. On the other hand, longer forecasts might veer away from the real weather, e.g., due to shortcomings in the model parametrizations, the initial conditions and/or inaccurate boundary conditions. The comparisons completed by Jourdier et al. (2023) suggest that it is better to use the forecasts with lead times of 4, 5 and 6 hours for each time step (00, 03, 06… 21). Thus, we have an hourly dataset from CERRA-"fc" (Figure 3, green boxes). It is currently available from September 1984 to June 2021.

Analysis		D																					
Day	H	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
D	0	A	+1	+2	+3	+4	+5	+6			+9			+12			+15			+18			+21		
D	3				A	+1	+2	+3	+4	+5	+6														
D	6							A	+1	+2	+3	+4	+5	+6											
D	9										A	+1	+2	+3	+4	+5	+6								
D	12													A	+1	+2	+3	+4	+5	+6			+9		
D	15																A	+1	+2	+3	+4	+5	+6		
D	18																			A	+1	+2	+3	+4	+5
D	21																						A	+1	+2

Figure 3. The analysis hours ("A") at each day (in red) and the forecasts chosen (in green) for ("A" +4, "A" +5, "A" +6) in the study for CERRA (From Jourdier et al., 2023)

1.5 ERA5 corrected by COSMO-REA6

We also use ERA5 bias-corrected by COSMO-REA6. The wind speed distribution (Cumulative distribution function) of ERA5 is corrected by that of COSMO-REA6 at a daily time step, keeping the diurnal cycles of ERA5. The dataset was produced by the Copernicus Climate Change Service Contract C3S2 412 Enhanced Operational Services for the Energy Sector and obtained from Reseau de Transport d'Electricité (RTE, personal communications).

Table 1: Summary of reanalyses characteristics.

Reanalysis	Spatial resolution	Temporal resolution	Height levels
ERA5	31 km	1 hour	10 m and 100 m
ERA5 corrected by COSMO-REA6	31 km		
COSMO-REA6	6 km		
CERRA	5.5 km		

1.6 LIDAR

LIDAR observations are available from Météo-France at 7 offshore stations (Figure 2). They are part of the plan of the Direction Générale de l'Énergie et du Climat (DGEC) to monitor wind speed for the development of offshore wind power projects. Wind data were acquired over a period of about 1 year using a buoy-mounted LIDAR (Light Detection and Ranging) system from LEOSPHERE (Windcube V2 Offshore 8.66 for buoys, model WLS866-4). The LIDAR sends an infrared laser pulse towards the atmosphere and measures the reflections off aerosol particles, which makes it possible to calculate the speed and direction of the wind based on the shift in their wavelength due to the Doppler effect. These measurements are taken at 11 different heights, from 40 to 200 meters (40, 50, 60, 80, 100, 120, 140, 150, 160, 180, 200 m).

2 Methods

2.1 Quality control of SYNOP measurements

Each "raw" observation exceeding the daily mean value by a factor of 3 times the daily standard deviation is removed as an outlier (Bentamy and Croize-Fillon, 2014). As the SYNOP measurements have several data gaps, we employ a data filtering method. First, sudden changes in data distribution were filtered by removing the older data. Those jumps happened when a new sensor was put in place. The new sensors are believed to be more precise and less subject to errors than the old ones, thus the older data is not used here. We found only one jump for Sete, around 1996. To deal with the missing values, we calculated the number of missing observations per month in each year. We allowed a maximum of 33% missing values in each month: it corresponds to two missing values per day for 3-hourly data, and eight for hourly data. The years with at least a month above this threshold were removed to prevent a bias in the seasonal cycle (Figure 4). The years where the number of missing values per month was constant were also excluded: after a manual check, those missing values happened daily at the same time. Thus, keeping them would lead to a bias in the diurnal cycle. This last issue occurred only in the 3-hourly data (before 1990) when the data collection was manual.

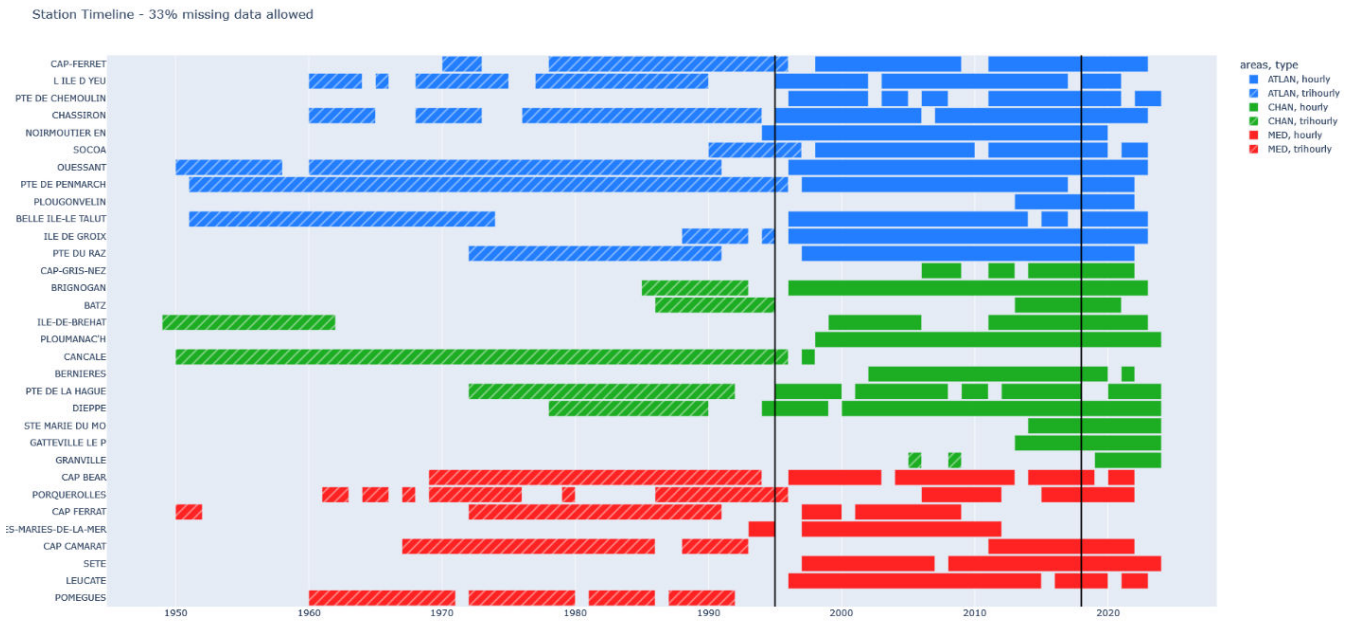


Figure 4 : Timeline of the stations after removing the missing values. Green, blue, red are for the English Channel, Atlantic and Mediterranean stations, respectively.

The stations with at least 10 years of data between 1995 and 2018 were maintained, resulting in a total of 24 stations (Figure 2). We chose 24 SYNOP stations around the French coastline: 7 stations for the English Channel coast, 11 for the Atlantic coast and 6 for the Mediterranean coast.

2.2 Comparison Methods

To compare with SYNOP measurements, the reanalysis grid points are selected according to a nearest neighbour approach. It turns out that this approach provides better results than bi-linear interpolation of the surrounding 4 grid cells for the coastal locations. It is better to avoid linearly interpolating land and ocean points together at the coastal transition. The reanalysis data are selected during the common time period 1995-2018. Moreover, times when any dataset or observation (SYNOP/ LIDAR) is missing are removed from all-time series. Thereafter, the evaluation is done for common timestamps.

The diurnal cycle, seasonal cycle, inter-annual variability, and commutative density function are evaluated at each SYNOP and LIDAR location. Thereafter, a station average along each coastline is also calculated. The coastal zones in France are divided into three seafronts: the English Channel, the Atlantic Ocean, and the Mediterranean Sea. We present box plots for both the mean and extreme percentiles. The error metrics (percentage bias: PBIAS, normalized standard deviation ratio: STDratio, normalized root mean square error: NRMSE, Pearson product-moment correlation coefficients: CORR) used here are defined as below:

$$PBIAS = 100 \times \frac{\sum_{i=1}^{N} M_i - O_i}{\sum_{i=1}^{N} O_i}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(M_i - O_i)^2}{N}}$$

$$NRMSE = \frac{RMSE}{O_{max} - O_{min}}$$

$$STDratio = \frac{\sigma_M}{\sigma_O}$$

$$CORR = \frac{\sum_{i=1}^{N}(M_i - \bar{M})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^{N}(M_i - \bar{M})^2}\sqrt{\sum_{i=1}^{N}(O_i - \bar{O})^2}}$$

Here, M and O present reanalysis and reference measurements, N is the length of overlapped time, and σ is standard deviation.

III. RESULTS

1 Evaluation of Surface Wind Speed (10 m)

Surface wind speed at 10 m provided by the four reanalyses is evaluated for the different coastlines of France in comparison to the in-situ measurements. The purpose is to assess the performance of each reanalysis, by considering the metrics described above, and to identify the one showing the best skill for most of the stations.

1.1 Distribution of error metrics over stations

1.1.1 Percentage bias (PBIAS)

Figure 5 provides the overview of percentage biases (PBIAS) for 7 stations on the English Channel coast, 11 stations on the Atlantic coast, and 6 stations on the Mediterranean coast (see Figure 2 for station location). After computing the PBIAS at each station for the hourly time series (at common times between the measurements and reanalyses) for the 1995-2018 period (refer to Appendix section for metrics at each station), the distribution of PBIAS values is presented through Box–Whisker plots for each reanalysis at each coastline. For the 7 stations along the English Channel coastline, the comparison with the SYNOP measurements shows that the median bias is ~2-3 % (<0.2 m/s) for the reanalyses (Figure 5, left panel). For CERRA, 50% of the bias values (First (Q1) to third quartile (Q3) range indicated by the filled boxes) are within 10% (0.5 m/s) with a systematic bias toward positive values. The inter-quartile range (Q3-Q1) is higher for other reanalyses compared to CERRA. Clearly, bias values have improved in ERA5 corrected by COSMO-REA6 with respect to ERA5. COSMO-REA6, although a regional reanalysis, has a larger bias spread.

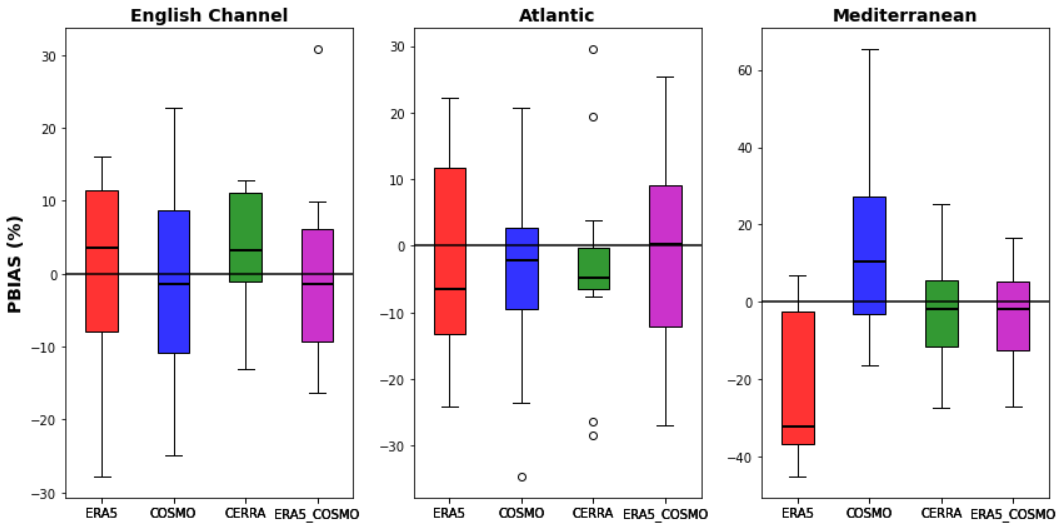


Figure 5. Box–Whisker plots for the bias distribution over the stations for 10 m wind speed from the reanalyses as compared to SYNOP. The red horizontal line shows median, and the box shows (25-75%) quartiles of the biases for different stations. Larger biases are indicated by the whiskers.

For the Atlantic (11 stations) and Mediterranean (6 stations) coastlines, CERRA has an inter-quartile range (Q3-Q1) within 10%. The percentage biases from ERA5 are the worst in the Mediterranean and have high values for the English Channel and Atlantic as well. The COSMO-REA6 performs well in the Atlantic, with a better mean but a larger interquartile spread than CERRA, and ERA5 corrected by COSMO-REA6 in the Mediterranean with similar results to CERRA. A larger underestimation by ERA5 (interquartile range: -2.4 to -0.16 m/s) over Southern France (also see Figure A1 in Appendix) is also reported by Jourdier (2020). Contrary to ERA5, COSMO-REA6 displays a larger overestimation (interquartile range: -0.19 to 1.17 m/s) for the Mediterranean coast. These findings hold true even when PBIAS values

are computed separately for the winter (December to February) and summer (June to August) seasons (Figure 18). In particular, the large overestimation by COSMO-REA6 is reduced in summer, and increased in winter (Figure 19). Overall CERRA presents the best results in terms of PBIAS, followed by ERA5 corrected by COSMO-REA6.

1.1.2 Normalized root mean square error (NRMSE)

In addition to the bias, which represents a systematic error, other error metrics, such as the Normalized root mean square error, correlation, normalized standard deviation (using hourly data) are computed. The NRMSE values of the reanalyses (Figure 6) are usually within 10%, except for ERA5 and COSMO-REA6 for the Mediterranean coastline, as high biases are also seen for these cases. The median NRMSE value from CERRA is <6% for the English Channel and Atlantic, and around 8% for the Mediterranean. The absolute RMSE values are <3 m/s for the English Channel and Atlantic coast and <5 m/s for the Mediterranean coast. Various studies which have evaluated reanalyses in different locations for wind speed found that most of the reanalyses show RMSEs higher than 3 m/s (Potisomporn et al., 2023). Overall CERRA presents the best results in terms of NRMSE, followed by COSMO-REA6.

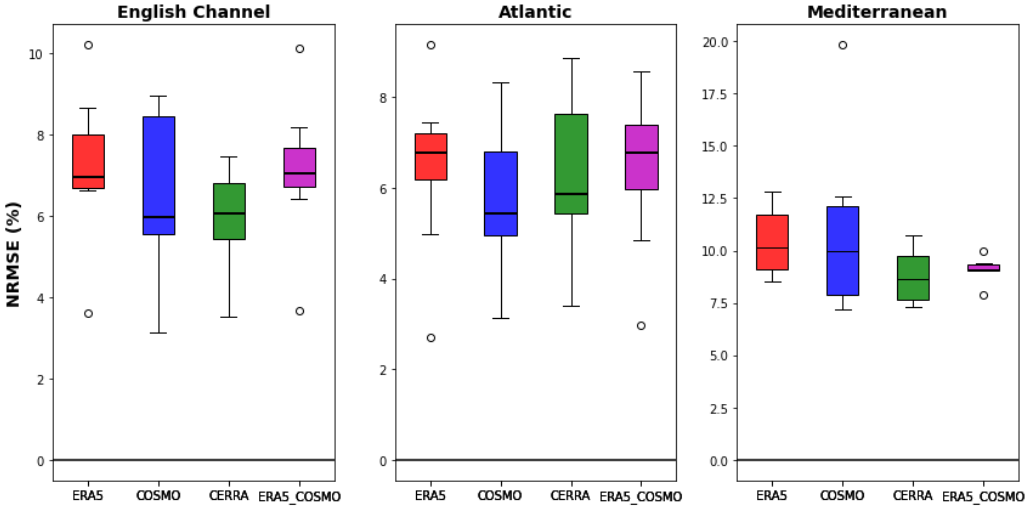


Figure 6: Box–Whisker plots for the normalized RMSE (%) distribution over the stations for 10 m wind speed from the reanalyses with compared to SYNOP

1.1.3 Correlation

In terms of correlations (Figure 7), the reanalyses show a decent correlation (> 0.75) with the measurements, which agrees with Potisomporn et al. (2023). The median value of CERRA (above 0.82) is always higher than that of the other reanalyses, except in the Atlantic, where COSMO-REA6 shows a slightly higher correlation. Slightly lower correlation values can be seen for ERA5 and ERA5 corrected by COSMO-REA6. The most eastern station in the Mediterranean, Cap Ferrat, has lowest correlation, especially in COSMO-REA6 (Figure A4). In general, lower correlations are observed along the Mediterranean coast (third quartiles drop to 0.7).

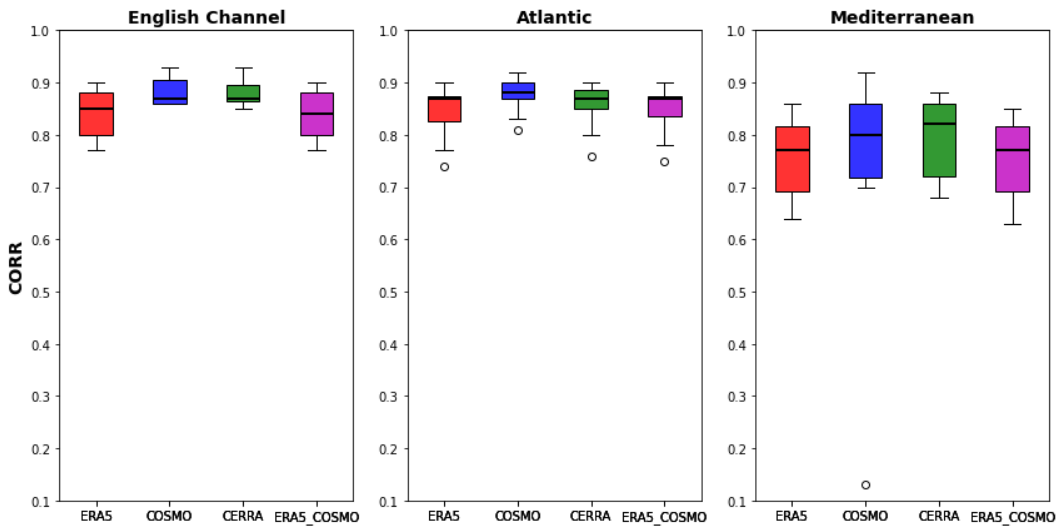


Figure 7. Box–Whisker plots for the distribution of Correlation Coefficients over all stations for 10 m wind speed from the reanalyses as compared to SYNOP. The red horizontal line shows the median, and the box shows (25-75%) quartiles. Larger values are indicated by the whiskers.

1.1.4 Normalized standard deviation

The normalized standard deviation depicts less variability (<1.0) in the reanalyses when compared to the SYNOP measurements (Figure 8). The median values are mostly below 1.0 for the reanalyses, except for COSMO-REA6, most notably along the Mediterranean coast. ERA5 has lower variability than the observations for all the coastal locations, and COSMO-REA6 has very high variability in the Mediterranean. However, CERRA represents the temporal variability well for the English Channel and Mediterranean coast. The first and third quartiles of the normalized standard deviation for CERRA are between 0.83 and 1.09, and for ERA5 corrected by COSMO-REA6 between 0.79 to 1.05. Potisomporn et al., (2023) explained ERA5's underestimation of the hourly wind variability caused by small-scale fluctuations that are not captured by the macro-scale driven ERA5 simulation.

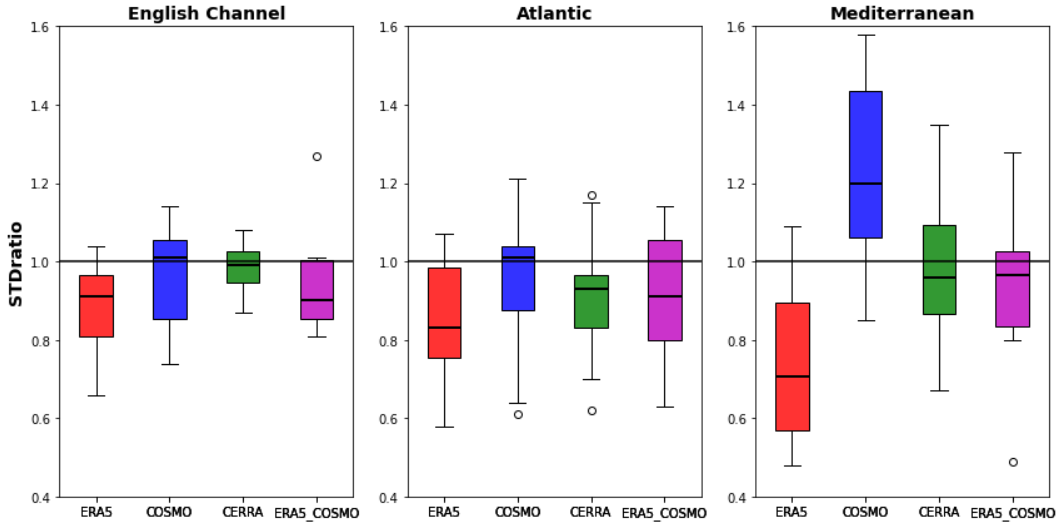


Figure 8. Box–Whisker plots for the distribution of Normalized standard deviation over all stations for 10 m wind speed from the reanalyses as compared to SYNOP. The red horizontal line shows median, and the box shows (25-75%) quartiles. Larger values are indicated by the whiskers.

1.1.5 Summary table of error metrics

The following tables (Table 2, Table 3 and Table 4) present the different scores for the 3 French seafronts. The best scores are indicated in bold and are highlighted in green. The last 2 lines present respectively the average of all stations for each score (note: for the PBIAS, it is the average of the absolute values that is calculated), and the number of stations for which the reanalysis obtains the best score.

Table 2: summary table of error metrics in the English Channel

Scores in English Channel	ERA5				COSMO				CERRA				ERA5 corrected by COSMO			
	PBIAS	NRMSE	STD ratio	CORR	PBIAS	NRMSE	STD ratio	CORR	PBIAS	NRMSE	STD ratio	CORR	PBIAS	NRMSE	STD ratio	CORR
Station	[%]	[%]	[]	[]	[%]	[%]	[]	[]	[%]	[%]	[]	[]	[%]	[%]	[]	[]
CAP-GRIS-NEZ	-27,8	10,2	0,7	**0,9**	-24,9	9,0	0,7	**0,9**	**-13,1**	**6,4**	**0,9**	**0,9**	-13,7	7,2	**0,9**	**0,9**
DIEPPE	14,0	7,4	**1,0**	0,8	-10,2	**5,7**	0,9	**0,9**	**9,7**	6,1	1,1	**0,9**	30,8	10,1	1,3	0,8
BERNIERES	-11,7	8,7	0,8	0,8	22,7	8,9	1,1	**0,9**	12,5	**7,2**	**1,0**	**0,9**	**-1,5**	8,2	0,8	0,8
PTE DE LA HAGUE	-4,1	3,6	0,9	**0,9**	**-1,4**	**3,2**	**1,0**	**0,9**	-2,4	3,5	1,0	**0,9**	2,5	3,7	**1,0**	**0,9**
ILE-DE-BREHAT	**3,6**	**7,0**	**0,9**	**0,9**	11,7	8,0	1,1	**0,9**	12,8	7,5	1,1	**0,9**	-4,8	7,0	0,9	**0,9**
PLOUMANAC'H	16,2	6,6	**1,0**	**0,9**	-11,4	5,4	0,9	**0,9**	**3,3**	**5,2**	**1,0**	**0,9**	-16,4	6,4	0,9	0,8
BRIGNOGAN	9,0	6,8	0,9	0,8	5,5	6,0	**1,0**	0,9	**0,2**	**5,7**	0,9	**0,9**	9,9	7,1	**1,0**	0,8
Seafront mean	**12,3**	**7,17**	**0,9**	**0,8**	**12,6**	**6,6**	**1,0**	**0,9**	**7,7**	**5,9**	**1,0**	**0,9**	**11,4**	**7,1**	**1,0**	**0,8**
Seafront best score	**1/7**	**1/7**	**3/7**	**4/7**	**1/7**	**2/7**	**2/7**	**5/7**	**4/7**	**4/7**	**3/7**	**7/7**	**1/7**	**0/7**	**3/7**	**3/7**

Table 3: summary table of error metrics in the Atlantic Ocean

Scores in Atlantic Ocean	ERA5				COSMO				CERRA				ERA5 corrected by COSMO			
	PBIAS	NRMSE	STD ratio	CORR	PBIAS	NRMSE	STD ratio	CORR	PBIAS	NRMSE	STD ratio	CORR	PBIAS	NRMSE	STD ratio	CORR
Station	[%]	[%]	[]	[]	[%]	[%]	[]	[]	[%]	[%]	[]	[]	[%]	[%]	[]	[]
OUESSANT-STIFF	-6,5	5,0	0,9	0,9	**-3,1**	**4,6**	**1,0**	**0,9**	-7,6	5,2	0,9	0,9	-3,9	4,9	**1,0**	0,9
PTE DU RAZ	**-9,8**	6,8	**0,8**	**0,9**	-12,5	**6,3**	**0,8**	**0,9**	-26,4	8,8	0,7	**0,9**	-10,9	6,8	**0,8**	**0,9**
PTE DE PENMARCH	-8,1	7,4	0,8	0,8	-0,9	**5,7**	**1,0**	**0,9**	-4,8	6,4	0,9	**0,9**	**0,4**	7,1	0,9	0,8
ILE DE GROIX	-17,4	6,3	0,8	**0,9**	**-2,2**	**5,5**	1,1	**0,9**	-5,6	5,7	**1,0**	**0,9**	-26,9	7,7	0,8	**0,9**
BELLE ILE-LE TALUT	12,8	6,9	**1,0**	**0,9**	**0,8**	**5,3**	**1,0**	**0,9**	-4,6	5,9	0,9	**0,9**	8,4	6,6	1,1	**0,9**
PTE DE CHEMOULIN	-24,1	9,2	0,6	**0,8**	-23,5	8,3	0,6	**0,8**	**-4,5**	**7,2**	**0,8**	**0,8**	-20,3	8,6	0,6	**0,8**
NOIRMOUTIER EN	22,2	7,4	**1,1**	**0,9**	**20,8**	**7,3**	**1,1**	**0,9**	29,5	8,9	1,2	**0,9**	25,4	8,0	**1,1**	**0,9**
L ILE D YEU	**10,8**	**2,7**	**1,1**	**0,9**	13,5	3,1	1,2	**0,9**	19,3	3,4	1,2	**0,9**	13,0	3,0	**1,1**	**0,9**
CHASSIRON	12,7	6,4	**1,0**	**0,9**	4,6	**5,4**	**1,0**	**0,9**	**3,8**	5,8	**1,0**	**0,9**	9,9	6,4	**1,0**	**0,9**
CAP-FERRET	-16,7	6,1	0,7	0,8	-6,7	**4,5**	**0,9**	**0,9**	**-5,5**	4,9	**0,9**	**0,9**	-13,4	5,6	0,8	0,8
SOCOA	**-2,3**	**7,0**	0,8	0,7	-34,6	8,3	0,6	**0,8**	-28,4	8,1	0,6	**0,8**	4,5	7,1	**0,9**	**0,8**
Seafront mean	**13,0**	**6,5**	**0,9**	**0,8**	**11,2**	**5,8**	**0,9**	**0,9**	**12,7**	**6,4**	**0,9**	**0,9**	**12,4**	**6,5**	**0,9**	**0,8**
Seafront best score	**3/11**	**2/11**	**5/11**	**7/11**	**4/11**	**8/11**	**7/11**	**11/11**	**3/11**	**1/11**	**4/11**	**10/11**	**1/11**	**0/11**	**6/11**	**8/11**

Table 4: summary table of error metrics in the Mediterranean Sea

Scores in Mediterranean Sea	ERA5				COSMO				CERRA				ERA5 corrected by COSMO			
	PBIAS	NRMSE	STD ratio	CORR	PBIAS	NRMSE	STD ratio	CORR	PBIAS	NRMSE	STD ratio	CORR	PBIAS	NRMSE	STD ratio	CORR
Station	[%]	[%]	[]	[]	[%]	[%]	[]	[]	[%]	[%]	[]	[]	[%]	[%]	[]	[]
LEUCATE	-36,9	12,1	0,7	**0,8**	6,6	9,0	1,2	**0,8**	-13,7	**8,5**	0,9	**0,8**	**4,1**	9,1	**1,0**	**0,8**
CAP BEAR	-36,6	12,8	0,5	0,8	-16,4	**7,2**	**0,9**	**0,9**	-27,3	10,1	0,7	**0,9**	**-14,1**	9,1	0,8	0,8
SETE	**6,9**	**8,5**	**1,1**	**0,7**	31,3	12,6	1,6	**0,7**	25,4	10,7	1,4	**0,7**	16,8	10,0	1,3	**0,7**
STES-MARIES-DE-LA-MER	6,0	8,9	0,9	**0,8**	14,6	10,9	1,2	**0,8**	**-4,7**	**7,3**	0,9	**0,8**	5,6	9,4	**1,0**	**0,8**
PORQUEROLLES	-27,8	9,8	0,8	**0,9**	-6,5	7,5	1,0	**0,9**	**1,0**	**7,3**	**1,0**	**0,9**	-7,9	7,9	0,9	**0,9**
CAP FERRAT	-45,0	10,5	0,5	0,6	65,2	19,8	1,5	0,1	**7,0**	**8,8**	**1,1**	**0,7**	-26,9	9,0	0,5	0,6
Seafront mean	**26,5**	**10,4**	**0,7**	**0,8**	**23,4**	**11,2**	**1,2**	**0,7**	**13,2**	**8,8**	**1,0**	**0,8**	**12,6**	**9,1**	**0,9**	**0,8**
Seafront best score	**1/6**	**1/6**	**1/6**	**4/6**	**0/6**	**1/6**	**1/6**	**5/6**	**3/6**	**4/6**	**2/6**	**6/6**	**2/6**	**0/6**	**2/6**	**4/6**

We can see that:

- in the English Channel, CERRA performs better than the other reanalyses, with an average normalized error (NRMSE) of less than 6%; note that COSMO-REA6 obtains the same scores as CERRA in terms of correlation and STD ratio.
- in the Atlantic Ocean, the scores are close, with normalized errors of the order of 5-6%; COSMO-REA6 is the best reanalysis.
- in the Mediterranean Sea, the differences are greater, and CERRA is again the best reanalysis, with an average normalized error of less than 10%.

Overall, in view of these scores, CERRA is therefore the best performing reanalysis on the scale of continental France.

1.2 Seasonal and Diurnal Cycles

To better evaluate the reanalyses, we then focus on the seasonal and diurnal cycles, which are important for wind farm planning and especially important when we want to study and improve climate models, the aim of which is to reproduce these cycles well.

The seasonal cycles, represented by monthly means for different French coasts (mean over the stations) are shown in Figure 9. The comparisons for individual stations are shown in the Appendix (section VI).

In comparison to SYNOP measurements, the shape of the seasonal cycle (left column) is well represented by each of the reanalyses, although biases exist. For the English Chanel, the seasonal cycles for the observations and the 4 reanalyses (Figure 9, top panel) show lower values in summer (June, July and August) around 5 m/s and higher values during winter (December, January and February) with average values above 7 m/s. We can see that CERRA is the reanalysis most in agreement with the observations in summer but presents too high values in winter. ERA5 and COSMO-REA6 are in agreement in winter but present lower values in summer. Overall, the biases in the reanalyses are small (within 0.5 m/s).

For the Atlantic coast (middle panel), the seasonal cycles of the reanalyses are very similar to each other but are lower than the measurements by around 0.3 m/s, especially in summer.

For the Mediterranean coast, larger biases are observed owing to the complexity of processes/ wind regimes. COSMO-REA6 has low biases during summer months, but high biases during winter months. The biases in CERRA and ERA5 corrected by COSMO-REA6 are below -0.5 m/s throughout the year. The wind speeds from CERRA are mostly higher than from ERA5, in agreement with Jourdier et al. (2023), who showed large differences, especially over mountains (where ERA5 tends to underestimate the wind speeds). Olsen et al. (2019) reported that ERA5 under-estimates the wind speed (-1.50 ± 1.30 m/s) and that the bias increases with the orographic complexity using wind-speed measurements from Vestas for almost 300 masts spread throughout Europe (including 30 in France).

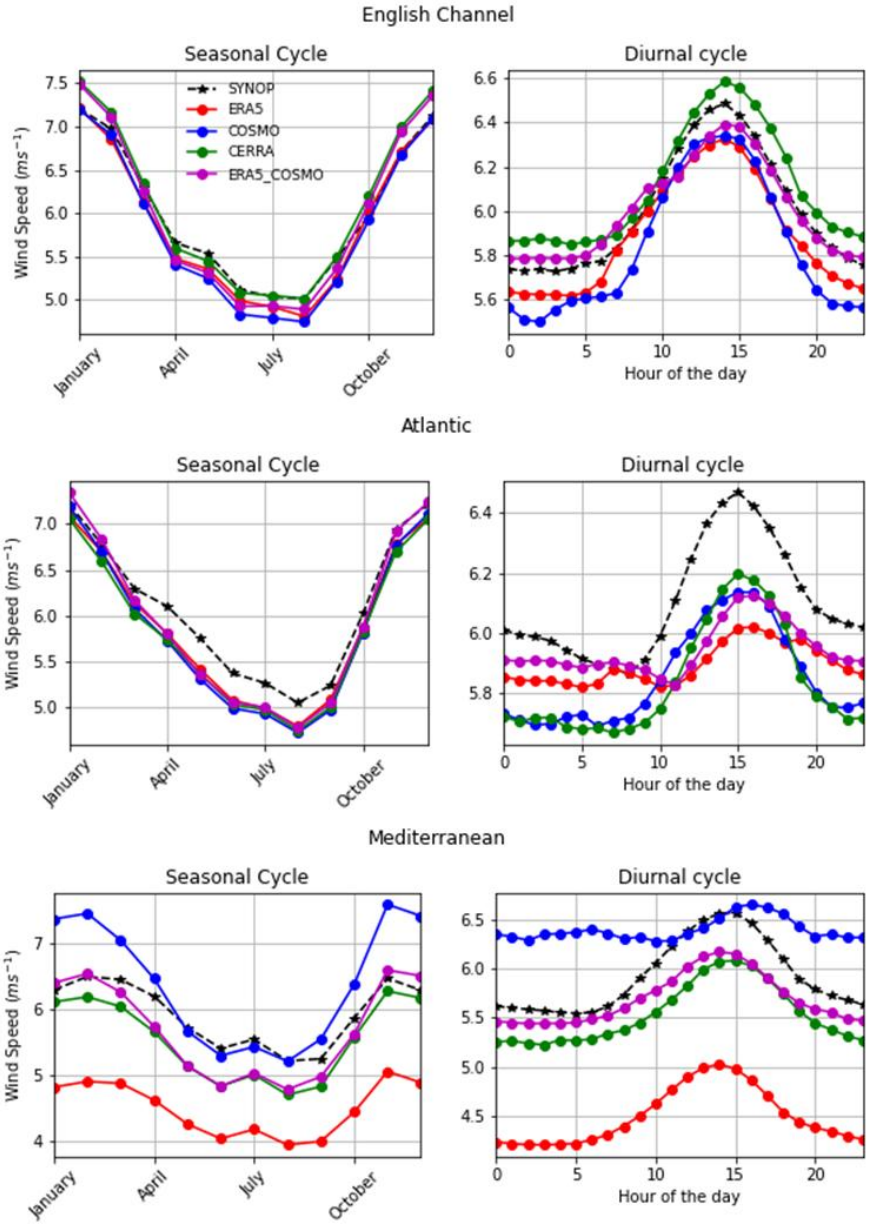


Figure 9. Comparison of Seasonal cycle and Diurnal cycle (mean over the stations) for different French coasts

The diurnal cycle biases are in a similar range as the seasonal cycle biases (Figure 9, right column). For the English Channel coast, both CERRA and ERA5 corrected by COSMO-REA6 are similar to the observations, but the shape of the cycle is better represented in CERRA. Although COSMO-REA6 and ERA5 show small underestimations, the shape of the diurnal cycle is represented relatively well.

Along the Atlantic coast, the shape of diurnal cycle in CERRA and COSMO-REA6 agrees well with the measurements, with a bias of -0.2 m/s. The diurnal cycles in ERA5 and ERA5 corrected by COSMO-REA6 are relatively flatter for the station mean.

For the Mediterranean coast, the diurnal cycle for CERRA and ERA5 corrected by COSMO-REA6 represents well the shape of the observed diurnal cycle (bias ~ -0.5 m/s). The form of the diurnal cycle in COSMO-REA6 is worse (low correlation in Figure 11), and ERA shows a larger (around -1.5m/s) underestimation. Jourdier (2020) emphasized that ERA5 shows very good skill in reproducing the diurnal cycle, even in complex environments: there is a bias, but the shape of the diurnal cycle is correct. The issue in diurnal cycle in COSMO-REA6 can be linked to difficulties with stable conditions,

with nocturnal low-level jets and with vertical mixing after sunrise, originating from the boundary layer turbulence scheme in the COSMO-REA6 model (Heppelmann et al., 2017).

We also present error metrics for seasonal and diurnal cycles at each station using box-whisker plots. The PBIAS in seasonal and diurnal cycles is similar to the PBIAS using hourly data (Figure 5), thus is not discussed here. The shape of the seasonal cycle in the reanalyses (as seen in Figure 9) can be better evaluated by its correlation with the observed seasonal cycle. Figure 10 displays that the correlation is always high (> 0.9) for the reanalysis seasonal cycles, except for a station in the Mediterranean (also see Figure A5). Nevertheless, reanalyses always show a decent correlation (> 0.70) with measurement (Potisomporn et al., 2023). Overall, considering the correlation, all reanalyses perform well for the English Channel and the Atlantic coast. Along the Mediterranean coast, CERRA has better correlation values for the seasonal cycle.

The diurnal cycle correlations (Figure 11) are not always as high as the seasonal cycle correlations (Figure 10). For the English Channel coast, the four reanalyses have high correlations (1st quartile > 0.9) with the measurements for the diurnal cycle, while in the Mediterranean, all but COSMO-REA6 (also clear from Figure 9) have high correlation values. Along the Atlantic coast, some low and even negative correlations are observed, but mostly for island locations (also see Figure A6). However, the median values are close to 0.9 for CERRA and COSMO-REA6 there. Moreover, the first and third quartiles of the correlations are 0.59 to 0.92 for COSMO-REA6, and 0.44 to 0.95 for CERRA. The third quartiles are below 0.85 for ERA5 and ERA5 corrected by COSMO-REA6.

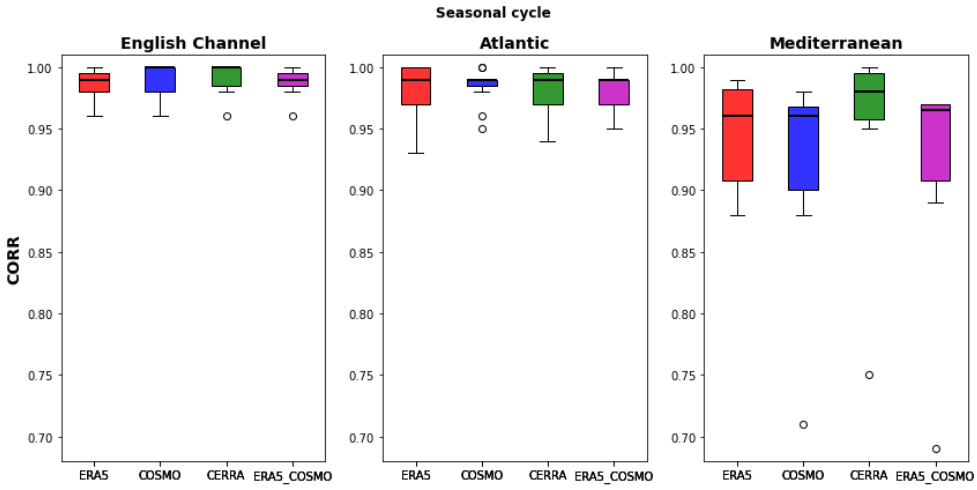


Figure 10. Box–Whisker plots for the distribution of Correlation Coefficients in Seasonal Cycle over all stations for 10 m wind speed from the reanalyses with SYNOP

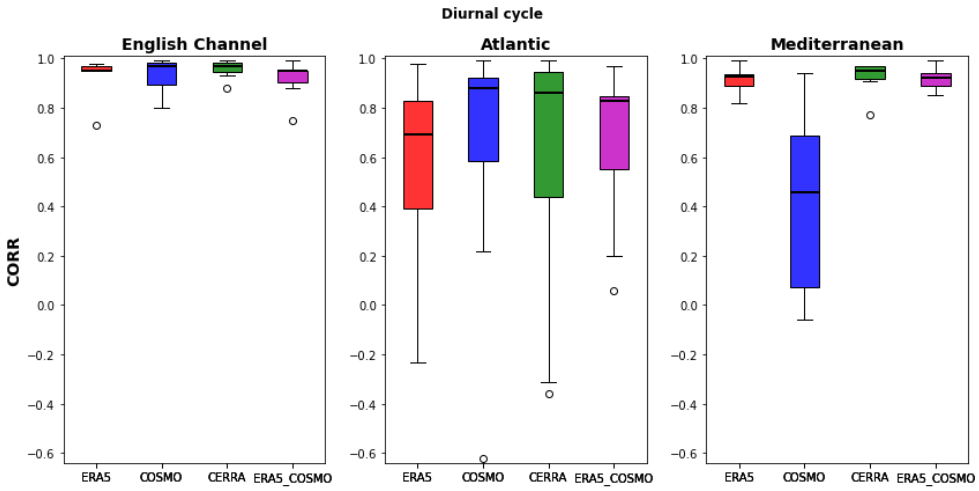


Figure 11. Box–Whisker plots for the distribution of Correlation Coefficients in Diurnal Cycle over all stations for 10 m wind speed from the reanalyses with SYNOP

1.3 Inter-annual Variability and Distribution

The reanalyses can capture the interannual variation (Figure 12), in agreement with Spangehl et al. (2023), who reported that ERA5, COSMO-REA6, and CERRA reproduce well the interannual variations in wind speed. It is important to note here that the measurements are not always temporally homogenous at the stations, creating gaps in the annual mean timeseries for a few stations (Figure 4). However, we compute the annual mean only using values at overlapping data periods between the measurements and reanalyses to ensure that the comparisons are not biased.

For the period around 2000, there are biases (~0.4 m/s) in the Atlantic coast station mean (primarily from 3 stations). During that period, there were much higher wind speeds in northern France, related to a very positive phase of the North Atlantic Oscillation (Jourdier, 2020). For the Mediterranean coastline, the underestimation by ERA5 and overestimation by COSMO-REA6 are also evident (Figure 12); but overall, a good correlation (>0.75) is found between the reanalyses and measurements for the annual mean time series The reanalyses show negative trends for the English Channel (around -0.08 m/s/ decade) and Atlantic Coast (around -0.12 m/s/decade). Such a trend is not clear along the Mediterranean coast where the trend seems on the contrary to be rather positive (Figure A7 and Figure A8). The negative trends in station-mean SYNOP measurements (Figure A8) are almost double those of the reanalyses, although caution should be taken for these trends as there are many missing values (Table A1).

The cumulative distribution function (CDF) shows that the distribution of wind speed in all the reanalyses resembles the observed distribution for the full range of wind speeds in the English Channel and along the Atlantic Coast (Figure 12). Similar observations have been reported for CERRA, COSMO-REA6 and ERA5 in the German Exclusive Economic Zone (EEZ) of the North Sea (Spangehl et al., 2023). However, for the Mediterranean coast, CERRA and ERA5 corrected by COSMO-REA6 are closer to the observed CDF, even if the CDF function shows a slight underestimation in CERRA for wind speeds higher than 6 m/s. The ERA5 distribution is shifted towards low values for all wind regimes. The distribution obtained from COSMO-REA6 shows a slight shift towards higher values for high wind regimes, which also agrees with Spangehl et al. (2023).

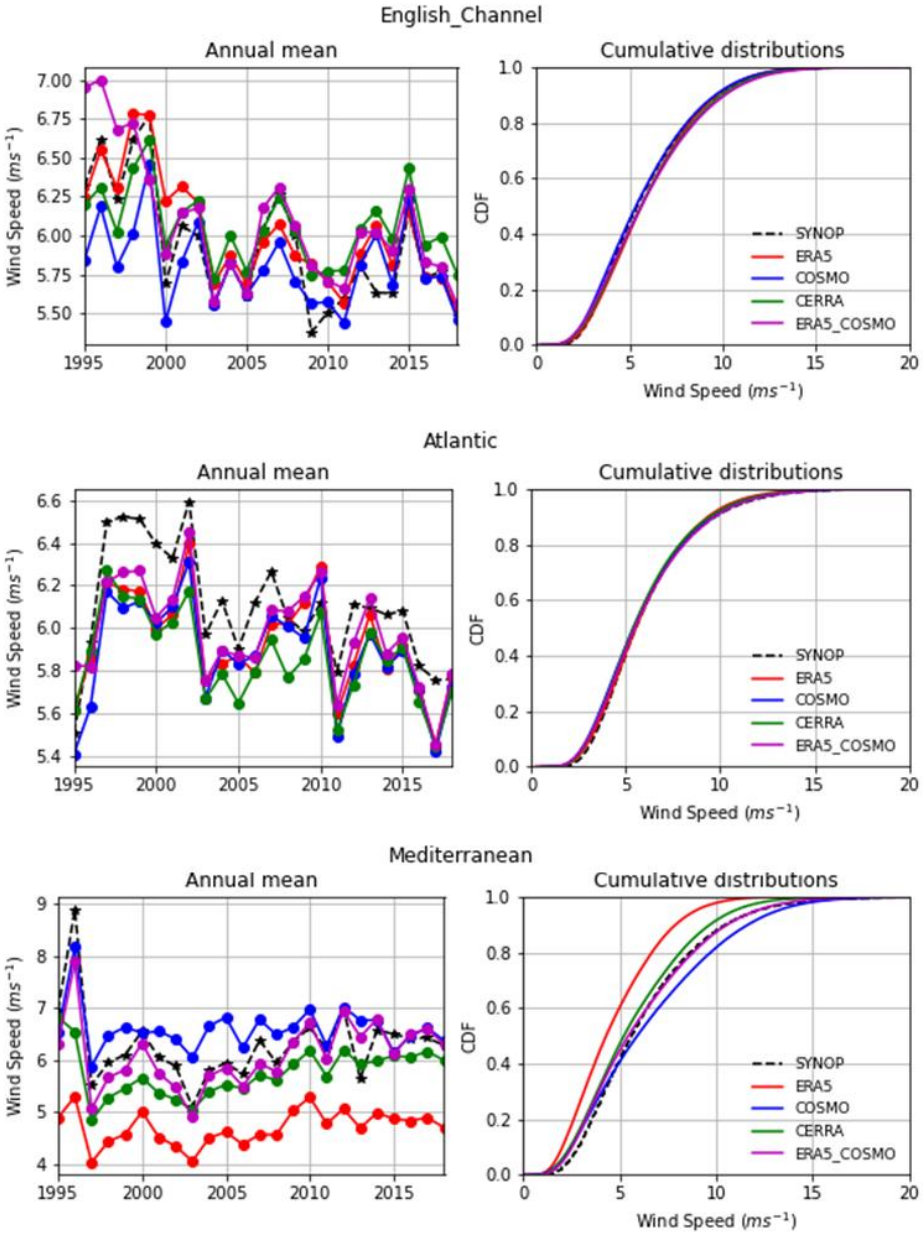


Figure 12. Comparison of annual mean and Cumulative distribution function (mean over the stations) for different French coasts

1.4 Low and high wind speeds

In this section, low and values of 10 m wind speeds (In the following, we will call these values "extremes", for simplicity and abuse of language) are evaluated, focusing on the 95th and 5th percentiles over the period 1995-2018. A comparison between the reanalyses and SYNOP measurements is made at 24 coastal locations (Figure 13). Here we choose percentiles from 10 m winds for extremes instead of winds beyond cut-in and cut-off limits. Indeed, the cut-in and cut-out wind speeds refer to the wind speed at the hub height of the wind turbine, which depends on wind turbine characteristics, and varies with the advancement in wind turbine technology.

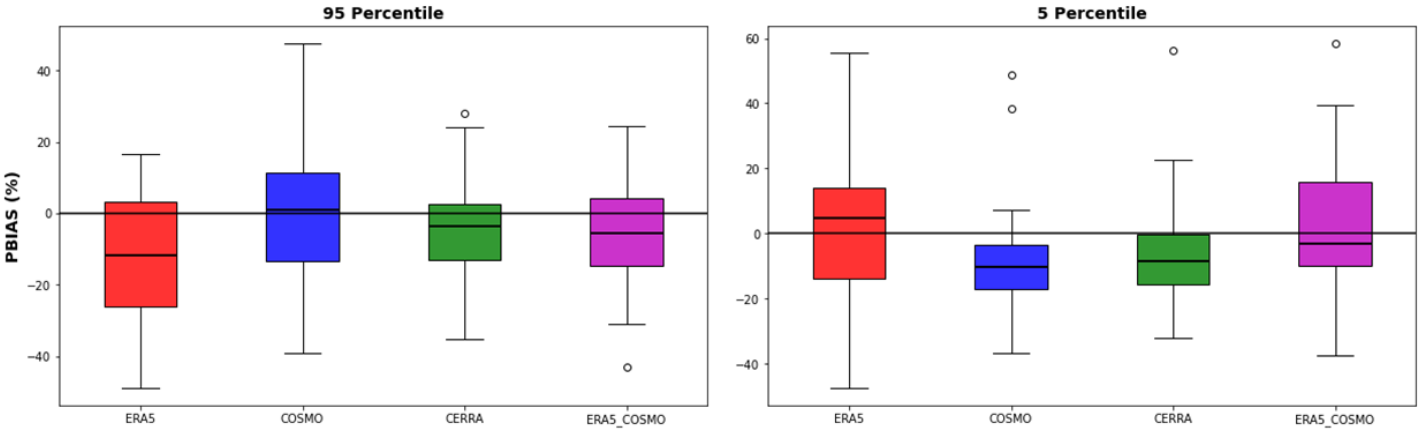


Figure 13. Box–Whisker plots for the distribution of bias values (%) for all 24 stations for higher (95th Percentile) and lower (5th Percentile) extremes

For the 95th percentile, ERA5 shows a large underestimation (Figure A9), and COSMO-REA6 has median value (of biases) close to 0. The ERA5 corrected by COSMO-REA6 has smaller biases than ERA5 itself. The first and third quartiles ranges are (-12.9% to 2.6%) for CERRA, (-13.3% to 11.4%) for COSMO-REA6 and (-14.7% to 4.1%) for ERA5 corrected by COSMO-REA6. The median values (Figure 13) for ERA5, COSMO-REA6, CERRA and ERA5 corrected by COSMO-REA6 are -11.6%, 0.9%, -3.5%, -5.4%, respectively.

For the 5th percentile, a systematic shift towards negative bias values is found for CERRA (median bias: -8.6%) and COSMO-REA6 (median bias: -10.2%), and the first and third quartiles ranges are -15.7% to -0.2% and -17.0% to -3.3%, respectively (Figure 13). A larger range of biases is observed for ERA5 and ERA5 corrected by COSMO-REA6 (Figure A10), although the median values are low (4.8% and -3.0%).

Figure 14 compares the annual percentile (95th and 5th) values for the reanalyses with those of the measurements by considering only the common time periods. One should note that there are gaps in the measured data. Overall, ERA5 underestimates high percentiles along all coasts, and overestimates the low percentiles (except in the Mediterranean), although by a smaller margin. For the English Channel and Atlantic coast, CERRA, COSMO-REA6 and ERA5 corrected by COMSO-REA6 are very close and have similar interannual variations for both high and low percentiles. In the Mediterranean, both the 5th and 95th percentile wind speeds are better represented by CERRA and ERA5 corrected by COSMO-REA6. For the correlations, higher values are observed for the higher percentiles, and lower values are observed for the lower percentiles.

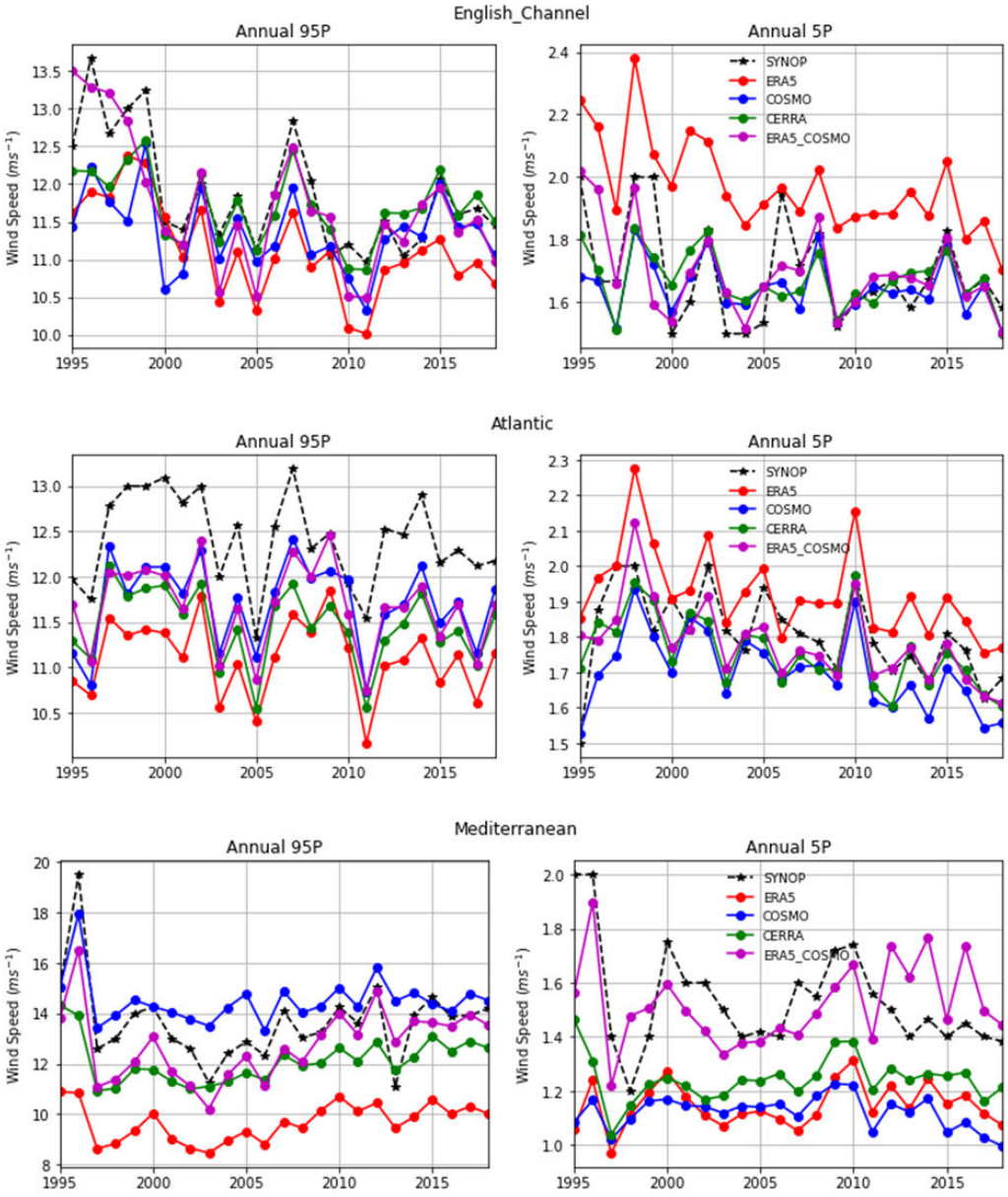


Figure 14. Interannual variability of the 95th percentile (left column) and 5th percentile (right column) wind speed (mean over the stations) for the different French coasts

Considering the trends, all of the reanalyses display negative trends in the English Channel and along the Atlantic coast, around - 0.1 m/s/decade for 95th Percentiles (Figure A11) and bit lower negative trend for 5th percentiles (Figure A12). Tian et al. (2019) reported the average trend of high winds (-0.26 m/s/decade) is more than twice as large as that of the mean speed winds (-0.09 m/s/decade) for Europe during 1979-2016 using observations from the Integrated Surface Database (ISD). The SYNOP measurements show slightly higher trends (Figure A8) when averaged over the selected stations (-0.42 m/s/decade and -0.24 m/s/decade for 95th percentile and mean, respectively). Measurement trends should be considered with caution because of missing values. The reanalyses show smaller trends than the observations.

2 Evaluation of 100 m Wind Speed with LIDAR

Reanalyses are evaluated based on the 10 m wind speed because observations at that height are available over an extended period. However, it is also necessary to evaluate the 100 m offshore wind speed in the reanalyses because reanalyses serve as the reference to evaluate and correct wind speeds from climate models (CMIP6: Coupled Model Intercomparison Project Phase 6 models) in the 2CNOW project as climate models are the only data source that can be used to evaluate future changes in wind speeds for offshore wind projects. The measurements for 100m wind speed data are from floating LIDAR and are available over shorter periods (12 to 18 months) than SYNOP 10m coastal stations.

From the data availability (Figure 15), we can see that all 4 reanalyses can be evaluated at stations like Groix and Oleron (Atlantic coast). For the other LIDAR stations, we are only confined to ERA5 and ERA5 corrected by COSMO-REA6, because these two reanalyses are the only ones presenting data after 2021.

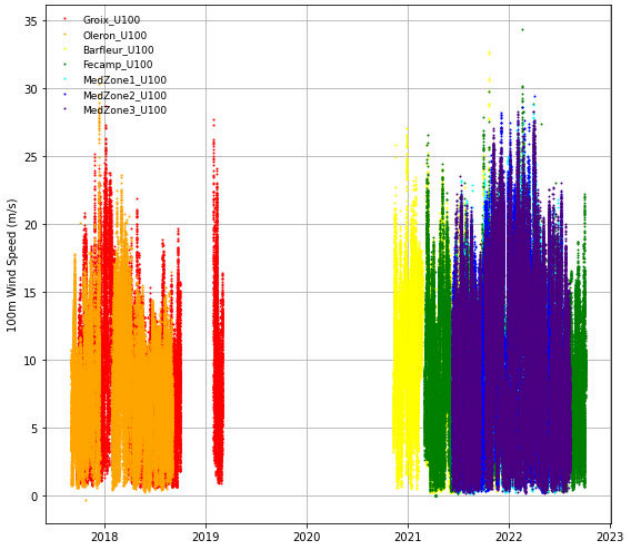


Figure 15. Time-series of LIDAR measurements at AO locations

We compute the error metrics during the common time period between the LIDAR measurements and the reanalyses. Overall, the reanalyses show low PBIAS (<10%) and NRMSE (<10%), high correlations (>0.9), and normalized standard deviations close to 1 for 100 m wind speed at all the stations (Table 5). Despite the relatively coarse ERA5 data resolution, good performances are noted when comparing to LIDAR measurements, as mentioned in Cañadillas et al. (2023) who reported that the ERA5 data shows high correlations with wind speed measurements from LIDAR located on Norderney Island near the German mainland. In Groix and Oléron, where all the reanalyses can be compared, we see once again that CERRA is the reanalysis which obtains the best scores overall.

Table 5. Summary table of error metrics at AO locations, for 100 m wind speed

Scores at AO locations (100 m)	ERA5				COSMO				CERRA				ERA5 corrected by COSMO			
	PBIAS	NRMSE	STD ratio	CORR	PBIAS	NRMSE	STD ratio	CORR	PBIAS	NRMSE	STD ratio	CORR	PBIAS	NRMSE	STD ratio	CORR
Station	[%]	[%]	[]	[]	[%]	[%]	[]	[]	[%]	[%]	[]	[]	[%]	[%]	[]	[]
BARFLEUR	**-3,2**	**5,2**	**1,0**	**0,9**									-4,0	6,2	**1,0**	**0,9**
FECAMP	**-1,9**	**5,5**	**1,0**	**0,9**									-8,7	7,4	**1,0**	**0,9**
MED Zone 1	-13,1	9,5	0,9	**0,9**									**-7,0**	**9,4**	**1,0**	**0,9**
MED Zone 2	-8,9	**8,0**	0,9	**0,9**									**-6,5**	8,3	**1,0**	**0,9**
MED Zone 3	-7,9	**8,6**	0,9	**0,9**									**-6,0**	9,1	**1,0**	**0,9**
GROIX	-3,3	**5,8**	**1,0**	**0,9**	-1,9	6,4	**1,0**	**0,9**	**0,4**	6,0	**1,0**	**0,9**	-7,4	6,6	**1,0**	**0,9**
OLERON	-5,7	6,5	0,9	**0,9**	-6,2	6,5	**1,0**	**0,9**	**-1,4**	**6,0**	**1,0**	**0,9**	-9,4	7,6	**1,0**	**0,9**

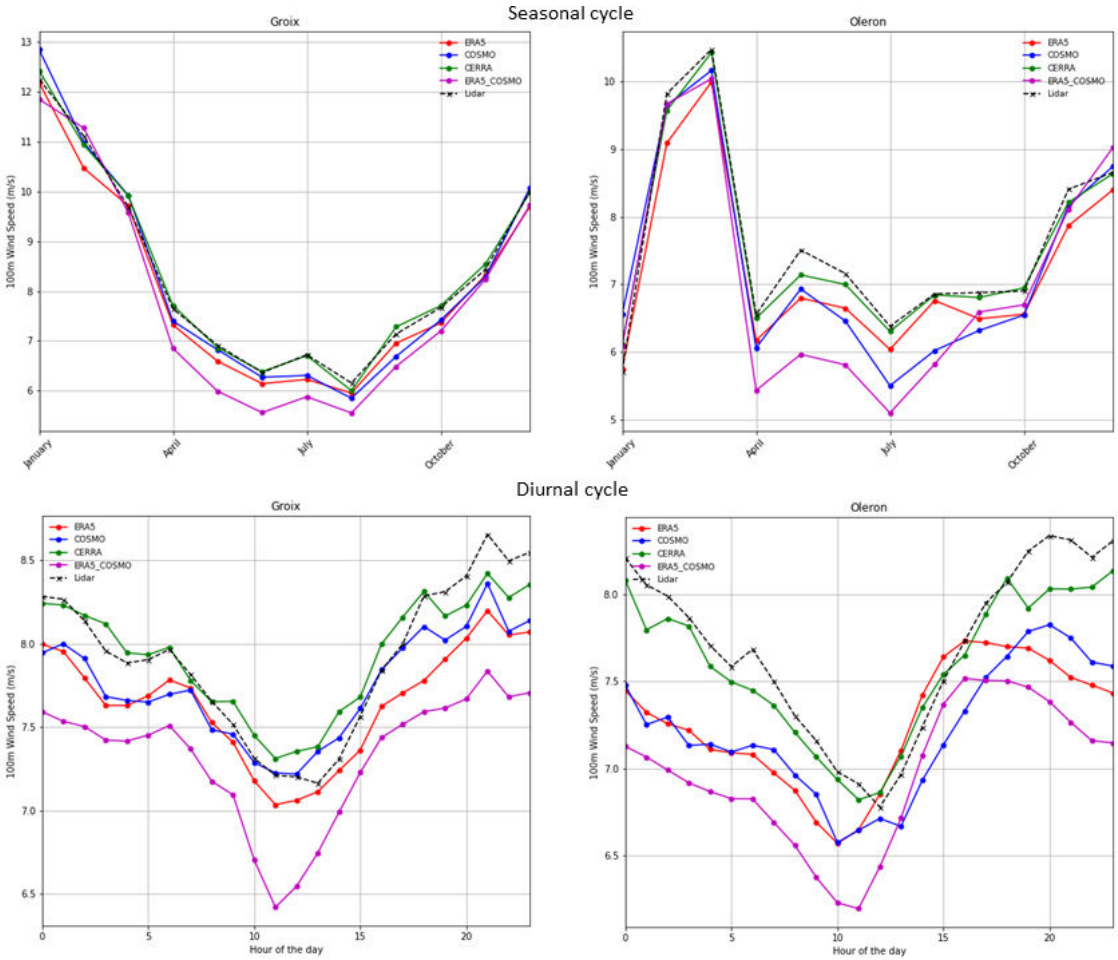


Figure 16. Comparison of seasonal and diurnal cycle between reanalyses and LIDAR measurements at GROIX and OLERON (AO locations)

These results are confirmed by diurnal and seasonal cycles (Figure 16). The shape of the seasonal cycle from reanalyses agrees well with the LIDAR observations. The biases are low (< 1- 1.5m/s) for all reanalyses, and CERRA has the smallest bias (<0.1 m/s). Similar negative biases (~ 1 m/s) are reported for ERA5 in comparison to LIDAR by Pronk et al. (2022). The regional reanalyses COSMO-REA6 and CERRA show small biases and the global reanalysis ERA5 slightly underestimates the wind speed measurements near 100 m at the FINO1 research platform (Spangehl et al., 2023). In addition, biases are smaller during winter months than summer months. Considering the shape of the diurnal cycles (Figure 16), the reanalyses show reasonable agreement with the LIDAR measurements. The negative bias values in diurnal cycles of the reanalyses are within the uncertainty range of the reanalyses. CERRA has the smallest bias, and ERA5 corrected by COSMO-REA6 has the largest bias.

IV. CONCLUSION

In this study, the quality of 10 m and 100 m wind speed estimates from four different reanalyses (ERA5, ERA corrected by COSMO-REA6, COSMO-REA6, and CERRA) is evaluated along various French coastlines. The reference datasets include SYNOP wind measurements at 10 m height (Météo-France, 24 coastal weather stations) and LIDAR measurements at 100 m height (Météo-France, 7 offshore stations). The longer overlapping data period (1995–2018) for 10 m wind speed allows for a thorough investigation of seasonal and diurnal cycles, interannual variability, distribution, and extreme values (low and high percentiles).

The reanalyses do not appear to perform consistently across the three main French coastlines. In particular, along the Mediterranean coast, larger discrepancies compared to the measurements (at 10 m height) are observed for ERA5 (underestimation) and COSMO-REA6 (overestimation). Overall, CERRA exhibits low bias across most stations along the French coastline. The reanalyses show high correlation values (>0.75), although slightly lower values are found along the Mediterranean coast. The normalized standard deviation indicates that CERRA performs well, ERA5 tends to underestimate variability at most locations, and COSMO-REA6 shows high variability in the Mediterranean region.

The seasonal cycle of 10 m wind speeds is generally well represented by all reanalyses, but ERA5 tends to underestimate and COSMO-REA6 to overestimate wind speeds during winter along the Mediterranean coast. Regarding the diurnal cycle, COSMO-REA6 does not accurately reproduce the observed shape along the Mediterranean coast, and neither ERA5 nor ERA5 corrected by COSMO-REA6 performs well along the Atlantic coast. CERRA, however, successfully captures the shape of the observed diurnal cycle of 10 m wind speed across all coasts.

Beyond a realistic average, the reanalyses must also reflect realistic variability. Diurnal variability of the wind resource is crucial for estimating the amount of energy produced during peak and off-peak hours, which has significant implications for balancing electricity generation and demand (Jourdier, 2020). The reanalyses are able to reproduce the interannual variations observed in the measurements. Low percentiles are underestimated by COSMO-REA6 and CERRA, while high percentiles are underestimated by ERA5.

The 100 m wind speed estimates from ERA5, COSMO-REA6, CERRA, and corrected ERA5 can reproduce the seasonal and diurnal cycles observed in the LIDAR measurements, although the analysis is limited by the short duration of available LIDAR data. Among these, CERRA shows the smallest bias. The best-performing reanalysis for each coast and each metric is summarized in Table 4, based on both the median and interquartile range.

CERRA performs well across many aspects (Table 6), and thus appears to be the most suitable reanalysis to serve as a reference dataset in France for future work aimed at evaluating and downscaling wind speed distributions from several Coupled Model Intercomparison Project Phase 6 (CMIP6) models.

Table 6. Best reanalysis for the coast and for different metrics

Metrics	English Channel	Atlantic	Mediterranean
PBIAS	**CERRA**	**CERRA**	corrected ERA5
Correlation	**CERRA**	COSMO	**CERRA**
NSTD	**CERRA**	COSMO	**CERRA**
Seasonal cycle (PBIAS)	**CERRA**	**CERRA**	**CERRA** and corrected ERA5
Seasonal cycle (Correlation)	**ALL**	**ALL**	**CERRA**
Diurnal Cycle (PBIAS)	**CERRA**	**CERRA**	**CERRA** and corrected ERA5
Diurnal cycle (Correlation)	**ALL**	COSMO	**ALL** but COSMO
Distribution	**ALL**	**ALL**	**CERRA** and Corrected ERA5
95th Percentile (PBIAS)	**CERRA**	**CERRA**	**CERRA** and Corrected ERA5
5 th Percentile (PBIAS)	**CERRA**	**CERRA**	**CERRA** and Corrected ERA5

V. BIBLIOGRAPHY

1. Bentamy, A., & Croize-Fillon, D. (2014). Spatial and temporal characteristics of wind and wind power off the coasts of Brittany. *Renewable Energy, 66*, 670–679. https://doi.org/10.1016/j.renene.2014.01.012

2. Bloomfield, H. C., Brayshaw, D. J., Shaffrey, L. C., Coker, P. J., & Thornton, H. E. (2016). Quantifying the increasing sensitivity of power systems to climate variability. *Environmental Research Letters, 11*, 124025. https://doi.org/10.1088/1748-9326/11/12/124025

3. Bollmeyer, C., Keller, J. D., Ohlwein, C., Wahl, S., Crewell, S., Friederichs, P., Hense, A., Keune, J., Kneifel, S., Pscheidt, I., Redl, S., & Steinke, S. (2015). Towards a high-resolution regional re-analysis for the European CORDEX domain. *Quarterly Journal of the Royal Meteorological Society, 141*, 1–15. https://doi.org/10.1002/qj.2486

4. Cañadillas, B., Wang, S., Ahlert, Y., Djath, B., Barekzai, M., Foreman, R., & Lampert, A. (2023). Coastal horizontal wind speed gradients in the North Sea based on observations and ERA5 reanalysis data. *Meteorologische Zeitschrift, 32*(3), 207–228. https://doi.org/10.1127/metz/2022/1166

5. Cannon, D. J., Brayshaw, D. J., Methven, J., Coker, P. J., & Lenaghan, D. (2015). Using reanalysis data to quantify extreme wind power generation statistics: A 33-year case study in Great Britain. *Renewable Energy, 75*, 767–778. https://doi.org/10.1016/j.renene.2014.10.024

6. Carvalho, D., Rocha, A., Gómez-Gesteira, M., & Santos, C. S. (2014a). WRF wind simulation and wind energy production estimates forced by different reanalyses: Comparison with observed data for Portugal. *Applied Energy, 117*, 116–126. https://doi.org/10.1016/j.apenergy.2013.12.001

7. Carvalho, D., Rocha, A., Gómez-Gesteira, M., & Santos, C. S. (2014b). Offshore wind energy resource simulation forced by different reanalyses: Comparison with observed data in the Iberian Peninsula. *Applied Energy, 134*, 57–64. https://doi.org/10.1016/j.apenergy.2014.08.018

8. Clarke, E., Doddy, S., Griffin, F., McDermott, J., Monteiro Correia, C., & Sweeney, C. (2021). Which reanalysis dataset should we use for renewable energy analysis in Ireland? *Atmosphere, 12*(5), 624. https://doi.org/10.3390/atmos12050624

9. Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., & Vitart, F. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society, 137*(656), 553-597. https://doi.org/10.1002/qj.828

10. Fan, W., Liu, Y., Chappell, A., Dong, L., Xu, R., Ekström, M., & Fu, T. M. (2020). Evaluation of global reanalysis land surface wind speed trends to support wind energy development using in situ observations. *Journal of Applied Meteorology and Climatology, 60*(1), 33–50. https://doi.org/10.1175/JAMC-D-20-0037.1

11. Gualtieri, G. (2022). Analysing the uncertainties of reanalysis data used for wind resource assessment: A critical review. *Renewable and Sustainable Energy Reviews, 167*, 112741. https://doi.org/10.1016/j.rser.2022.112741

12. Heppelmann, T., Steiner, A., & Vogt, S. (2017). Application of numerical weather prediction in wind power forecasting: Assessment of the diurnal cycle. *Meteorologische Zeitschrift, 26*(4), 319–331. https://doi.org/10.1127/metz/2017/0820

13. Hersbach, H., Bell, B., Berrisford, P., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society, 146*, 1999–2049. https://doi.org/10.1002/qj.3803

14. Jourdier, B. (2020). Evaluation of ERA5, MERRA-2, COSMO-REA6, NEWA, and AROME to simulate wind power production over France. *Advances in Science and Research, 17*, 63–77. https://doi.org/10.5194/asr-17-63-2020

15. Jourdier, B., Diaz, C., & Dubus, L. (2023). Evaluation of CERRA for wind energy applications. *EMS Annual Meeting, 2023*, Bratislava, Slovakia, 4–8 September 2023. https://doi.org/10.5194/ems2023-311

16. Miao, H., Dong, D., Huang, G., Hu, K., & Tian, Q. (2020). Evaluation of northern hemisphere surface wind speed and wind power density in multiple reanalysis datasets. *Energy, 200*, 117382. https://doi.org/10.1016/j.energy.2020.117382

17. Olsen, B. T., Hahmann, A., Žagar, M., Hristov, Y., Mann, J., Kelly, M., & Badger, J. (2019). Mapping the European wind climate: Validation of the New European Wind Atlas. *EMS Annual Meeting, 2019*, 9–13 September 2019, Lyngby, Denmark. https://backend.orbit.dtu.dk/ws/portalfiles/portal/198181819/EMS2019_V2.pdf

18. Potisomporn, P., Adcock, T. A., & Vogel, C. R. (2023). Evaluating ERA5 reanalysis predictions of low wind speed events around the UK. *Energy Reports, 10*, 4781-4790. https://doi.org/10.1016/j.egyr.2023.11.035

19. Pronk, V. N., Bodini, M., Optis, J. K., Lundquist, J. K., Moriarty, P., Draxl, C., Purkayastha, E., & Young, E. (2022). Can reanalyses products outperform mesoscale numerical weather prediction models in modeling the wind resource in simple terrain? *Wind Energy Science, 7*, 487–504. https://doi.org/10.5194/wes-7-487-2022

20. RTE (2020). *Analyses mensuelles*. https://www.rte-france.com/fr/eco2mix/analyses-mensuelles (Last accessed 10 January 2020)

21. RTE (2021). https://assets.rte-france.com/prod/public/2022-02/CP_RTE_Bilan-electrique-2021_1.pdf

22. Schimanke, S., Ridal, M., Le Moigne, P., Berggren, L., Undén, P., Randriamampianina, R., Andrea, U., Bazile, E., Bertelsen, A., Brousseau, P., Dahlgren, P., Edvinsson, L., El Said, A., Glinton, M., Hopsch, S., Isaksson, L., Mladek, R., Olsson, E., Verrelle, A., & Wang, Z. (2021a). CERRA sub-daily regional reanalysis data for Europe on height levels from 1984 to present. *CDS [data set]*. https://doi.org/10.24381/cds.38b394e6

23. Schimanke, S., Ridal, M., Le Moigne, P., Berggren, L., Undén, P., Randriamampianina, R., Andrea, U., Bazile, E., Bertelsen, A., Brousseau, P., Dahlgren, P., Edvinsson, L., El Said, A., Glinton, M., Hopsch, S., Isaksson, L., Mladek, R., Olsson, E., Verrelle, A., & Wang, Z. (2021b). CERRA sub-daily regional reanalysis data for Europe on single levels from 1984 to present. *CDS [data set]*. https://doi.org/10.24381/cds.622a565a

24. Spangehl, T., Borsche, M., Niermann, D., Kaspar, F., Schimanke, S., Brienen, S., Möller, T., & Brast, M. (2023). Intercomparing the quality of recent reanalyses for offshore wind farm planning in Germany's exclusive economic zone of the North Sea. *Advances in Science and Research, 20*, 109–128. https://doi.org/10.5194/asr-20-109-2023

25. Tian, Q., Huang, G., Hu, K., & Niyogi, D. (2019). Observed and global climate model-based changes in wind power potential over the Northern Hemisphere during 1979–2016. *Energy, 167*, 1224–1235. https://doi.org/10.1016/j.energy.2018.11.027

VI. APPENDIX

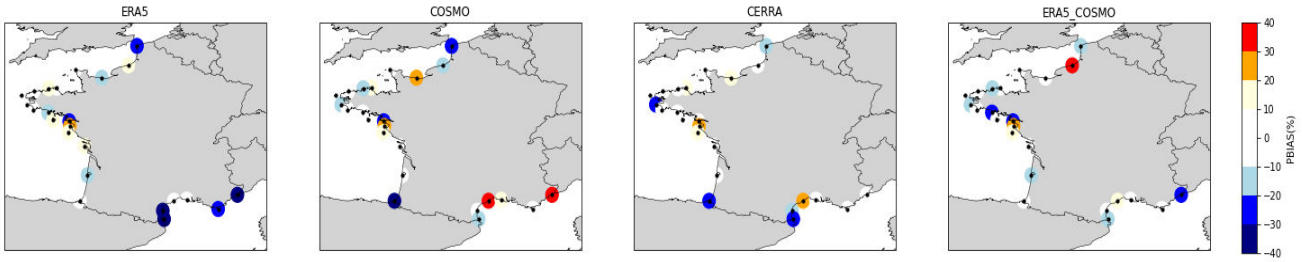


Figure 17. Maps showing percentage biases (PBIAS) for 10 m wind speed from the reanalyses with compared to SYNOP measurements.

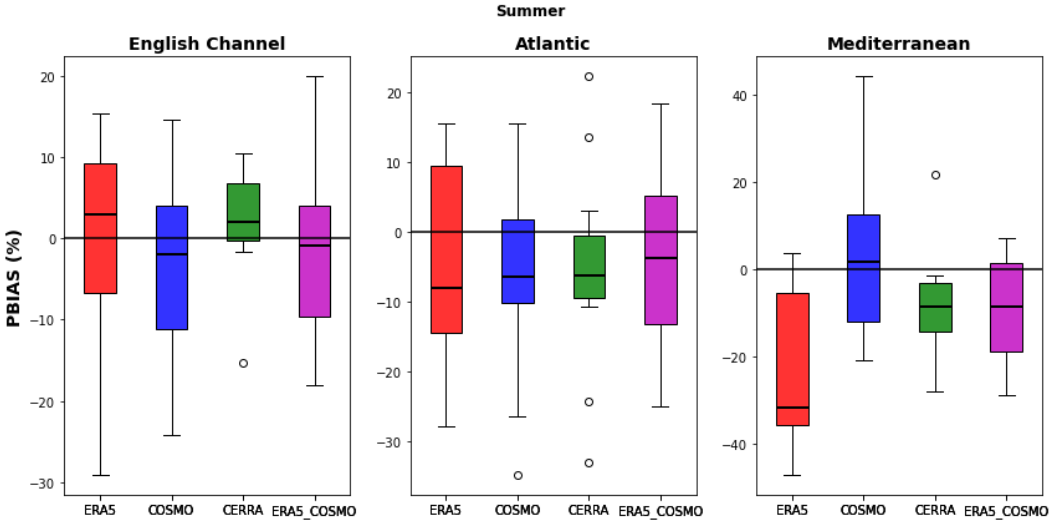


Figure 18. Box–Whisker plots for the bias distribution over the stations for 10 m wind speed from the reanalyses as compared to SYNOP but just including summer months: June-August. The red horizontal line shows median, and the box shows (25-75%) quartiles of the biases for different stations. Larger biases are indicated by the whiskers.

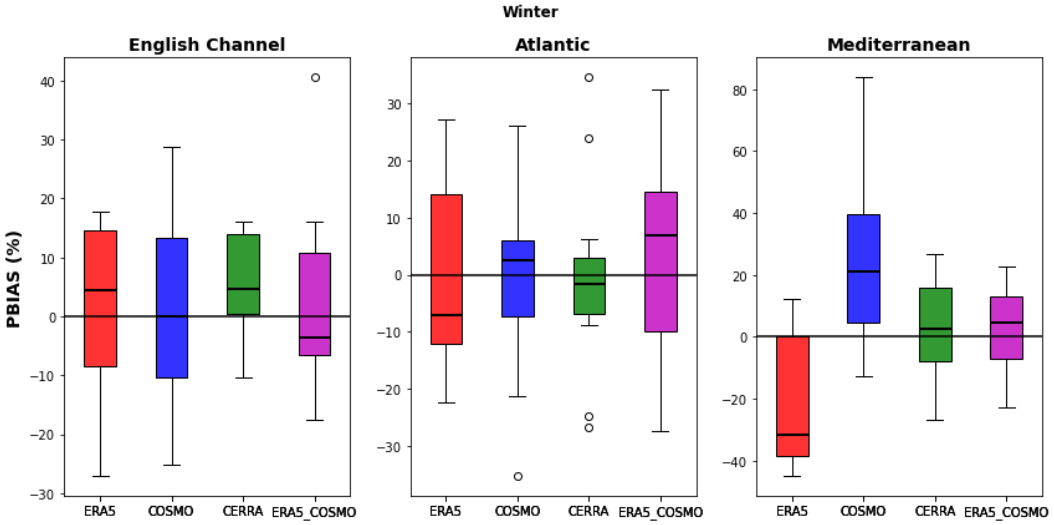


Figure 19. Box–Whisker plots for the bias distribution over the stations for 10 m wind speed from the reanalyses as compared to SYNOP but just including winter months: December-February. The red horizontal line shows median, and the box shows (25-75%) quartiles of the biases for different stations. Larger biases are indicated by the whiskers.

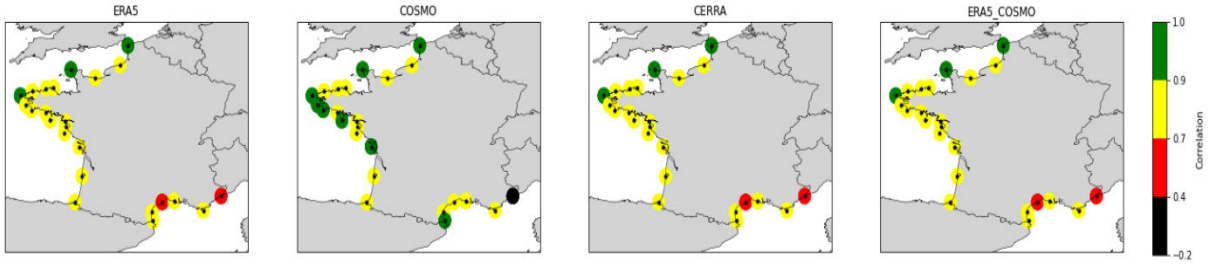


Figure 20. Maps showing Pearson correlation coefficient for 10 m wind speed from the reanalyses and SYNOP measurements. This is using hourly time series at common timesteps between reanalyses and measurement.

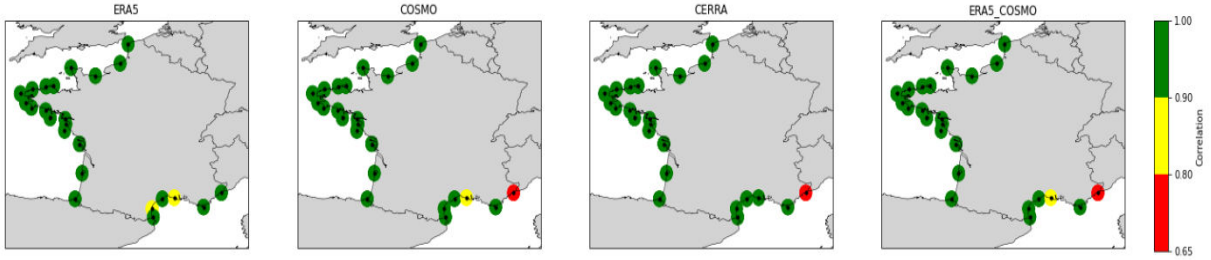


Figure 21. Maps showing Pearson correlation coefficient between seasonal cycles from the reanalyses and SYNOP measurements (10 m wind speed)

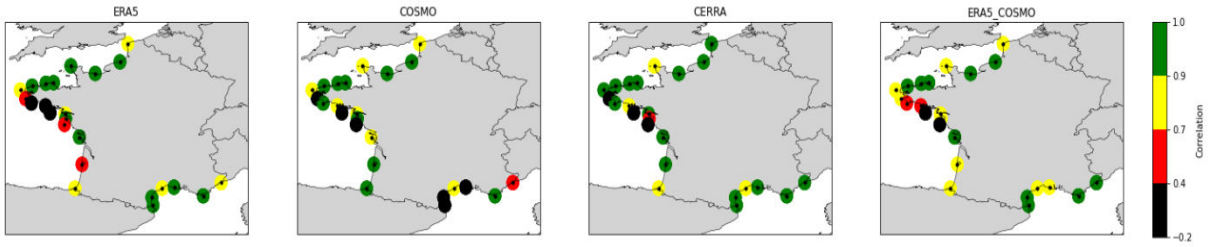


Figure 22. Maps showing Pearson correlation coefficient between diurnal cycles from the reanalyses and SYNOP measurements (10 m wind speed)

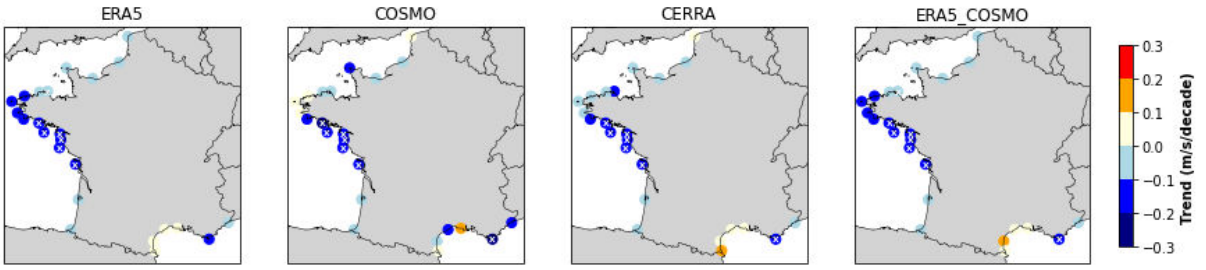


Figure 23. Linear trends (m/s/decade) from annual mean wind speed (10 m height). Significance is shown by "X".

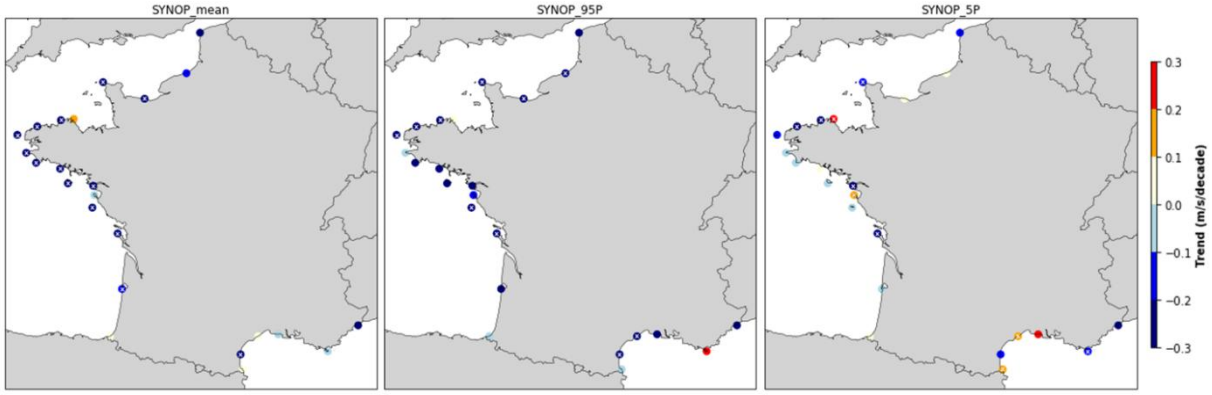


Figure 24. Linear trends (m/s/decade) in wind speed (10 m height) from SYNOP measurements. The trends values should be taken with caution because of missing values. The trends are calculated on annual mean, annual 95th percentiles and annual 5th percentiles. Significance is shown by "X".

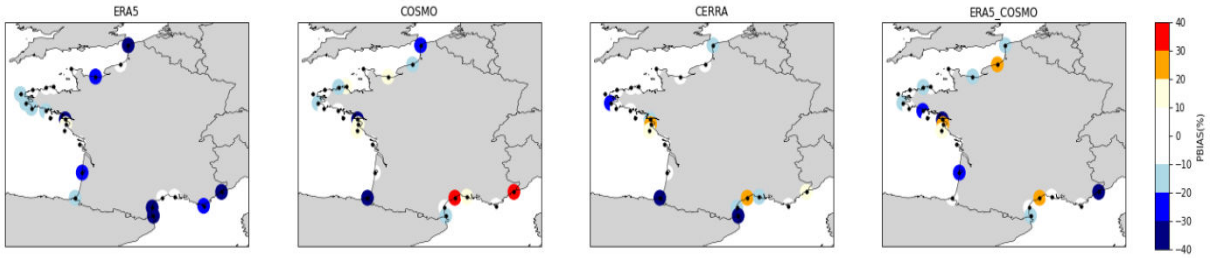


Figure 25. Maps showing percentage biases (PBIAS) in 95th Percentiles for 10 m wind speed from the reanalyses with compared to SYNOP measurements

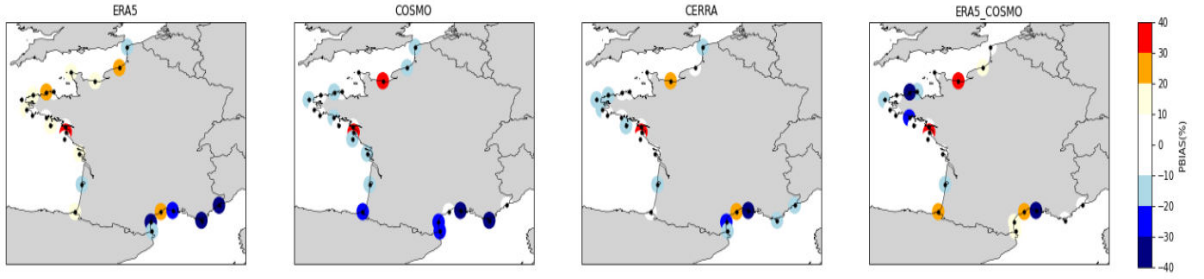


Figure 26. Maps showing percentage biases (PBIAS) in 5th Percentiles for 10 m wind speed from the reanalyses with compared to SYNOP measurements

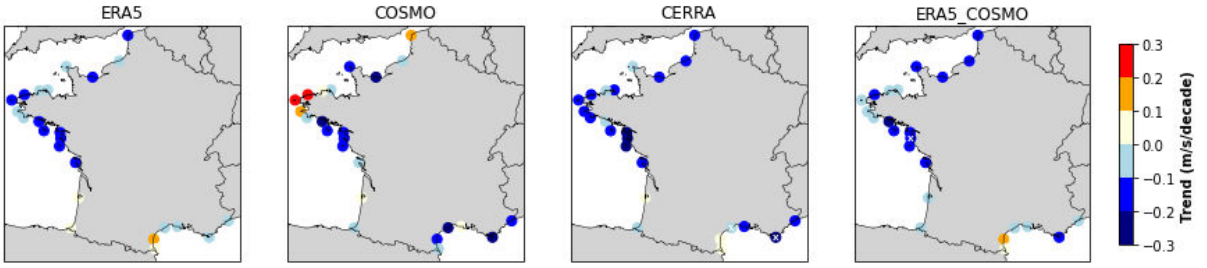


Figure 27. Linear trends (m/s/decade) in annual 95th Percentile wind speed (10 m height). Significance is shown by "X".

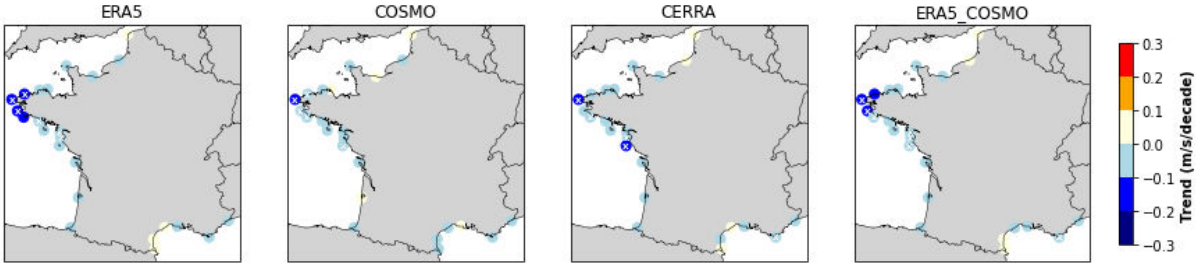


Figure 28. Linear trends (m/s/decade) in annual 5th Percentile wind speed (10 m height). Significance is shown by "X".

Table 7. Percentage of missing SYNOP data (10 m wind speeds) per station for the period 1995-2018

Station	Percentage of missing data (%)
BELLE ILE-LE TALUT	12.67
BERNIERES	29.45
BRIGNOGAN	4.43
CAP BEAR	13.52
CAP FERRAT	54.88
CAP-FERRET	17.28
CAP-GRIS-NEZ	58.66
CHASSIRON	4.95
DIEPPE	5.18
ILE DE GROIX	4.81
ILE-DE-BREHAT	38.41
ILE D'YEU	9.09
LEUCATE	8.84
NOIRMOUTIER EN	0.47
OUESSANT	4.84
PLOUMANAC'H	12.65
PORQUEROLLES	55.91
PTE DE CHEMOULIN	25.46
PTE DE LA HAGUE	17.66
PTE DE PENMARCH	8.83
PTE DU RAZ	8.84
SETE	13.03
SOCOA	9.53
STES-MARIES-DE-LA-MER	38.11

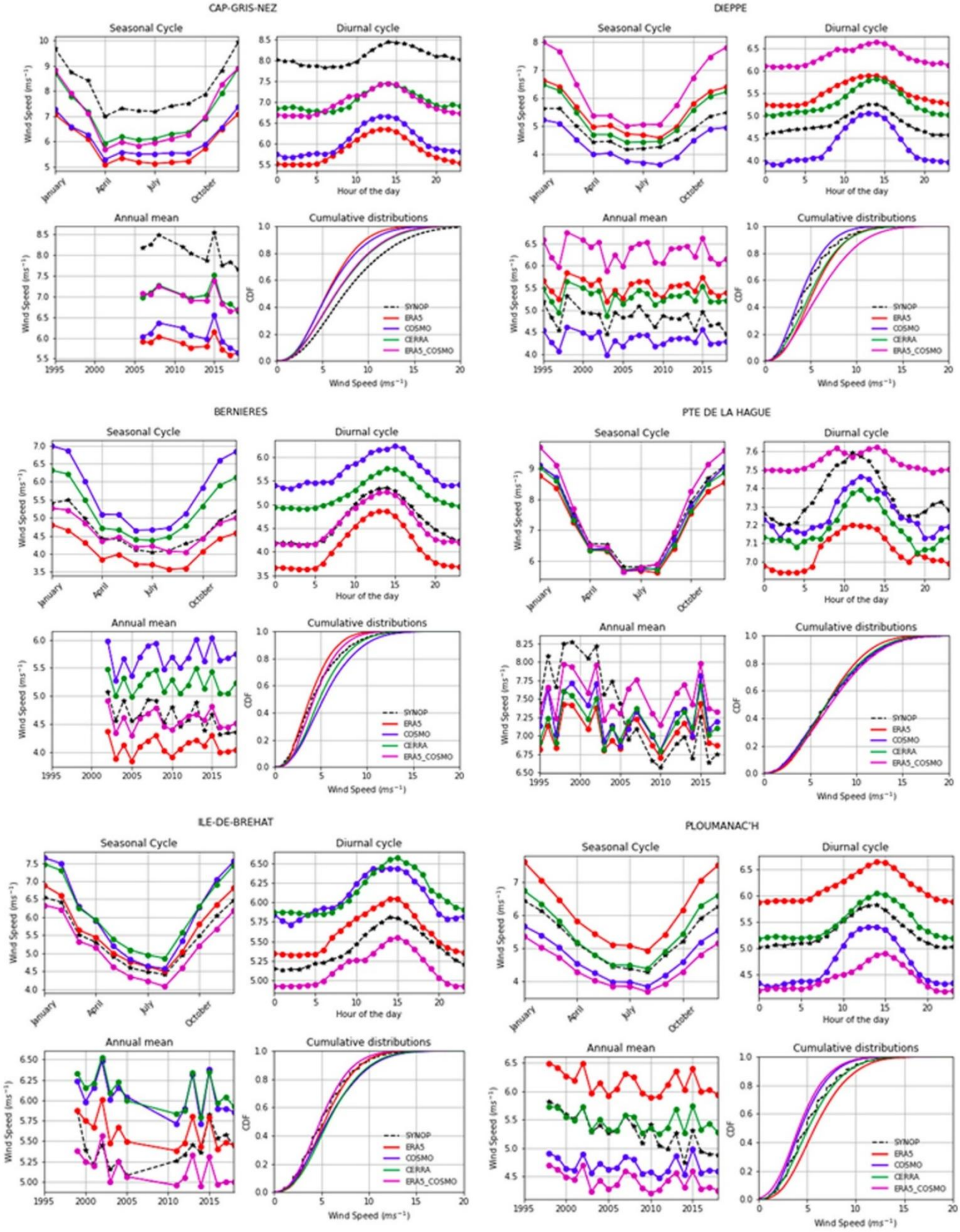
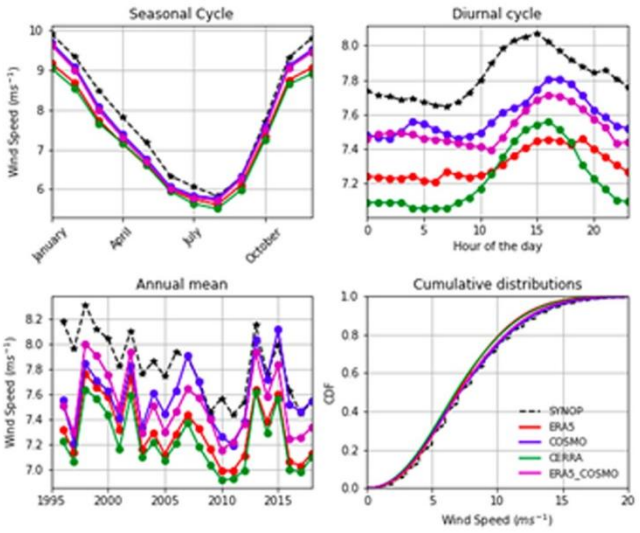
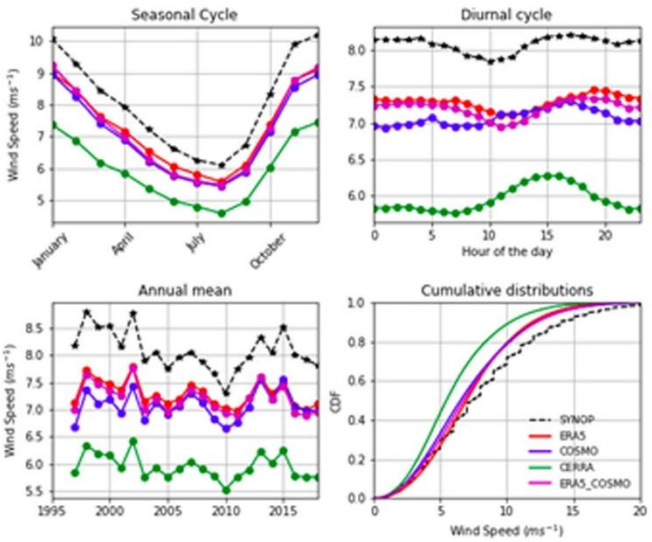


Figure 29. Evaluation of surface wind speed (10 m) at each station: English Channel

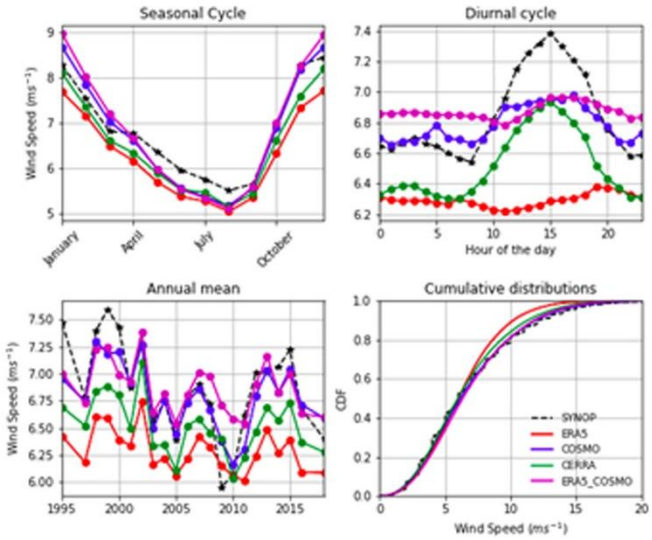
OUESSANT-STIFF



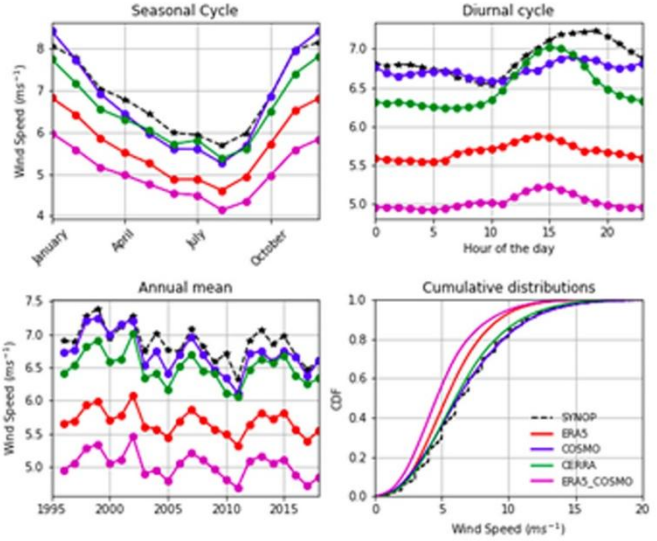
PTE DU RAZ



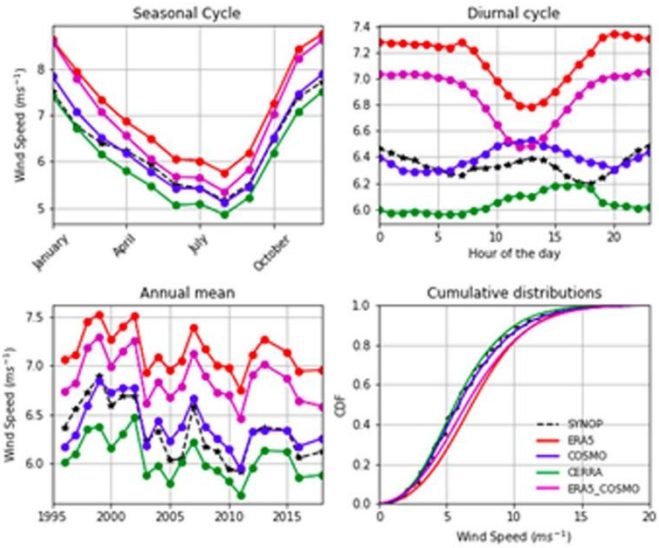
PTE DE PENMARCH



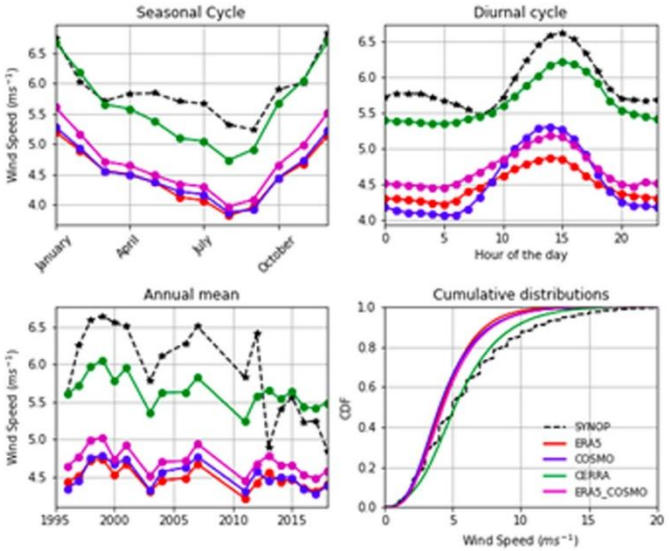
ILE DE GROIX



BELLE ILE-LE TALUT



PTE DE CHEMOULIN



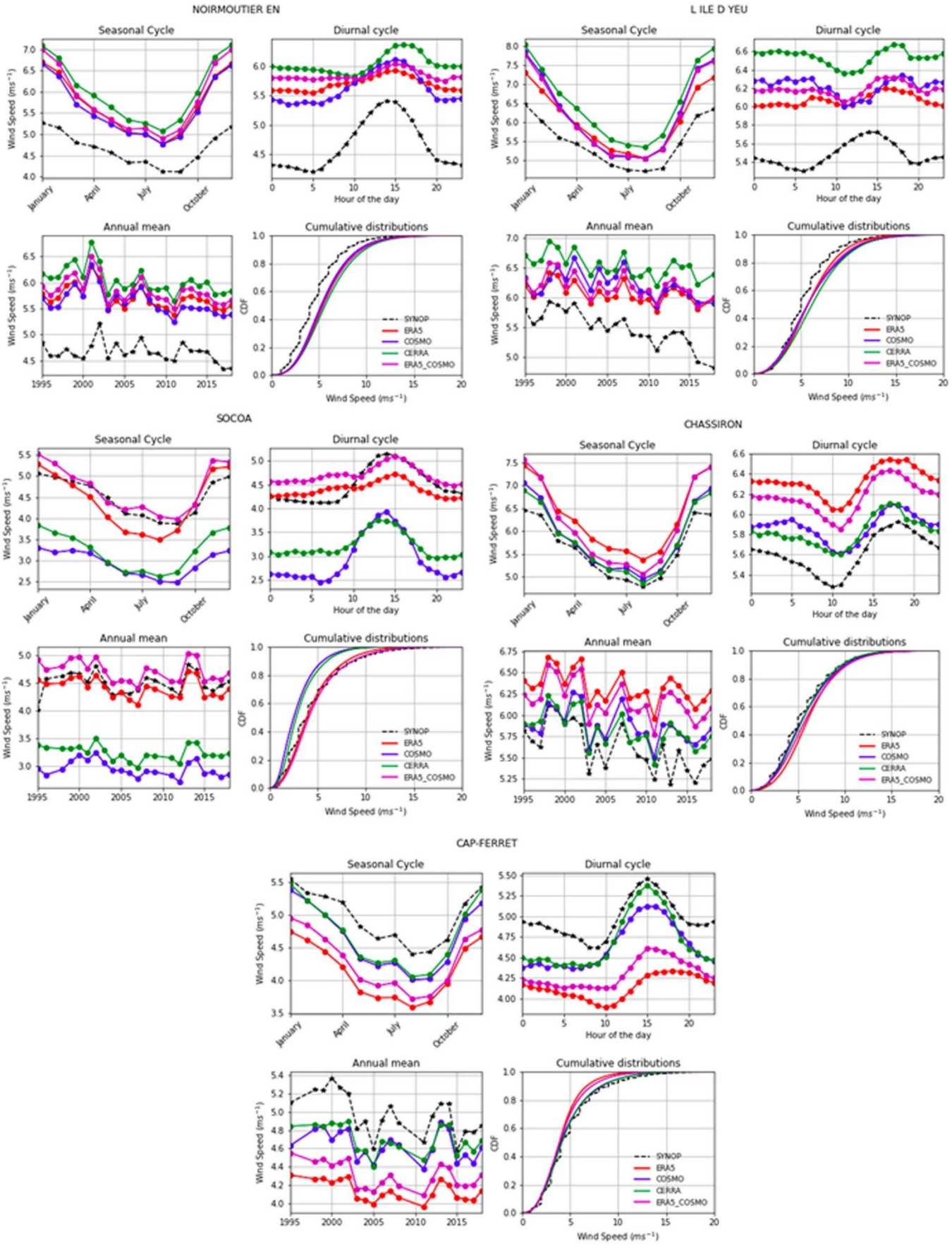


Figure 30. Evaluation of surface wind speed (10 m) at each station: Atlantic

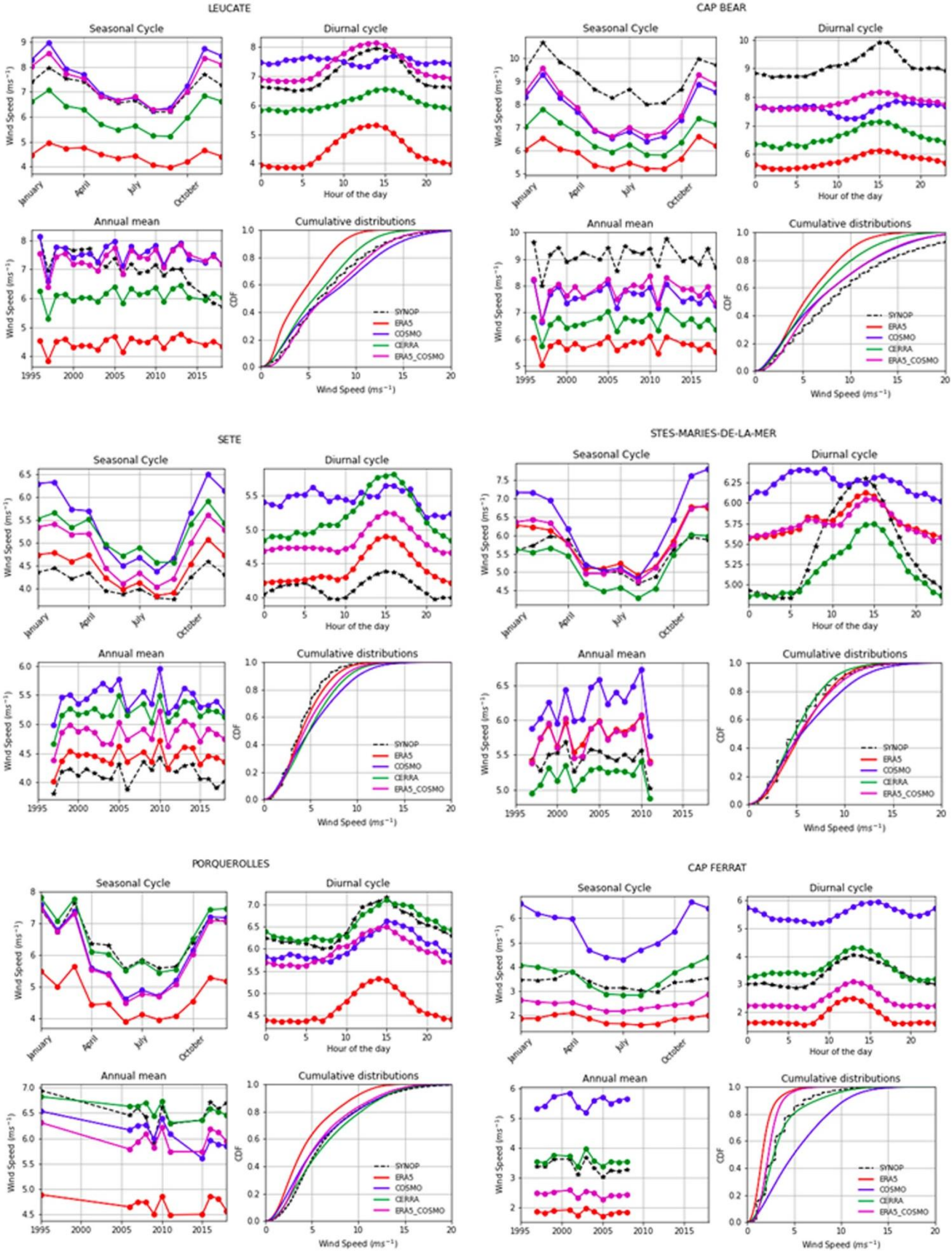


Figure 31. Evaluation of surface wind speed (10 m) at each station: Mediterranean