Global Atlas for Siting Parameters Ocean and Coasts - Ensemble calculation of 50-year return winds

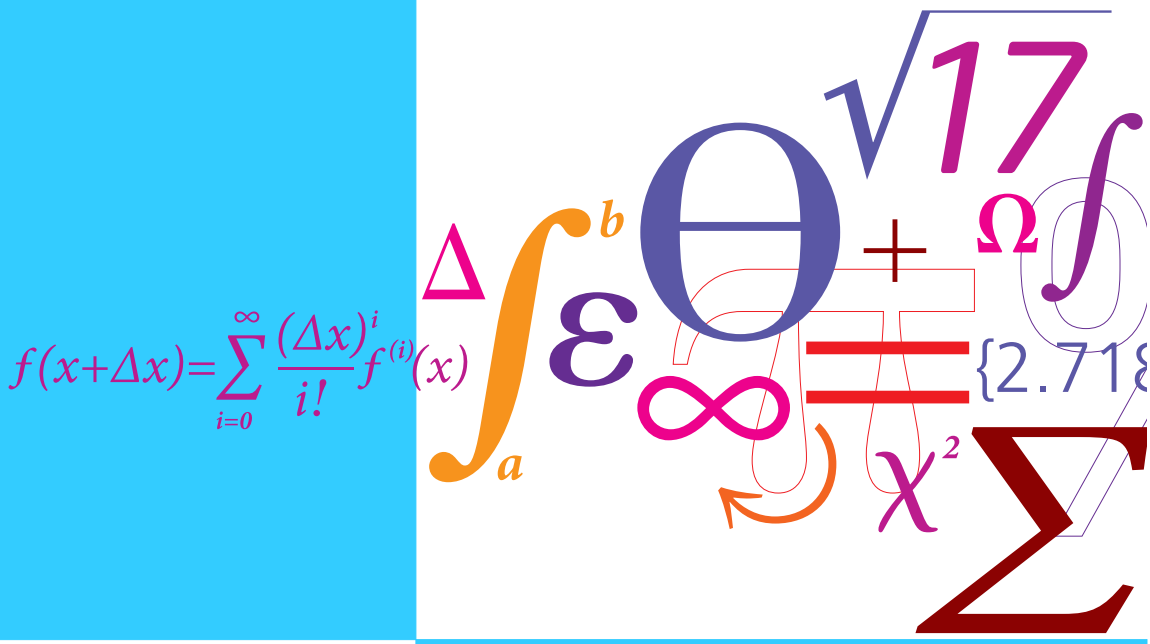
Imberger, Marc; Larsén, Xiaoli Guo; Davis, Neil

Publication date:
2024

Document Version
Publisher's PDF, also known as Version of record

Global Atlas for Siting Parameters Ocean and Coasts - Ensemble calculation of 50-year return winds

Report
E-0244



Marc Imberger (maim@dtu.dk)
Xiaoli Guo Larsén
Neil Davis

DTU Wind and Energy Systems
ISBN: 978-87-87335-72-0

April 2024

1 Introduction

This report shows some results and describes the methodology behind the ensemble calculation of 50-year return winds for oceans and coastal areas during the Global Atlas for Siting Parameters Ocean and Coast (GASPOC[1]) Project. These data complement the high-resolution onshore/nearshore global atlas for siting parameters[2]. The enhancements include the expansion of data coverage to oceans, the inclusion of two more reanalysis products, and the further improvement/enhancement of the methodology for coastal and tropical cyclone affected areas. The input reanalysis data is described in Section 2, the workflow is described in Section 3, and the results are presented in Section 4 including final outputs and possible applications. The rest of the report is rounded out with a summary section (Section 5), and information on where to obtain the data (Section 6). More in-depth analysis related to the reported findings is documented in the Appendix.

2 Processing reanalysis products including CFSR, MERRA2, CFDDA and ERA5

Three reanalysis products were selected as input for calculating the ensemble of global 50-year return wind siting parameters. The products included are the Climate Forecast System Reanalysis (CFSR, Saha et al. 2010) data set, which was used for the Global Atlas for Siting Parameters (GASP[3],Larsén, Davis, Hannesdóttir, Kelly, Olsen, et al. 2021) project, the European Centre for Medium-Range Weather Forecasts Reanalysis 5 (ERA5, Hersbach et al. 2020) data set and the Modern-Era Retrospective Analysis for Research and Applications version 2 (MERRA2, Gelaro et al. 2017) data set. A summary of key parameters is given in Table 1.

The Climate Four-Dimensional Data Assimilation (CFDDA, Rife et al. 2014) data set, which has shown good performance as supporting background reanalysis product for 50-year return wind calculations in WAsP Engineering (WEng), was also investigated for inclusion. However, it was found that the 10-m diagnostic wind speed, which is key to estimating the 50-year return wind was found to have significant errors in a band of $\pm$ 3 degrees latitude around the Equator (cf. Appendix A.2). Therefore, the CFDDA dataset was omitted from the pool of reanalysis products used in this study, since it is not suitable for the near-global coverage aimed for in this project. It must be noted that CFDDA product is neither compromised outside the critical band around the equator, nor for wind speeds calculated on the original model vertical levels.

Table 1: Summary of key parameters for the three main reanalysis products (CFSR, MERRA2 and ERA5) and CFDDA.

Name	Period available[†]	Period used[†]	Grid spacing (lat x lon)	Output frequency
CFSR	1979 - 2010	1979 - 2010	0.3125 x 0.3125 deg	1-hourly
MERRA2	1980 - now	1980 - 2020	0.5 x 0.625 deg	1-hourly
ERA5	1950 - now	1979 - 2020	0.25 x 0.25 deg	1-hourly
CFDDA	1985 - 2005	1985 - 2005	0.4 x 0.4 deg	1-hourly

[†]limits included

The following key variables have been processed and downloaded: time series of horizontal wind speed components at 10-m, time series of surface characteristics (depending on availability either surface roughness length, friction velocity or surface stress), land-sea masks and topography. This processed data set provides the input for the workflow further described in Section 3.

3 Workflow for calculation of near-global 50-year return winds

3.1 Overview

The methodology is based on the workflow previously developed for the Global Atlas for Siting Parameters (further referred to as GASP workflow). The largest difference is around the treatment of coastal zones and tropical cyclone affected areas (tc-areas). An updated flow diagram is presented in Figure 1.

[1] `https://www.gaspoc.com/`

[2] `https://science.globalwindatlas.info/#/map` (Larsén, Davis, Hannesdóttir, Kelly, Svenningsen, et al. 2022)

[3] `https://data.dtu.dk/articles/dataset/Global_Atlas_of_Siting_Parameters_V1/14753349/1`

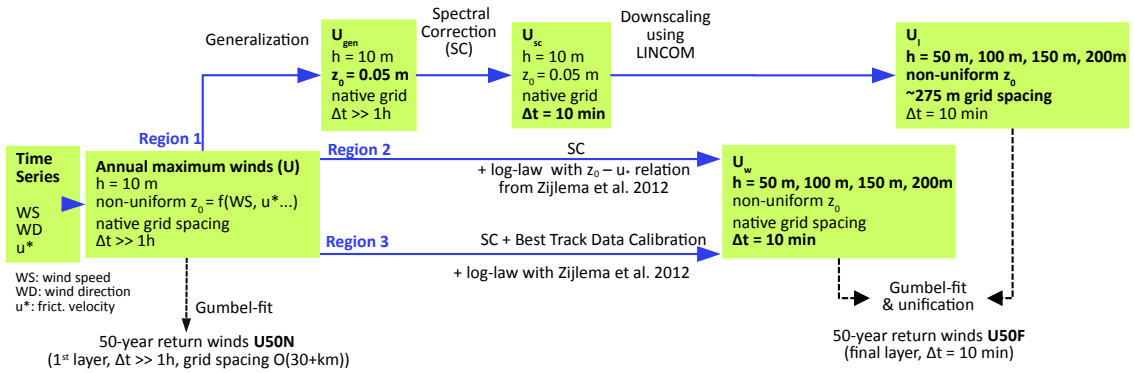


Figure 1: Updated flow diagram of the calculation pipeline starting from the global reanalysis product (left) to the final product (right) with the distinct paths for the three major geographical regions: onshore and non-tropical cyclone (non-tc) affected coastal areas with distances between 0 km and 200 km to shore (Region 1), non-tc affected offshore regions with shore distances larger than 50 km (Region 2), offshore regions (all distances) affected by tropical cyclones (Region 3). Overlapping regions between Region 1 and 2 (50 km - 200 km) are unified using the methodology described in Section 3.2.

As displayed in Figure 1, the workflow is defined differently for three geographical regions: Region 1, contains onshore and non-tropical cyclone (non-tc) affected coastal areas within 200-km of the shore; Region 2, includes non-tc affected offshore regions more than 50 km from the shore; and Region 3, for all offshore regions affected by tropical cyclones. The overlapping regions between Region 1 and 2 (50 km - 200 km) are unified using the methodology described in Section 3.2. All regions require the derivation of annual maximum winds (cf. Figure 1). For ERA5, additional processing was required to filter unreasonable high values in isolated locations due to numerical instabilities (cf. Appendix A.1). The detailed breakdown of the calculations performed within each region is as follows:

- Region 1:
  1. Apply DTUs generalization approach to the annual maximum winds following the methodology described e.g. in Troen and Lundtang Petersen (1989) or Badger et al. (2014) to obtain the generalized extreme winds.
  2. Determine spectral correction factors r_{sc} following the methodology in Larsén, Ott, et al. (2012) to correct for missing variability in the modeled wind speed time series.
  3. Downscale the generalized extreme wind using the LINCOM model (Astrup, Jensen, et al. 1996; Astrup and Larsen 1999) to target resolution of approx. 275 m and apply spectral correction factor.
- Region 2:
  1. Determine spectral correction factors r_{sc} as in Region 1.
  2. Use log-law relationship in combination with a surface roughness-friction velocity relation (Zijlema et al. 2012) to extrapolate annual maximum wind speeds to target heights and apply spectral correction factors. This is the same method as used in GASP (more details about this step can be found in Larsén, Davis, Hannesdóttir, Kelly, Svenningsen, et al. 2022).
- Region 3:
  1. Use spectral correction adapted for tropical cyclone affected areas to obtain correction factors. The method is conceptually based on Larsén and Ott (2022) but has been adapted further as part of the project (more details in Section 3.3).
  2. Use log-law relationship for extrapolation like in Region 2 and apply adapted spectral correction factors.

For the individual regions, 50-year return winds are calculated via Gumbel-fitting and subsequently unified as described in Section 3.2.

Since GASPOC has special focus on coastal and ocean applications, efforts have been made to enhance the representation of those areas in the calculation pipeline which are described in the subsections below.

3.2 Treatment of nearshore coastal areas not affected by tropical cyclones (tc).

In GASP, nearshore coastal areas, both onshore and offshore, not affected by tropical cyclones (non-tc areas) were modeled using a blending technique that combined the model results from the onshore and offshore pipeline of the original GASP workflow (not shown). This ensured a minimum level of consistency in the transition zone in coastal regions. This was done by first using the Earth System Modeling Framework's (ESMF, Hill et al. 2004) "creep filling" technique (Kara et al. 2007) to extrapolate the offshore pipeline results across the coastline. The extrapolated results were used for all offshore regions between 200 km and 0 km offshore, and replaced near-coast onshore model results for the following conditions: 1) The location was within 50 km to the coastline and the terrain elevation is below 500 m; and 2) the extreme wind estimate of LINCOM was higher than the creep filled extrapolated data. This was due to unreasonably high wind speeds in the coastal region when using the LINCOM model.

In GASPOC, LINCOM downscaling was used for the non-tc affected coastal areas. Over water, LINCOM models the surface roughness length using a Charnock-like relationship between the friction velocity (u_*) and roughness length (z_0). The Charnock coefficient varies with wind speed and fetch length (derived from distance to coast from high-resolution land-sea mask), according to a pre-calculated lookup-table. The look-up table combines empirical relations derived from Johnson et al. 1998 and Hasselmann et al. 1979, and converges to a simple Charnock-relation with α=0.011 for open water (Johnson et al. 1998). This methodology not only ensures a more consistent treatment of the offshore-onshore transition (since it's now being part of the same modeling chain), but also replaces the pure extrapolation-driven creep-filling method with a more physics based approach in those areas. The LINCOM derived results are used without modification for the first 50-km offshore. Between 50 km and 200 km offshore, a weighted linear combination following Equation 1 of model results from Region 1 (100% at 50 km) and Region 2 (100% at 200 km) is used to transition between the two models. The scaling factor γ decreases linearly from 1 (at 50 km) to 0 (at 200 km).

$$U_{50,combined} = \gamma U_{50,Region1} + (1-\gamma) U_{50,Region2} \tag{1}$$

For tc-affected coastal regions, the workflow still relies on a spatial extrapolation of the non-generalized, tc-corrected 50-year return winds obtained from the reanalysis product based on the creep-filling technique offshore.

3.3 Expansion of tropical cyclone affected areas

The original method of defining tc-affected areas has been revisited and extended. In GASP, tc-affected areas were defined using three focus areas, based on fixed latitude/longitude bounds from Larsén and Ott 2022. Here, we extend the region using data from the International Best Track Archive for Climate Stewardship (IBTrACS, Knapp, Kruk, et al. 2010; Knapp, Diamond, et al. 2018). To create a likelihood region of tropical cyclones, 33 years of IBTrACS track records (1990 to 2023, Fig. 2a) were used as input to a non-parametric statistical method (Kernel Density Estimation, KDE) to derive a proxy for the spatial probability density function in space. The following settings have been used: A Gaussian kernel shape with bandwidth of 0.04, Haversine distance metric (to acknowledge the spherical nature of the underlying dataset) and the Ball Tree algorithm for the nearest neighbor search. Once evaluated, a cut-off minimum likelihood threshold of 0.008 is defined to identify tc-affected locations (colored area in Figure 2b). It must be noted that this methodology does not distinguish between weak and strong tropical cyclones, but only considers their presence or absence in IBTrACS.

While the updated method provides a data-driven expansion of what is considered a tc-affected area, it is acknowledged that the updated region will likewise not fully encapsulate all possible locations of tropical cyclones.

3.3.1 Imposing upper limits for tc-aware spectral correction method in ERA5

In addition to the spatial expansion of the tc-affected areas, the extension of the tc-correction method described in Larsén and Ott 2022 has been adapted to work with the other reanalysis products. Without adaption, application to the ERA5 reanalysis resulted in a significant over-correction of the 50-year return winds with values above 100 ms^{-1} at 100 m. This was fixed by introducing an additional upper limit for b_2 (Equ. 7 in Larsén and Ott

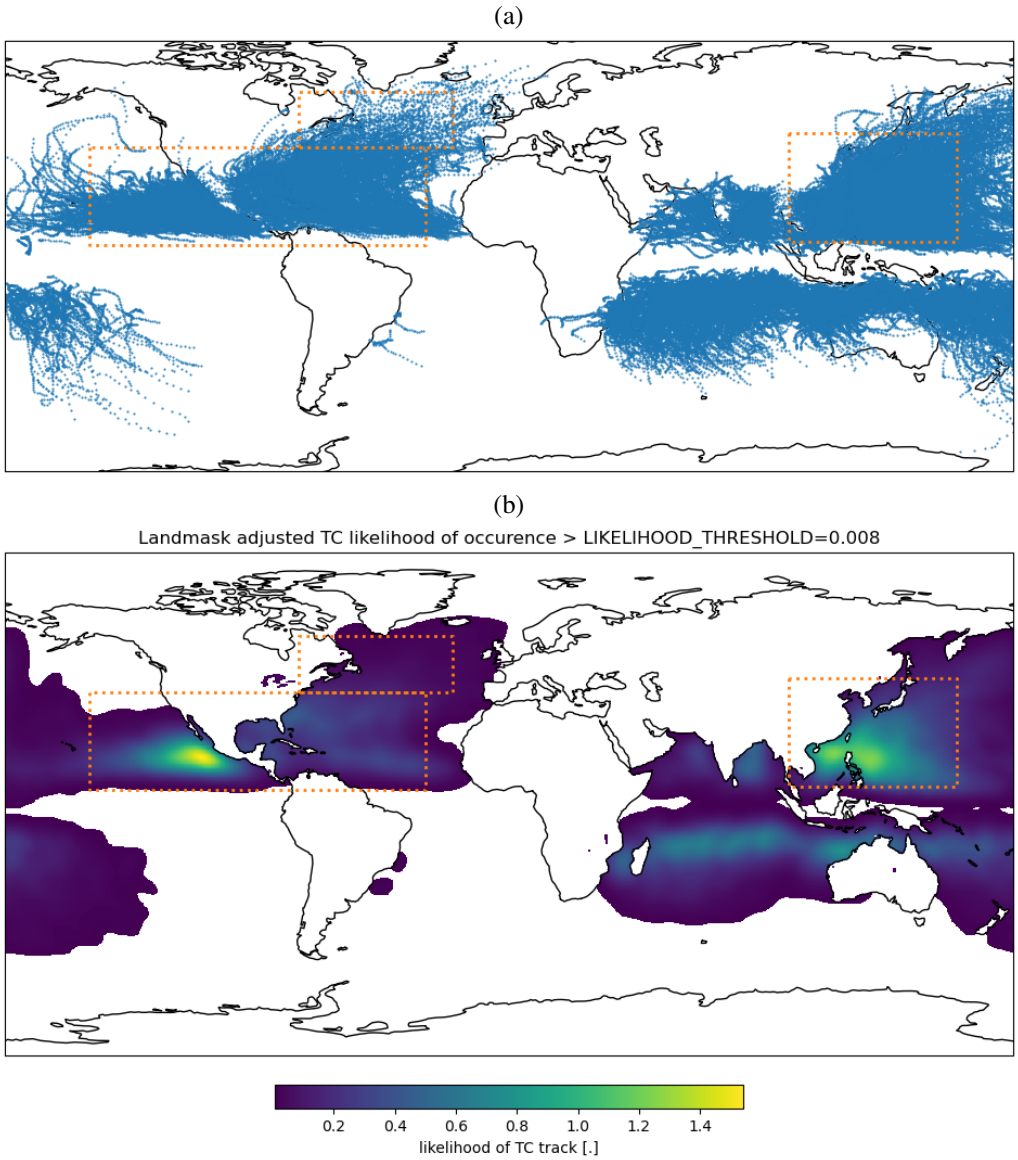


Figure 2: (a) Point cloud created from tropical cyclone best track records from IBTrACS between 1990 and 2023 (blue). (b) Likelihood of tropical cyclone presence derived from Kernel Density Estimate. Values below the cut-off minimum threshold of 0.008 and estimates over land have been removed. The dashed orange lines represent the former tc-affected area as defined in Larsén and Ott 2022.

2022) during the calibration process with the best track data. For more background details, the reader is referred to the description in Appendix A.3. It must be noted that this change did not impact the calibration of the CFSR or MERRA2 products since both products remained below the imposed threshold for all points.

4 Results

4.1 Ensemble of 50-year return winds

Figure 3 depicts the ensemble of the 50-year return estimates from CFSR, MERRA2 and ERA5 at two (50 m and 150 m) of the four modelled heights. A visual inspection already reveals differences in spatial features and magnitudes in the three products. This is particularly visible in the width of low-wind regions west of both South America and Africa, the Indian Ocean, the southern high-latitudes, and the typical tropical-cyclone hot spot areas in the North Atlantic and Western Pacific/Philippine Sea. While the absolute magnitudes are relevant for the estimation of siting parameters, the ensemble method also provides the opportunity to derive spread and other statistical metrics to identify critical regions of strong agreement/disagreement, which is demonstrated in the following section.

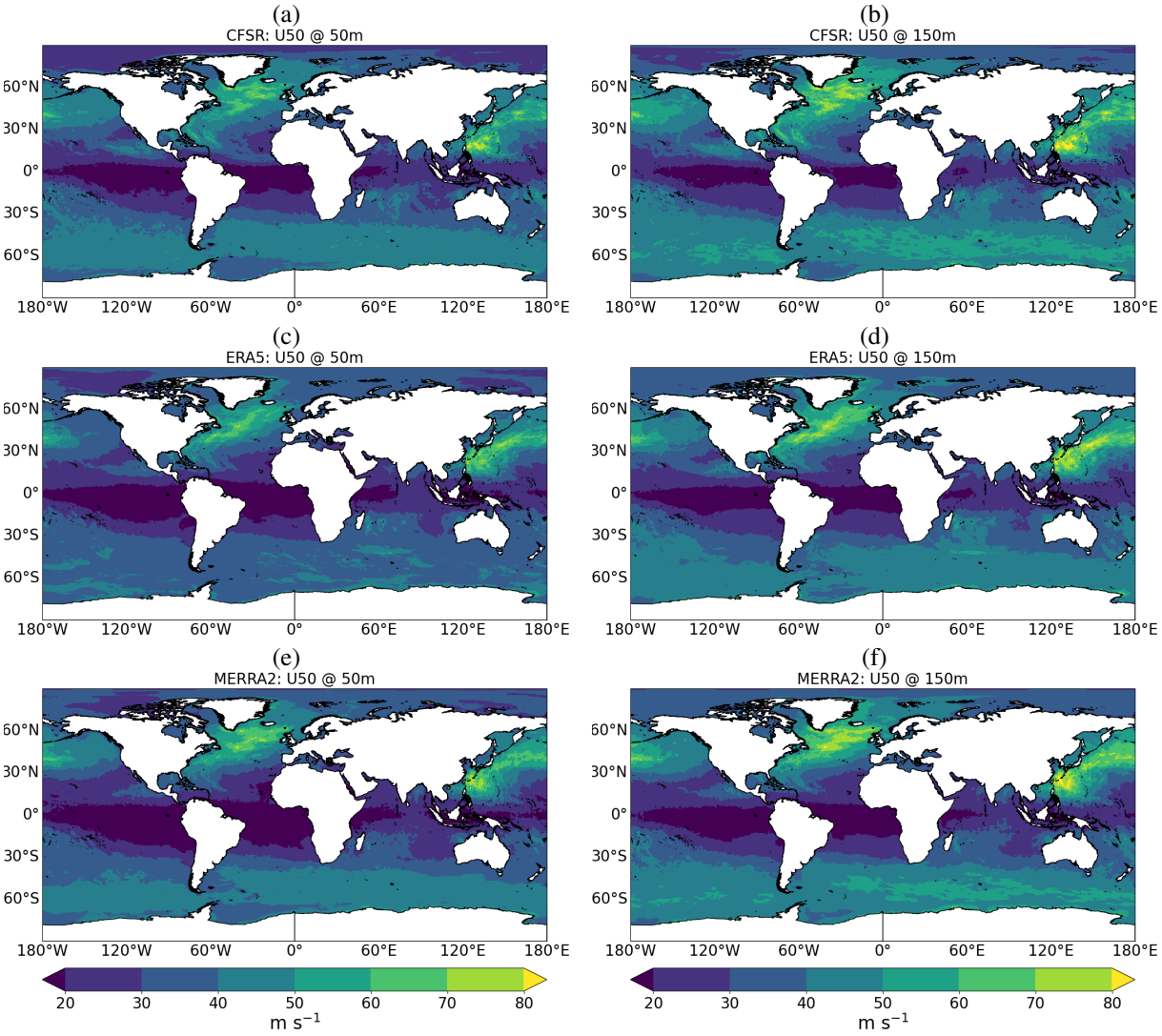


Figure 3: Offshore 50-year return winds (U50) derived from CFSR (a,b), ERA5 (c,d) and MERRA2 (e,f) at 50 m (left column) and 150 m (right column)

$$r = \frac{U_{50,\text{median}} - U_{50,\text{mean}}}{U_{50,\text{max}} - U_{50,\text{min}}} = \frac{U_{50,\text{II}} - \frac{1}{3}\left(U_{50,\text{I}} + U_{50,\text{II}} + U_{50,\text{III}}\right)}{U_{50,\text{III}} - U_{50,\text{I}}} \quad \text{with} \quad U_{50,\text{I}} \leq U_{50,\text{II}} \leq U_{50,\text{III}} \tag{2}$$

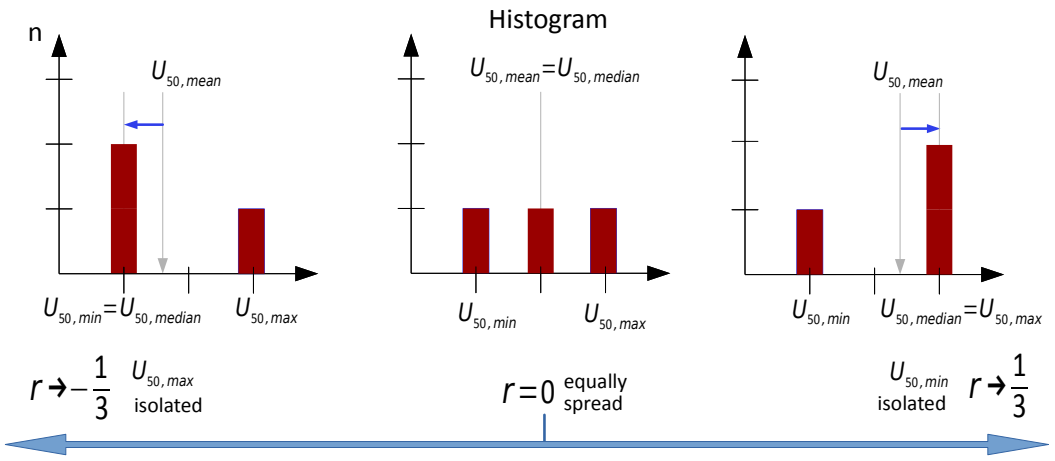


Figure 4: Illustration of the discrepancy index concept as defined in Equation 2 for the three extreme cases: $r = 1/3$ (left): locations where two products agree in their estimate while the estimate of the 3rd product provides an isolated maximum value, $r = 0$ (center): the three estimates are distributed equidistantly, $r = 1/3$ (left): locations where two products agree in their estimate while the estimate of the 3rd product provides an isolated minimum value.

4.2 Variability, Discrepancy and Agreement Classification

Within the scope of the project, four major metrics have been used to identify critical regions: 1) absolute spread, the reanalysis product that provides the 2) minimum and 3) maximum value at a given location and 4) a discrepancy index (Eq. 2). The discrepancy index is used as a proxy to describe how isolated a single product is from the other two products. Figure 4 illustrates the concept of the discrepancy index. When analyzed together, they can provide a good overview of the variability, consistency and robustness of 50-year return wind estimates at a particular location.

Global versions of the above mentioned parameters for the example height of 150 m are depicted in Figure 5. While regions of small absolute spread (Figure 5a) are located in the low-wind regions in the Pacific and Atlantic west of South America and west of Africa, the largest absolute spread between the three models are in the tropical cyclone affected areas and in particular the central Atlantic region north of the Equator (Figure 5a). At the same time, this region is characterized by a highly negative discrepancy factor (r < -0.25), indicating that the model that provides the maximum estimate (CFSR in this case, cf. Figure 5d) is isolated from the other two models, which provide similar results (cf. illustration in Figure 4). While this does not say anything about the quality of the estimates, it does provide insight in terms of how isolated a certain ensemble member is. In other high spread regions (like the high-latitudes in the southern hemisphere), the discrepancy index tends to indicate the opposite (highly positive, r > 0.25) suggesting that the product providing the minimum estimate provides the isolated estimate (in this case ERA5, cf. Figure 5c). Interestingly, the spatial distribution of products providing minima and maxima estimates reveal significant global patterns. CFSR has the highest values in almost all offshore regions, except the Arctic and Southern Ocean (Figure 5d). The minimum estimates show a strong latitude dependency. At low-latitudes MERRA2 has the lowest estimates in the low-latitudes, while ERA5 and CFSR provide the lowest estimates in mid-latitude and high-latitude regions, respectively.

A combination of absolute spread and discrepancy can be used to identify critical regions where the spread or discrepancy are especially large. One such classification is illustrated in Figure 6 which uses a 6-class classification scheme (Table 3) based on spread and discrepancy. While Figure 6 provides high granularity, it is mostly the critical cut-off values for the different turbine classes that are of relevance for wind energy siting applications. Especially critical are the areas, where estimates do not agree on the suitable turbine class. To calculate the turbine class, the

upper threshold values for the reference wind speed for the IEC turbine classes I, II, III, T and S (cf. Table 2) have been taken to map each point in the 50-year return wind layers to a turbine class. This is then used to check if all three products suggest the same IEC class at a given location (i.e. agreement) or not (i.e. disagreement), for each grid cell. The outcome is depicted in Figure 7 highlighting critical areas with model disagreement. In areas with disagreement the choice of reanalysis product has a significant impact on the use of the estimated 50-year return wind as a siting parameter.

Wind turbine class	U_{50} threshold [m/s]
T	57
I	50
II	42.5
III	37.5
S	> 57

Table 2: Basic 10-min equivalent, 50-year return wind (U_{50}) upper limits for each of the main IEC wind turbine class types. Class S has been assigned to areas with fall under none of the previous four categories. Values taken from IEC 61400-1 (International Electrotechnical Commission 2019).

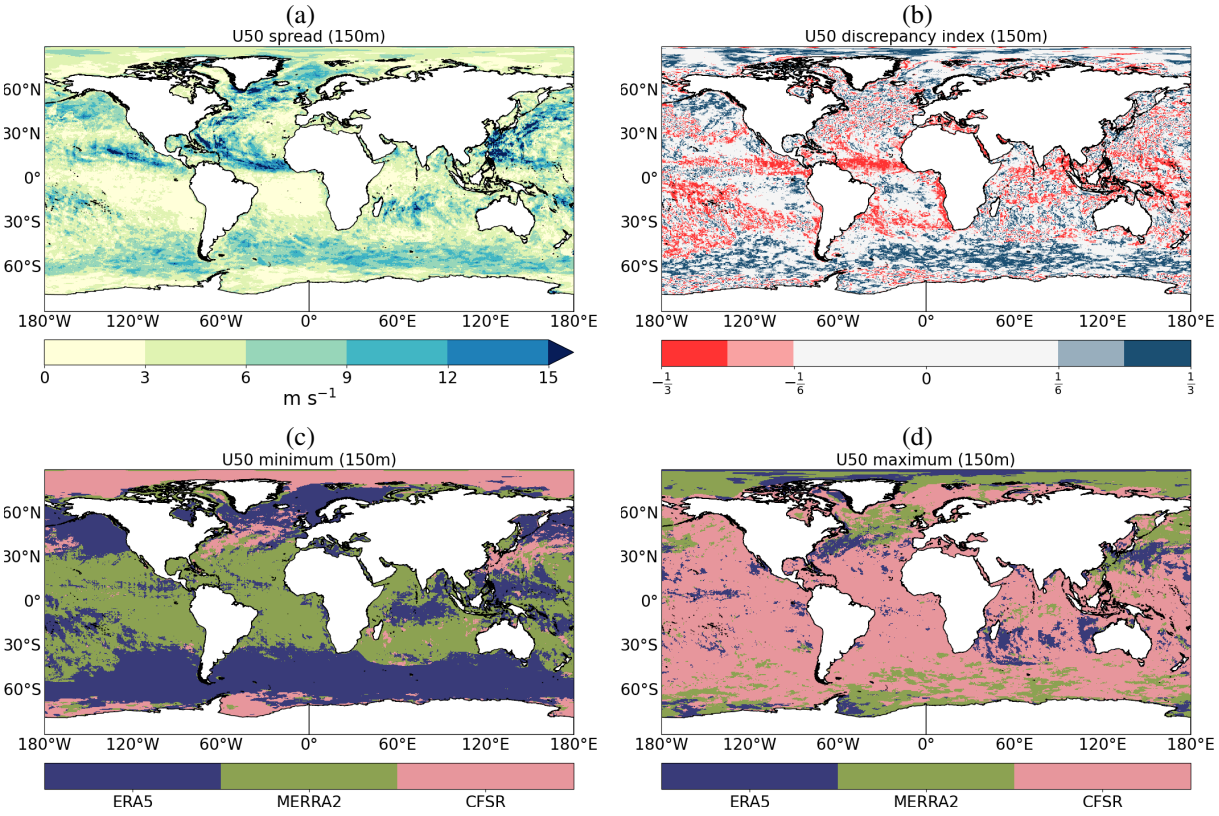


Figure 5: Overview of metrics for classifying variability and discrepancy at 150 m: (a) absolute spread of 50-year return winds, (b) discrepancy index as well as a depiction of which product provides the (c) minimum and (d) the maximum estimate at a given location.

Class ID	Definition
1	small absolute spread ($s \leq 2.5$ ms^{-1}) with identified isolated ensemble member ($\|r\| > 1/6$)
2	small absolute spread ($s \leq 2.5$ ms^{-1}), no identified clear isolated ensemble member ($\|r\| < 1/6$)
3	medium absolute spread (2.5 ms^{-1} $< s \leq 5$ ms^{-1}) with identified isolated ensemble member ($\|r\| > 1/6$)
4	medium absolute spread (2.5 ms^{-1} $< s \leq 5$ ms^{-1}), no identified isolated ensemble member ($\|r\| < 1/6$)
5	high absolute spread ($s > 5$ ms^{-1}), with identified isolated ensemble member ($\|r\| < 1/6$)
6	high absolute spread ($s > 5$ ms^{-1}), no identified isolated ensemble member ($\|r\| > 1/6$)

Table 3: Definition of the six classes behind the area classification map in Figure 6 using thresholds for absolute spread (s) and discrepancy index (r).

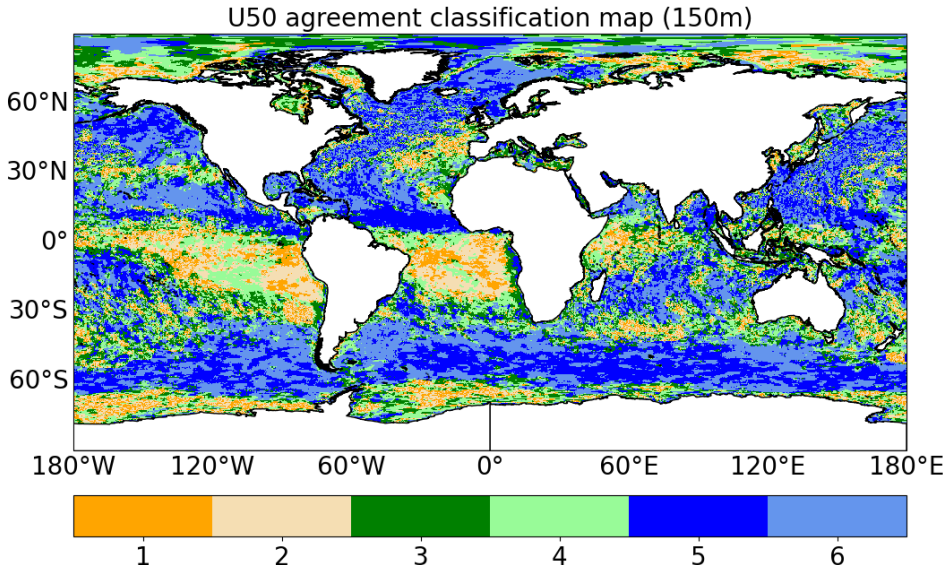


Figure 6: Six-class area classification map based on spread and discrepancy index at 150 m following the class definition in Table 3.

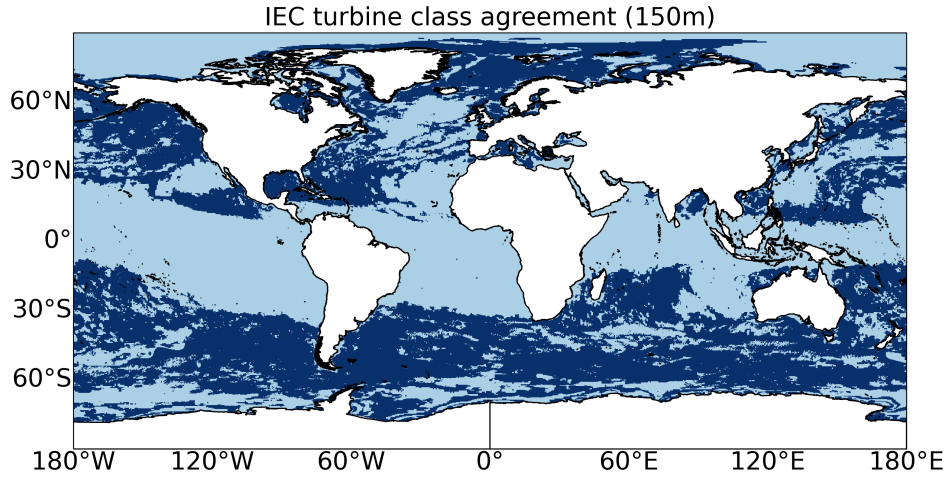


Figure 7: Turbine classes: areas where model derived IEC classes (I, II, III, T and S, cf. Table 2) of CFSR, MERRA2 and ERA5 agree (light blue) and disagree (dark blue).

4.3 Detailed Perspectives

In the following sections, two areas of interest are selected for a more in-depth perspective on the data set, namely the North Sea region (Section 4.3.1) and Taiwan Strait (Section 4.3.2). Both regions are relevant for offshore wind energy development, but they also provide good example cases to highlight some of the insights the GASPOC 50-year return wind data layers can provide.

4.3.1 North Sea

Over a large part of the North Sea, there is a major discrepancy between the ERA5 and the MERRA2 and CFSR based 50-year return wind estimates (Figure 8). Over most of the region, the ERA5 results are the lowest (Figure 8f). The discrepancy index is highly positive for this region (Figure 8e), suggesting that the product providing the minimum estimate (i.e. ERA5) is isolated (cf. definition in Figure 4). With the exception of the area North of Scotland, there is no clear spatial pattern regarding the maximum estimate, with the CFSR and MERRA2 results fairly equally spread (Figure 8g). The discontinuities observed near the Shetland Islands (Figure 8a-c) compared to the rest of the North Sea area are a consequence of the tc-adjusted spectral correction method being applied there. In this region, MERRA2 provides the maximum estimate but the highly negative discrepancy suggests that this is an isolated estimate. The largest spread is observed around 54.5N, 0.3W, along the Eastern coast of England. The large spread is caused by the artificially high and low estimates in the CFSR product in this area (Figure 2a,d). This is caused by a highly non-linear relationship between the annual maximum winds and the associated friction velocity, which lead to errors in the generalization process (more detailed analysis in Appendix A.4). Similar areas of large spread affected by this imprints can be found at several places around the globe.

From a wind turbine siting perspective, the agreement on the adequate IEC class estimate between the products is quite low (Figure 2h) making it a very critical region when it comes to uncertainties associated with 50-year return winds (low robustness of the class estimation).

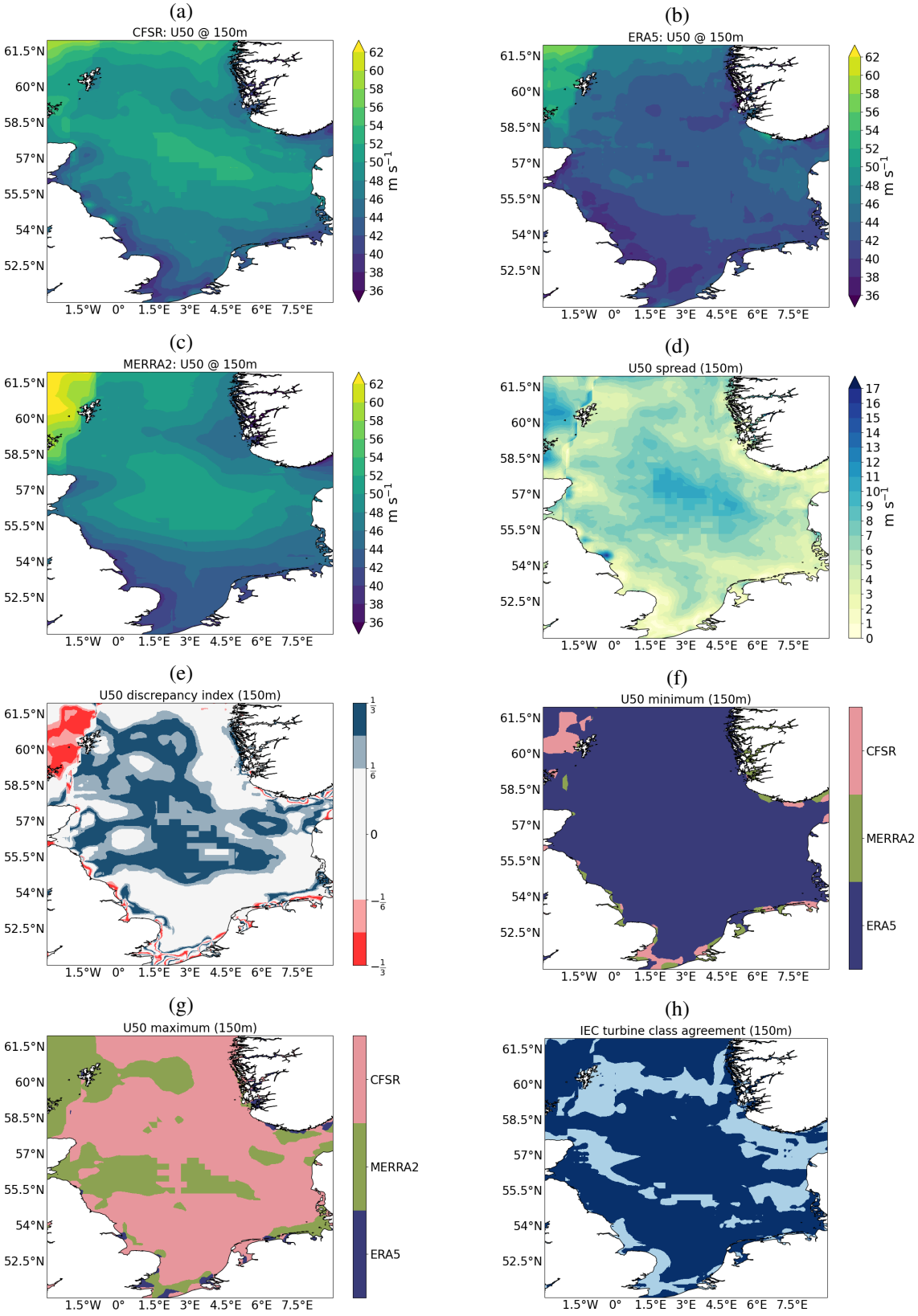


Figure 8: North Sea focus area at height of 150 m: offshore 50-year return winds (U50) derived from (a) CFSR, (b) ERA5 and (c) MERRA2. (d) Absolute spread between U50s, (e) discrepancy index, (f) smallest and (g) largest estimate at a given location out of the three model estimates. (h) Regions of IEC turbine class agreement (light blue) and disagreement (dark blue).

4.3.2 Taiwan Strait

The Taiwan Strait is part of the tc-affected zone where the tc-adjusted spectral correction method is used (Region 3 in workflow, Section 3). Near-coastal details are less detailed (Figure 9a-c) compared to non-tc affected regions due to the reliance on the creep filling extrapolation technique in this area (as described in Section 3.2). In the northern part of the strait, large absolute spread with magnitudes above 17 ms^{-1} is observed (Figure 9d), while the spread is less pronounced in the southern part. The large spread is partly driven by MERRA2, which shows an isolated (Figure 9e) maximum value (cf. Figure 9g). ERA5 shows the smallest estimates over the western part of the strait (Figure 9f), while CFSR shows the lowest values around Taiwan. Similar to the North Sea, agreement in suitable IEC class is low (Figure 9f), indicating a high uncertainty when it comes to the robustness of the choice of a suitable IEC turbine class.

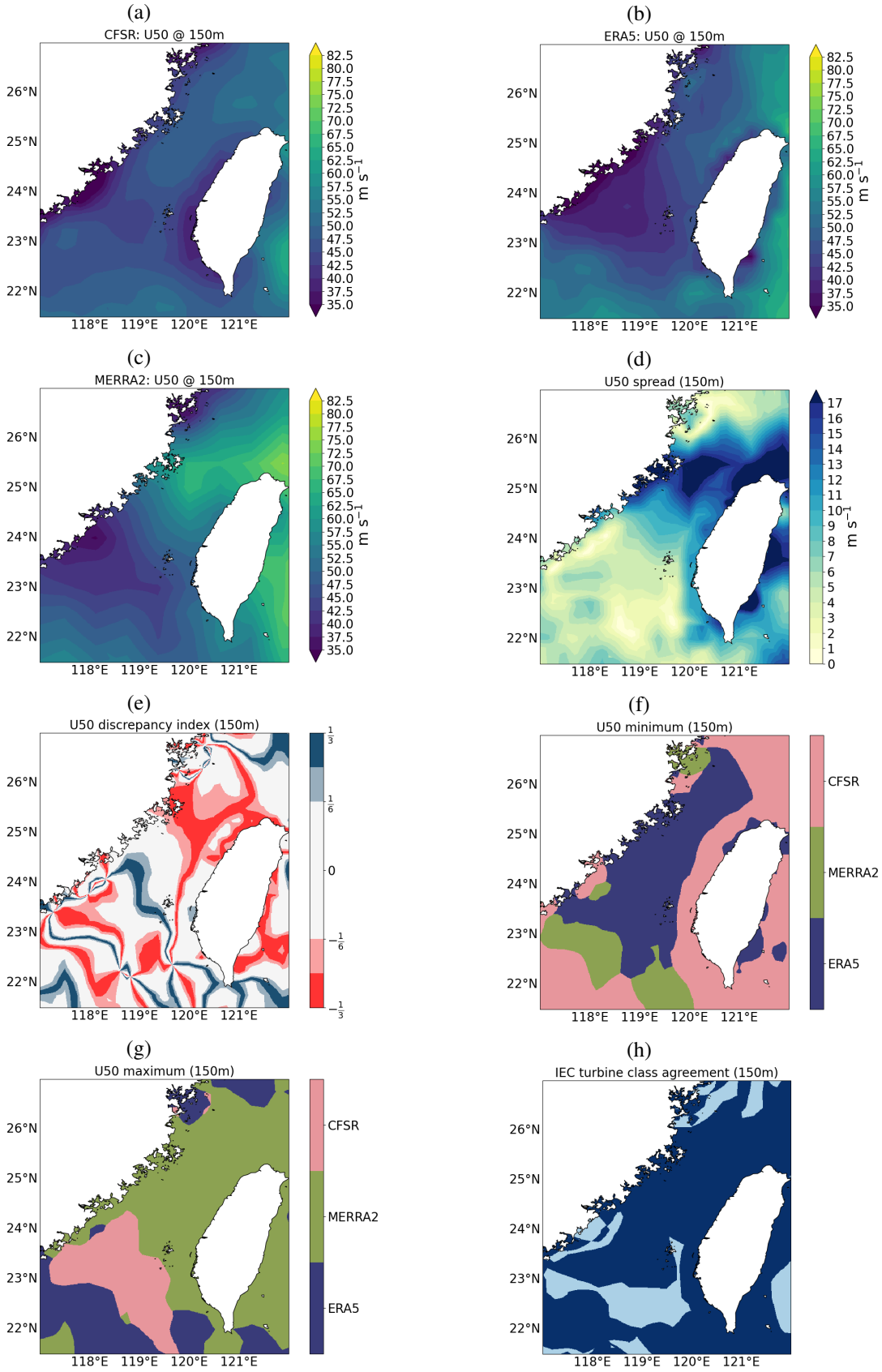


Figure 9: Taiwan Strait focus area at height of 150 m: offshore 50-year return winds (U50) derived from (a) CFSR, (b) ERA5 and (c) MERRA2. (d) Absolute spread between U50s, (e) discrepancy index, (f) smallest and (g) largest estimate at a given location out of the three model estimates. (h) Regions of IEC turbine class agreement (light blue) and disagreement (dark blue).

5 Summary & Conclusion

Within the scope of the GASPOC project, near-global 50-year return winds data layers, with focus over coastal and ocean locations based on three different reanalysis products (CFSR, ERA5 and MERRA2) have been created. For this purpose, the workflow developed during the creation of the onshore-focused Global Atlas for Siting Parameters has been further enhanced by improving the near-coastal treatment of non-tc affected sites, and extending the tc-adjusted spectral correction method from Larsén and Ott 2022.

When comparing the results from the different reanalysis products, significant differences were found. One example that is important for offshore wind turbine siting is the robustness of IEC class estimates. These were found to deviate across boundaries in many regions of the globe, including in the North Sea and Taiwan Strait, which were highlighted in this report. It is expected that these results and datasets will provide insights for high-level guidance, and may also be of value as complementing material for more refined wind turbine siting studies.

6 Data Availability

The calculated 50-year return wind layers, which build the foundation of the presented analysis, are made available for download on `data.dtu.dk` and will be visualized at `https://science.globalwindatlas.info/#/map`. Results from each reanalysis product dataset, CFSR, MERRA2 and ERA5, are provided at four heights, 50 m, 100 m, 150 m and 200 m, and for two resolutions:

- A high resolution layer with approximately 275 m grid spacing contains data for the coastal region, between 0 km and 200 km distance from the shoreline. It is made available on the same grid and spatial extension as the Global Wind Atlas for convenience.
- A coarser 0.25x0.25 degree grid is provided for far-shore data, more than 200 km from the coastline. This dataset includes data points within the coastal zone. However, it is not recommended to use it in those regions.

7 Acknowledgments

The work behind this report has been funded by the EUDP Global Atlas for Siting Parameters Oceans and Coasts (GASPOC) project (J. no. 64020-1043). The authors acknowledge the computational resources provided on the Sophia HPC Cluster at the Technical University of Denmark (Technical University of Denmark 2019).

References

Astrup, P., N.O. Jensen, and Torben Mikkelsen (1996). *Surface roughness model for LINCOM*. English. Denmark. Forskningscenter Risoe. Risoe-R 900(EN). ISBN: 87-550-2187-5.

Astrup, P. and Søren Ejling Larsen (1999). *WAsP engineering flow model for wind over land and sea*. English. Denmark. Forskningscenter Risoe. Risoe-R 1107(EN). ISBN: 87-550-2529-3.

Badger, Jake, Helmut Frank, Andrea N. Hahmann, and Gregor Giebel (2014). "Wind-Climate Estimation Based on Mesoscale and Microscale Modeling: Statistical–Dynamical Downscaling for Wind Energy Applications". In: *Journal of Applied Meteorology and Climatology* 53.8, pp. 1901–1919. DOI: 10.1175/JAMC-D-13-0147.1.

Campos, Ricardo M., Carolina B. Gramcianinov, Ricardo de Camargo, and Pedro L. da Silva Dias (2022). *Assessment and Calibration of ERA5 Severe Winds in the Atlantic Ocean Using Satellite Data*. en. DOI: 10.3390/rs14194918. URL: http://dx.doi.org/10.3390/rs14194918.

CFDDA Documentation (2014). *Additional Information regarding Input Datasets, Input Conditions, and Data Processing*. URL: https://rda.ucar.edu/datasets/ds604.0/docs/Initial_Input_Conditions_and_Data_Processing_Info.pdf.

Gelaro, Ronald et al. (2017). "The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2)". In: *Journal of Climate* 30.14, pp. 5419–5454. DOI: 10.1175/JCLI-D-16-0758.1.

Grell, Georg A, Jimy Dudhia, David R Stauffer, et al. (1994). "A description of the fifth-generation Penn State/N-CAR Mesoscale Model (MM5)". In.

Hansen, Brian Ohrbeck, Xiaoli Guo Larsén, Mark C. Kelly, Ole Steen Rathmann, Jacob Berg, Andreas Bechmann, Anna Maria Sempreviva, and Hans Ejsing Jørgensen (2016). *Extreme Wind Calculation Applying Spectral Correction Method – Test and Validation*. English. DTU Wind Energy E. Denmark: DTU Wind Energy.

Hasselmann, K. et al. (1979). *Measurements of wind-wave growth and swell decay during JONSWAP*. Tech. rep. Deutsches Hyrdographisches Insititut.

Hersbach, Hans et al. (2020). "The ERA5 global reanalysis". In: *Quarterly Journal of the Royal Meteorological Society* 146.730, pp. 1999–2049. DOI: https://doi.org/10.1002/qj.3803.

Hill, C., C. DeLuca, Balaji, M. Suarez, and A. Da Silva (2004). "The architecture of the Earth System Modeling Framework". In: *Computing in Science Engineering* 6.1, pp. 18–28. DOI: 10.1109/MCISE.2004.1255817.

International Electrotechnical Commission (2019). *IEC 61400-1:2019: Wind energy generation systems – Part 1: Design requirements*. Technical Report. International Electrotechnical Commission. URL: https://www.iec.ch/standards/61400-1-ed-4-0.

Johnson, H. K., J. Højstrup, H. J. Vested, and S. E. Larsen (1998). "On the Dependence of Sea Surface Roughness on Wind Waves". en. In: *Journal of Physical Oceanography* 28.9, pp. 1702–1716. DOI: 10.1175/1520-0485(1998)028<1702:otdoss>2.0.co;2.

Kara, A. Birol, Alan J. Wallcraft, and Harley E. Hurlburt (2007). *A Correction for Land Contamination of Atmospheric Variables near Land–Sea Boundaries*. en. DOI: 10.1175/jpo2984.1. URL: http://dx.doi.org/10.1175/JPO2984.1.

Knapp, Kenneth R., Howard J. Diamond, James P. Kossin, Michael C. Kruk, and Carl J. Schreck (2018). *International Best Track Archive for Climate Stewardship (IBTrACS) Project, Version 4*. DOI: 10.25921/82TY-9E16. URL: https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.ncdc:C01552.

Knapp, Kenneth R., Michael C. Kruk, David H. Levinson, Howard J. Diamond, and Charles J. Neumann (2010). "The International Best Track Archive for Climate Stewardship (IBTrACS)". en. In: *Bulletin of the American Meteorological Society* 91.3, pp. 363–376.

Larsén, Xiaoli Guo, Jake Badger, Andrea N. Hahmann, and Niels G. Mortensen (2012). "The selective dynamical downscaling method for extreme-wind atlases". en. In: *Wind Energy* 16.8, pp. 1167–1182. DOI: 10.1002/we.1544. URL: http://dx.doi.org/10.1002/we.1544.

Larsén, Xiaoli Guo, Neil Davis, Ásta Hannesdóttir, Mark Kelly, Bjarke Olsen, Rogier Floors, Morten Nielsen, and Marc Imberger (2021). *Calculation of Global Atlas of Siting Parameters*. English. E-Report-0208. Denmark. URL: https://orbit.dtu.dk/en/publications/calculation-of-global-atlas-of-siting-parameters.

Larsén, Xiaoli Guo, Neil Davis, Ásta Hannesdóttir, Mark Kelly, Lasse Svenningsen, René Slot, Marc Imberger, Bjarke Tobias Olsen, and Rogier Floors (2022). "The Global Atlas for Siting Parameters project: Extreme wind, turbulence, and turbine classes". en. In: *Wind Energy* 25.11, pp. 1841–1859. DOI: 10.1002/we.2771.

Larsén, Xiaoli Guo and Søren Ott (2022). "Adjusted spectral correction method for calculating extreme winds in tropical-cyclone-affected water areas". English. In: *Wind Energy Science* 7.6, pp. 2457–2468. ISSN: 2366-7443. DOI: 10.5194/wes-7-2457-2022.

Larsén, Xiaoli Guo, Søren Ott, Jake Badger, Andrea N. Hahmann, and Jakob Mann (2012). "Recipes for Correcting the Impact of Effective Mesoscale Resolution on the Estimation of Extreme Winds". In: *Journal of Applied Meteorology and Climatology* 51.3, pp. 521–533. DOI: 10.1175/JAMC-D-11-090.1.

Rife, Daran L, James O Pinto, Andrew J Monaghan, Chris A Davis, and John R Hannan (2014). *NCAR Global Climate Four-Dimensional Data Assimilation (CFDDA) Hourly 40 km Reanalysis*. DOI: 10.5065/D6M32STK.

Rife, Daran L, James O Pinto, Andrew J Monaghan, Christopher A Davis, and John R Hannan (2010). "Global Distribution and Characteristics of Diurnally Varying Low-Level Jets". In: *Journal of Climate* 23.19, pp. 5041–5064. DOI: 10.1175/2010JCLI3514.1.

Saha, Suranjana et al. (2010). "The NCEP Climate Forecast System Reanalysis". In: *Bulletin of the American Meteorological Society* 91.8, pp. 1015–1058. DOI: 10.1175/2010BAMS3001.1.

Technical University of Denmark (2019). *Sophia HPC Cluster*. DOI: 10.57940/FAFC-6M81. URL: https://dtu-sophia.github.io/docs/.

Troen, Ib and Erik Lundtang Petersen (1989). *European Wind Atlas*. English. Risø National Laboratory. ISBN: 87-550-1482-8.

Zijlema, M., G.Ph. van Vledder, and L.H. Holthuijsen (2012). "Bottom friction and wind drag for wave models". en. In: *Coastal Engineering* 65, pp. 19–26. DOI: 10.1016/j.coastaleng.2012.03.002.

A Appendix

A.1 Mitigation of unrealistically high wind speeds in ERA5s 10-m diagnostic for extreme wind calculations

During the calculation of the 50-year winds, unrealistic wind speeds, in the order of several hundreds of meters per second, have been observed at specific locations over the globe. The location and severity depend on the analyzed wind sector, e.g. Figure 10 depicts an example for the 50-year return winds at the 10-m height for winds from the south (inflow $180° \pm 15°$). A closer analysis of the underlying 10-m wind speed diagnostic revealed that this is caused by spike-like jumps, as depicted in Figure 11. The existence of such structures was acknowledged in the paper describing the ERA5 dataset (Hersbach et al. 2020) and attributed to instabilities in the 4D-Var tangent-linear physics. A practical mitigation strategy for past events was also suggested (removal of critical points where max(U,V) > 50 m s^{-1}, cf. ERA5 Online Documentation[4]). However, this is not sufficient when looking at extreme winds, since neighboring time-steps, around the spike, are also affected, and have higher than expected winds even if the wind speed is below the cutoff (cf. blue curve in Figure 11). In this study, an alternative filtering strategy was developed, which removes not only the critical points, but also all values within a time window centered around each critical point (cf. orange curve in Figure 11). A window of $\pm$ 6 hours around each critical point has been found to be sufficiently large for this purpose. In addition to that, a second filter was applied in locations with maximum winds above 50 m/s where the relative difference between the 99.9-percentile and actual maximum value exceeds 200% (i.e. maximum is twice as large as the 99.9-percentile). In those locations, the 99.9-percentile is used instead of the maximum value.

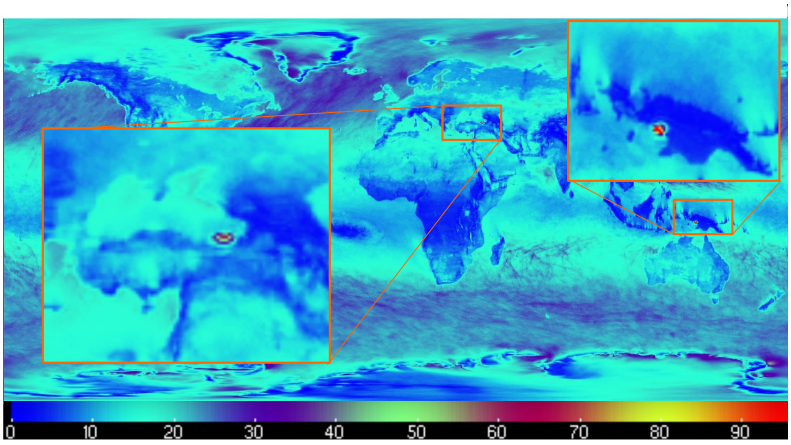


Figure 10: Depiction of global 50-year return winds (in m/s) at 10-m height determined from ERA5 using data from the southern wind direction sector (inflow 180 ± 15 deg) only: Two critical points (Caucasus region and near New Guinea) suffering from unrealistically high wind speeds are highlighted.

[4] `https://confluence.ecmwf.int/display/CKB/ERA5%3A+large+10m+winds`, last accessed: 16 Dec 2021

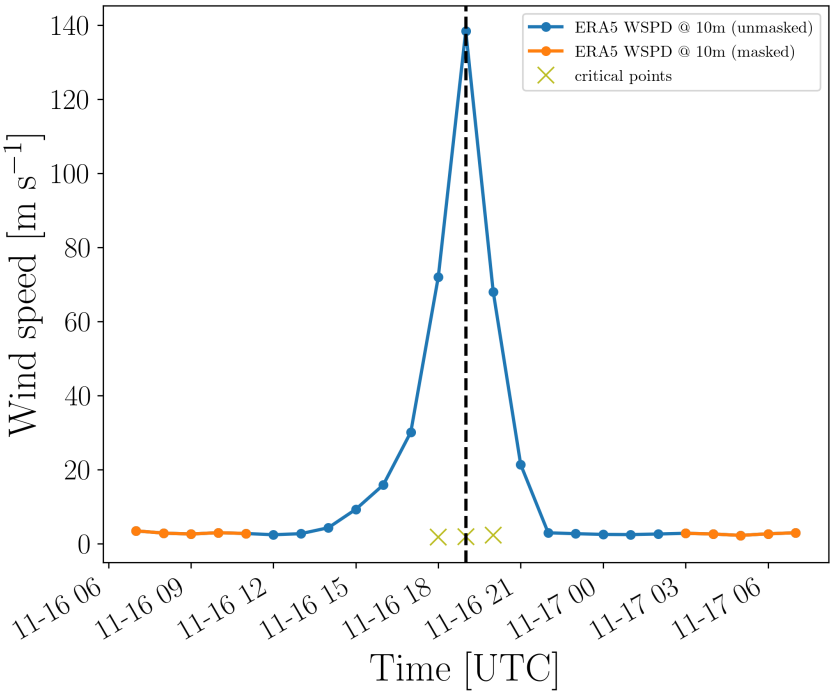


Figure 11: Extract of the 10-m wind speed time series (blue) from ERA5 during 16/17 November 2019 at the critical point near the island of New Guinea depicted in Figure 10. Yellow "x": critical points as detected by the proposed mitigation strategy in Hersbach et al. 2020. Orange line: resulting time series using temporal filtering strategy with 6-hour time window around each critical point.

A.2 Grid Effects on 10-m Wind Speed Diagnostic around Equator Region in CFDDA

Within the 10-m wind speed diagnostic of the CFDDA data set, an artificially low annual extreme wind speed was seen over a small band around the Equator (Figure 12a). While especially pronounced over the ocean, it is visible over all longitudes and affects both water and land points. A more detailed investigation of the underlying numerical model setup revealed that the affected region coincides with the location of a numerical blending zone (bright yellow band in Figure 12b). This blending zone is a result of the way the global CFDDA dataset is constructed: two, separately integrated, simulations of the PSU/NCAR mesoscale model Version 5 (MM5, Grell et al. 1994), in twin polar stereographic grids configuration centered at the two poles (Figure 12b), are joined around the equator via a Cressmann-type interpolation method (Rife et al. 2010). While the exact origin of the grid effect on the 10-m wind speed diagnostic was not identified, it overlaps exactly with the area where the interpolation method was applied. It must also be noted that the region of low wind speeds was not observed for wind speeds calculated using the model's vertical levels. This strongly indicates that the artificial effect was limited to the 10-m wind model diagnostic, and did not affect any prognostic variables in the model. However, since the extreme wind speed calculation and subsequent generalization is built on the 10-m wind diagnostic, and since it unclear how such effects could be removed retrospectively, the CFDDA reanalysis was dropped from this study.

It should be noted that 10-m CFDDA wind speeds have been used as the long-term reference wind data for implementing the spectral correction method in WAsP Engineering (WEng) in a set of key geographic areas around the globe since 2016. In February 2024, the CFDDA dataset in WEng was updated to near-global coverage, excluding the numerical blending zone [5]. It must be noted that areas outside this critical band are not affected and CFDDA data provides reasonable estimates within WEng as shown in validation studies (e.g. Hansen et al. 2016).

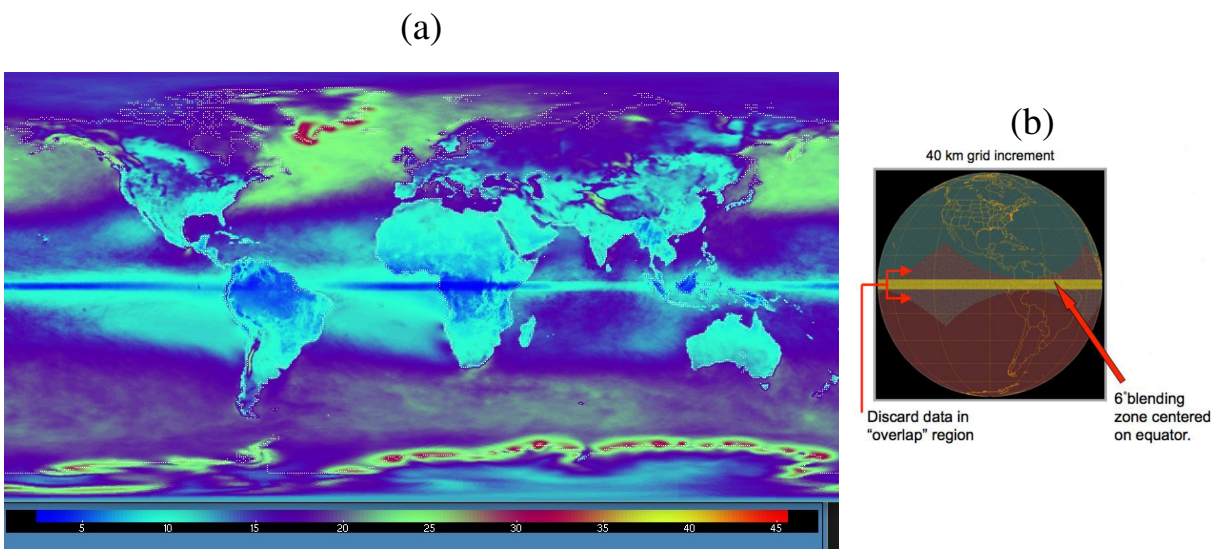


Figure 12: (a) Average of annual maximum wind speed at 10-m height (in m s^{-1}) over the full CFDDA data set period (1985-2005). (b): Depiction of the twin polar stereographic grid configuration with overlap region and blending zone of the global configuration of the PSU/NCAR mesoscale model Version 5. Image source: CFDDA Documentation 2014.

[5] near-global coverage for spectral correction method, `https://www.wasp.dk/news/nyhed?id=e105f574-437a-4a55-9636-d58b3202ed2c`, last accessed 01/03/2024

A.3 Modification of the TC-adjusted spectral correction method of Larsén and Ott 2022 to prevent over-correction in tc-affected areas in ERA5

During the application of the tc-adjusted spectral correction method to ERA5, unphysical overestimations of the 50-year return winds are observed. Investigation showed that the large over-correction is caused mainly by the underestimation of extreme wind speed in ERA5 (which has also been reported previously elsewhere, e.g. Campos et al. 2022). When applying the tc-correction algorithm from Larsén and Ott 2022, without modifications, a very high offset value (b_2, more details below) is calculated to align the 50-year return wind observed in the best-track data at the evaluation point P1 with the underestimated values from ERA5 during the calibration step (Equ. 7 in Larsén and Ott 2022). As a result, massive spectral offsets of the spectral tail with values of $n > 20$ over most of the calibration region were introduced (Figure 13a) out-competing any regional differences in spectral energy encoded in the underlying time series.

This resulted in unreasonably high 50-year return wind estimates of 100 m/s and higher (Figure 13b), which were significantly larger than those from the CFSR reanalysis (Figure 13c,d). This was corrected by introducing an offset limit within the calibration step that relates the ratio between uncorrected best-track data derived 50-year return winds (U50BT) and uncorrected reanalysis derived 50-year return winds (U50) with U50. This is further referred to as the r-u relation following the notation in Larsén and Ott 2022. By limiting the maximum offset of b_2 in Equ. 7 from Larsén and Ott 2022 to 1.3 times b_1, and thus capping vertical shifts in the r-u relation (Figure 14), more stable and physically meaningful results were obtained (Figure 13e,f). While this offset limitation prevents the absolute magnitude of ERA5 at P1 from matching the magnitude of the best-track data, the improved overall spatial distribution of 50-year return wind within the calibration domain compared to other reanalysis products like e.g. CFSR (cf. Figure 15) increases the confidence in introducing such a limit to stabilize the tc-adjusted spectral correction methodology.

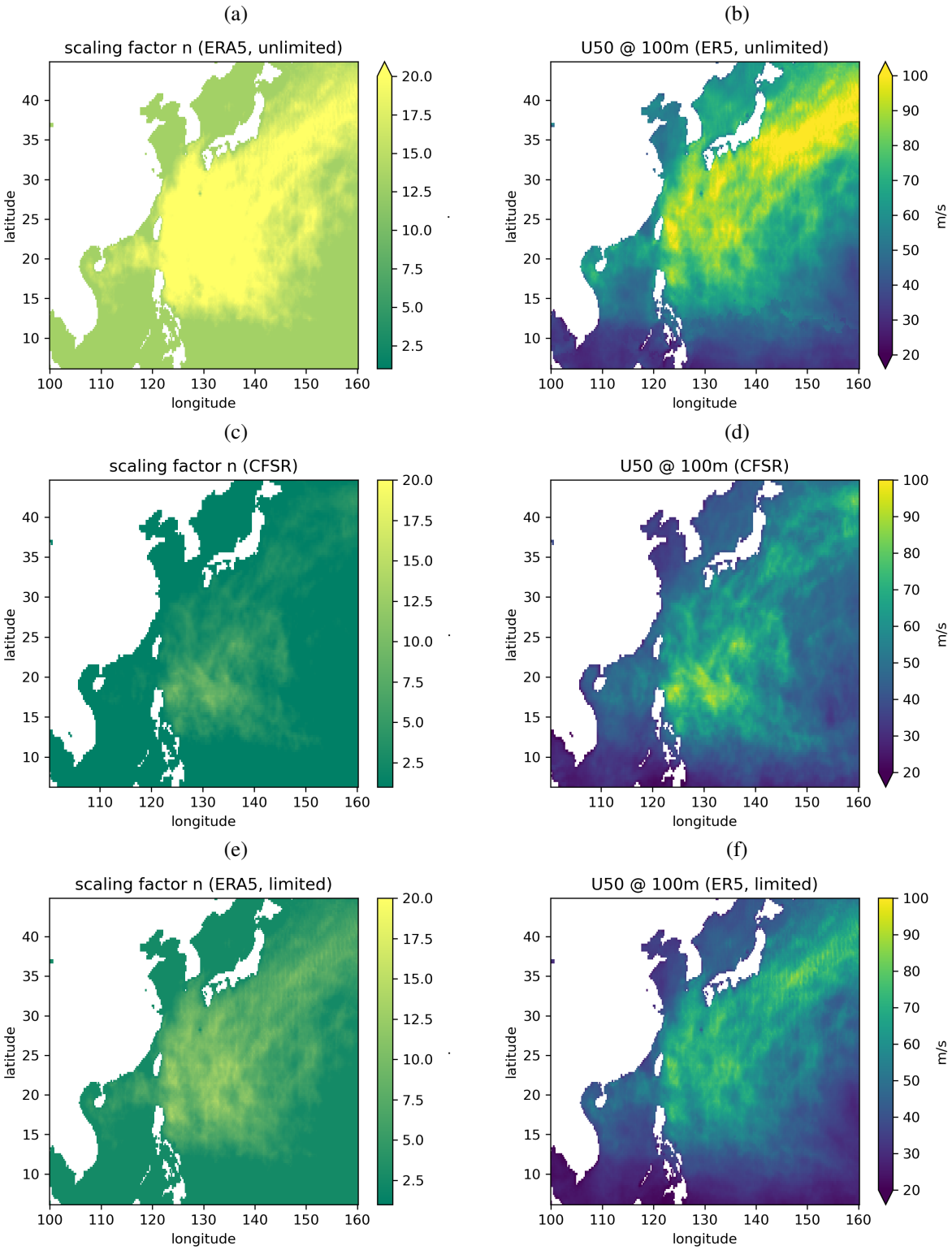


Figure 13: Scaling factor for tc-adjusted spectral correction n (left column) and spectral corrected 50-year return winds at 100 m (right column) obtained from workflow using unlimited offset shift during tc-correction workflow in ERA5 (a,b) vs.upper-bound limited offset shift (e,f). For comparison, values obtained from CFSR (c,d) are depicted as well.

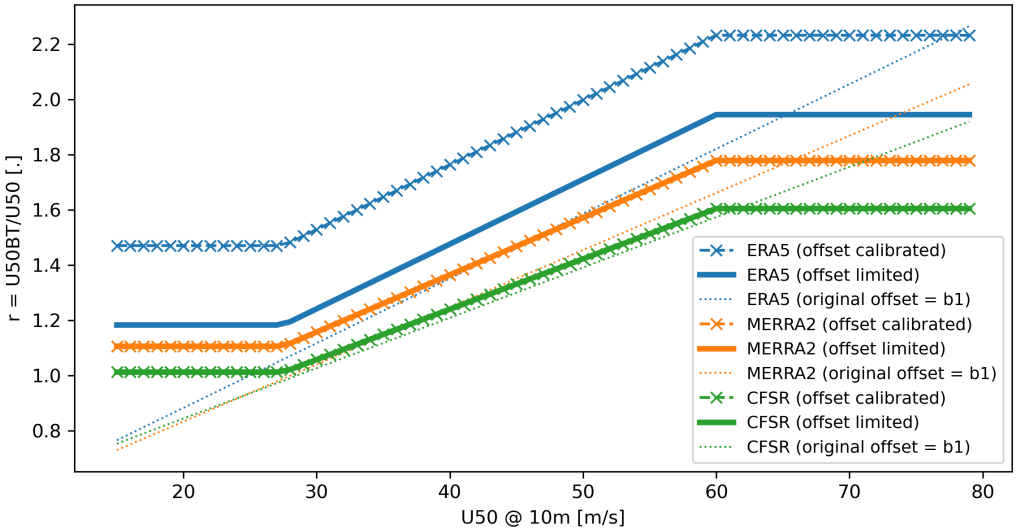


Figure 14: Illustration of offset-limitation on calibrated r-u relationship as derived in Larsén and Ott 2022 for the three reanalysis products ERA5 (blue), MERRA2 (orange) and CFSR (green): unmodified regression line (dotted line, following Eq. 6 in Larsén and Ott 2022) is calibrated using best-track data at P1 ('x'-marked dashed line, Equ. 7 in Larsén and Ott 2022). The introduction of an upper limit on the possible offset shift during calibrations limits the r-u relation to the formulation as depicted by the solid lines lines (Equ. 7 in Larsén and Ott 2022 with $b_2 \leq 1.3b_1$). Note that for MERRA2 and CFSR, calibrated and offset-limited lines are identically since upper-limit threshold on free parameter b_2 is not exceeded.

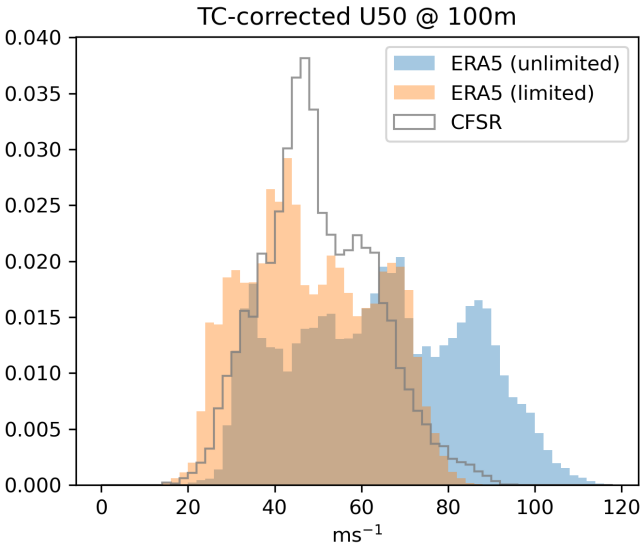


Figure 15: Normalized histogram of 50-year return winds (U50) at 100 m over the calibration region (Eastern dashed frame in Fig. 2a): tc-aware spectral corrected U50 from ERA5 before and after the imposed threshold limits on b_2 are depicted in blue and orange, respectively. Gray: U50 from CFSR over the same region.

A.4 Near-coastal generalization challenges in CFSR associated with high non-linearity between annual maximum winds & friction velocity.

As part of the generalization process, a map of local roughness lengths $z_{0,loc}$ is required to estimate and remove the local roughness effects (Badger et al. 2014). When generalizing data from numerical weather prediction (NWP) models, like those used for the reanalysis products, $z_{0,loc}$ needs to be representative of the roughness length used in the model. How this is determined depends on the NWP model setup and the application for which the downscaled winds will be used. To reconstruct an empirical $z_{0,loc}$ for extreme wind conditions, we use friction velocity and 10-m wind speed information from the reanalysis products following the methodology in Larsén, Davis, Hannesdóttir, Kelly, Olsen, et al. 2021 and Larsén, Badger, et al. 2012. The methodology distinguishes between onshore (derivation based on logarithmic wind profiles applied to annual maximum winds) and offshore locations (Charnock-like relationship adapted for extreme sea state conditions).

In the ideal case, $z_{0,loc}$ is valid for all n annual maximum winds. This means that the functional relationship between the i-th annual maximum wind, $U_{amm,i}$, and it's associated modeled friction velocity, u_*, can be described adequately by the logarithmic wind profile in Equ. 3.

$$U_{amm,i} = \frac{u_{*,i}}{\kappa} \ln\left(\frac{10\text{m}}{z_{0,loc}}\right) \tag{3}$$

Only in this case, will the generalization be bias-free in terms of over-correction (artificial enhancement of the affected annual maximum wind during the generalization) and under-correction (artificial damping of the affected annual maximum wind during the generalization).

While the generalization based on Equ. 3 performs adequately at most locations, and particularly well for MERRA2 ($U_{amm,i}$-u_* relation almost perfectly linear), a large spread is observed in some locations. This is particularly seen when using the CFSR product. Figure 16 shows the relationship between U_* and wind speed for a sample location in the East of the UK. The MERRA2 data falls along the gray line representing the linear relationship between the two functions based on the log law. However the ERA5 and particularly the CFSR are more scattered around the line. For annual maximum winds above the gray line, the generalization estimates a too large speed-up of the wind (over-correction), while it's too low for annual maximum winds below the gray line (under-correction). This misinterpretation translates to an artificial increase/decrease of the corresponding annual maximum wind after reintroducing the high-resolution roughness features during the LINCOM downscaling step. Depending on the rank of the over/under-corrected annual maxima, compared to the other years in the dataset, this bias can greatly impact the Gumbel-fit, creating regionally overly high or low final 50-year return wind estimates as seen in e.g. Section 4.3.1.

Reasons for the log-law detachment of $U_{amm,i}$ and u_* can be diverse. For CFSR, it is likely due to the way the friction velocity has been derived in the project. CFSR reconstructs u_* from hourly-averaged surface stress instead of instantaneous values, which might have filtered out some of the high-frequency variability to accurately estimate the friction velocity associated with a particular extreme event. In an attempt to correct for the bias seen in the results, a wind-sector dependent formulation of Equ. 3 and subsequent averaging was tried (not shown). Unfortunately this did not improve the accuracy of $z_{0,loc}$.

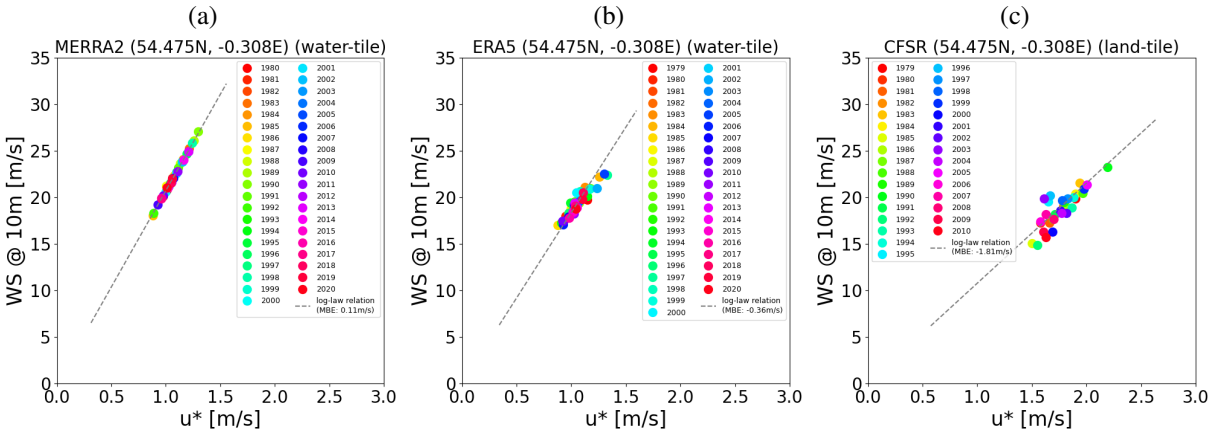


Figure 16: Annual maximum winds derived from 10 m winds on native grid over friction velocity u_* (colored dots). The dashed line shows the U_{amm}-u_*-relationship described by the log-law relation with surface roughness length $z_{0,loc}$ (Equ. 3). Plots are obtained at the native grid point of the given reanalysis product closest to 54.48N, 0.31E. Abbreviations: MBE: mean bias error between Equ. 3 and colored dots.

Technical University of Denmark

DTU Risø Campus
Frederiksborgvej 399
4000 Roskilde
Tel: + 45 46 77 50 85
https://www.vindenergi.dtu.dk/english