Product Description Document (PDD)

Anemometer dataset

Contributors: Ivan Manso-Narvarte, Lohitzune Solabarrieta, Ainhoa Caballero, Asier Anabitarte and Jose A. Fernandes-Salvador

Version	Changes made by	Nature of changes
V_2023_07	I. Manso-Narvarte, L. Solabarrieta, A. Caballero, A. Anabitarte, J.A. Fernandes-Salvador	Document generation

1 Introduction

Anemometers are tools designed to obtain wind direction and speed. Wind data have different utilities in the maritime sector, either for the estimation of fuel consumption or for the calculation of routing algorithms models. Moreover, they also have oceanographic applications. The measured data can be sent to public repositories or used for calibration or validation of other data available in public repositories and to assimilate into models. This document describes a dataset built in the framework of the SusTunTech project from wind data collected onboard several vessels in the Northwestern Indian Ocean from June 2021 to October 2023.

2 Data provenance

The wind data come from anemometers placed on vessels, which measure the speed and direction of the apparent wind relative to the vessel. Therefore, it is necessary to compensate for the vessels' speed and heading to establish the true wind speed and direction (relative to the true North). The data from the anemometer are transferred by serial cable to an Ethernet converter, which in turn are sent to a Ratatosk server, where information of other instruments onboard is also gathered, to finally send them in near real-time. Throughout this process, Ratatosk will receive relative wind data collected by the anemometer and use the information received from other instruments to calculate the true wind speed and direction. In case GPS signals fail, Ratatosk will use previous data to estimate the ship's speed and if gyrocompass data are missing, it will use GPS data to estimate the direction. To reduce vessels' disturbances in the measurements, anemometers were positioned on the foreside of the bow pole as far as possible from the vessel structure and at around 7.5 m from the deck (Figure 1).

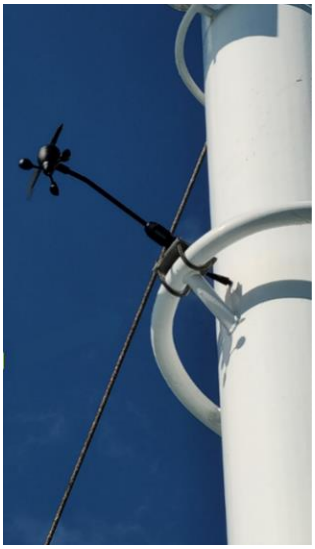

Figure 1. *Anemometer installation in one vessel.*

3 Dataset description

1 Hz resolution data were collected from June 2021 to October 2023, in the Northwestern Indian Ocean, providing a huge number of observations. After onboard correcting the observations for referencing them to the true North, the spurious data were removed. Then, the possible effects of the pole or other structures near the anemometers in the measurements were analysed. Measurements were adapted, when necessary, as explained in the following section.

4 Data QC, processing, and validation

The true speed and direction data were quality controlled (QC) based on the QARTOD manual of the U.S. Integrated Ocean Observing System (2017) selecting the QC filters that best suited the data and adapting them if needed. The QC filters used are listed in Table 1 and were applied in a sequential way.

***Table 1**. Sequentially applied filters for QC raw data.*

FILTER	DESCRIPTION
F1. Time	Check that the dates were realistic and increasing in time. The dates that were not realistic were removed and they were sorted if they did not increase in time. If dates were duplicated the first one was removed.
F2. Location	Check for realistic locations: latitudes (longitudes) above or below 90° (180°) and -90° (-180°), respectively, were removed.
F3. Realistic values	Check for direction and speed values that were not realistic. Direction values below or above 0 rad and 2π rad were removed, while speed values higher than 100 m/s were removed as well.
F4. Flat line	When some anemometers fail, the result can be a continuously repeated observation of the same value. Therefore, when an observation had the same value or almost the same value (within a range of 0.001 m/s and 0.001 rad for speed and direction respectively) as the previous 1 and/or 2 observations it was removed.
F5. Spikes	Check for outliers found for values higher or lower 8 times the standard deviation from the mean within a moving window of 200 observations for each observation. Outliers were removed. This filter was applied to speed data and not for the direction due to the gap from 0 and 2π radians.
F6. Visual inspection	Visual inspection of the data series for ensuring that flat lines and outliers were removed. Uncommon patterns that might remain in the data series were also checked.

After the QC, 4.09% of the data were removed. Once the observations were QC, the effects of the pole or any other structure on the wind measurements were analysed. To this end, marked changes in the intensity of the average wind speed when the wind blew from specific directions in the reference system of the vessel (i.e. apparent wind) were searched. Then, for dates corresponding to problematical apparent wind directions data were removed. Note that this does not mean that specific true directions are removed but apparent directions that can have any true direction depending on the heading of the vessel. In Figure 2 the average true wind speed with respect to the direction in the reference system of two of the vessels (herein vessel 1 and vessel 2) shows a decrease in the intensity when the wind blows from the bow and stern of the vessel. Given the dimensions of the pole and the distance of the anemometer to it, the measurements of the winds blowing from the stern and within a range of 25° (i.e. 180° ± 25°) should be affected by the pole as observed in the figure. In addition, a decrease in the wind

intensity was observed for winds blowing from the bow, probably due to the perturbations caused by the structures located nearby in that direction, even if they are located at a different height than the anemometer. Thus, based on the intensities observed (as in Figure 2), winds coming from the bow within a range of 35° (i.e. 0° ± 35°) were also removed. All these removals caused the elimination of 19.28% of the remaining data.

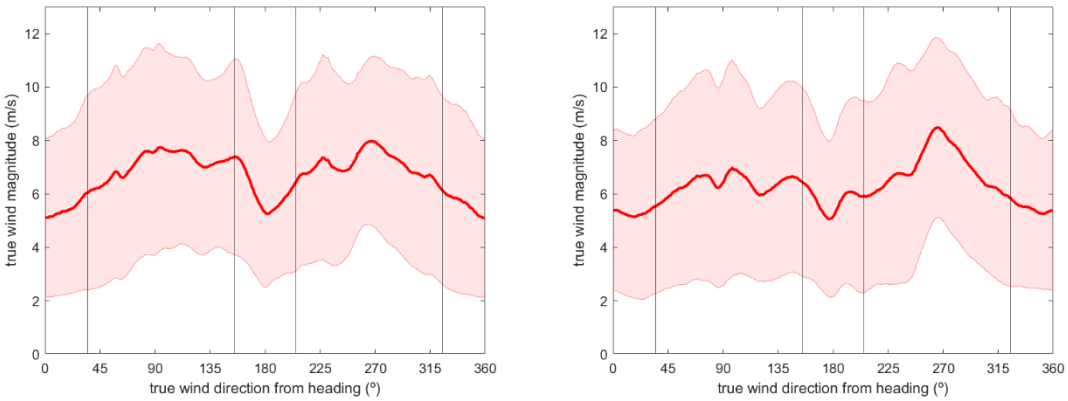

Figure 2. *Average true wind (Y-axis) relative to each wind direction with respect to the vessel heading (i.e. in the system of reference of the vessel; X-axis) for vessels 1 (left) and 2 (right). The shaded areas show the standard deviation of the averages. The vertical black lines show the angle thresholds used for the removal of measurements affected by the structures of the vessel.*

Finally, anemometer data were compared against the WIND_GLO_PHY_L4_NRT_012_004 product from Copernicus Marine Environment Monitoring Service (CMEMS) (https://doi.org/10.48670/moi-00305). This product provides hourly Level 4 information with a resolution of 0.125°. The product is a result of a combination of numerical models and satellite observations. In addition, ERA5 hourly data (DOI: 10.24381/cds.adbb2d47) were used in order to have another source. This product is the fifth generation ECMWF reanalysis for the global climate and weather and combines model data with observations from across the world with a resolution of 0.25°. Although these products heavily rely on modelled data and do not provide temporal high-resolution wind data (compared to the 1 Hz data measured by the anemometers), the comparisons enable to assess the quality of the anemometer data in general terms.

In order to make the datasets comparable, anemometer data were adapted by hourly averaging wind and position values and decomposing speed and direction into zonal (U) and meridional (V) components. Then CMEMS and ERA5 data were interpolated to the anemometer positions and correlations and root mean square differences (RMSD) were computed. Only the above-mentioned two vessels (in Figure 2) were used for the comparisons. In order to remove the exceeding variability in the anemometer data, these were additionally smoothed. Note that the anemometer data used for the comparisons ranged from the beginning of the series of each vessel until 4 October 2022 (using more than one year of data in both cases).

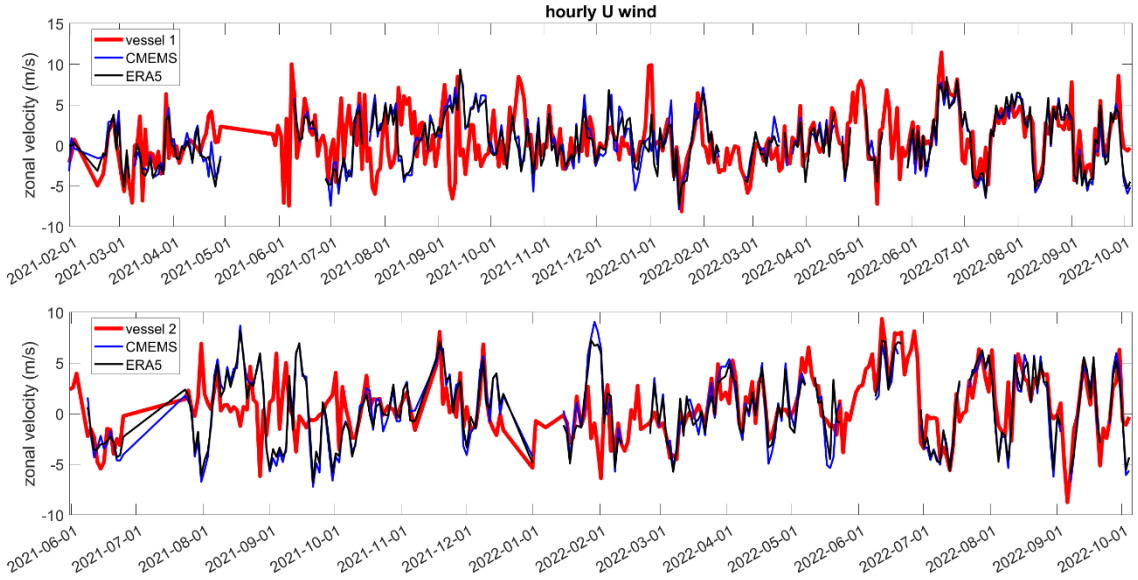

Figure 3. *Hourly zonal (top panel) and meridional (bottom panel) wind speed (in m/s) time series of vessel 1 (red), CMEMS (blue) and ERA5 (black).*

The data series (Figures 3 and 4) of both wind components show a fair agreement between anemometers and CMEMS/ERA5 data. The main seasonal patterns are detected in the three datasets.

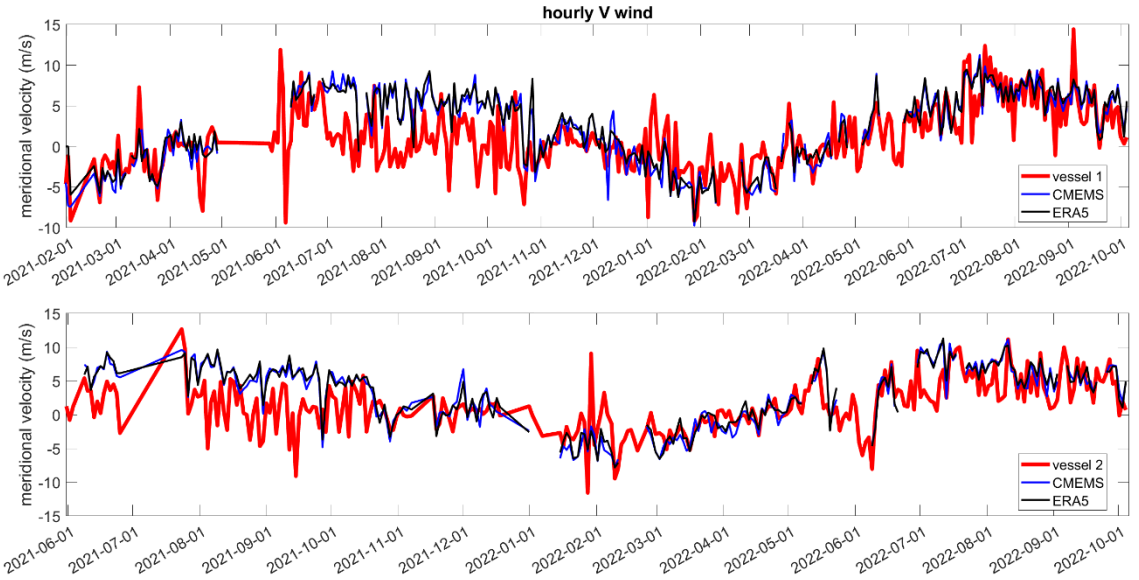

Figure 4. *Hourly zonal (top panel) and meridional (bottom panel) wind speed (in m/s) time series of vessel 2 (red), CMEMS (blue) and ERA5 (black).*

Moreover, the correlations also show a relatively good agreement (between 47 and 65%) and the RMSD are not higher than 4 m/s (see Table 2). Given that CMEMS/ERA5 data are strongly based on models and that they are being compared to observations, the veracity of the data measured by the anemometers is shown in general terms.

Table 2. *Correlation and RMSD values of the comparisons between anemometer of vessels 1 and 2 vs CMEMS and ERA5 for U and V wind components.*

	Correlation (%)				RMSD (m/s)			
	Vessel 1 vs CMEMS	**Vessel 1 vs ERA5**	**Vessel 2 vs CMEMS**	**Vessel 2 vs ERA5**	**Vessel 1 vs CMEMS**	**Vessel 1 vs ERA5**	**Vessel 2 vs CMEMS**	**Vessel 2 vs ERA5**
U	47.14	50.29	51.79	55.26	3.42	3.25	3.32	3.10
V	63.08	63.51	64.08	64.95	3.90	3.88	3.81	3.75

Since the anemometers are located at around 10 m over the water, this is the height that is going to be considered for the dataset. For the final products, wind measurements from all the vessels were decomposed into zonal and meridional components, temporally averaged and merged.

5 Product description

The wind dataset is created as an individual NetCDF file. Note that datasets published in data repositories, due to the standards of these repositories, may contain minor differences in the global attributes that do not affect the data. The general characteristics of these datasets are described in Table 3.

Table 3. *General description of the wind dataset.*

Spatial Coverage	Northwestern Indian Ocean
Spatial Resolution	Pointwise
Temporal Coverage	From June 2021 to March 2023
Temporal Resolution	0.5 hour
Variables	Zonal and meridional wind velocity; and Metadata
File format	NetCDF 1.4
Dataset location	
Access to the dataset	Open access

Table 4 shows the variables that are embedded in the dataset, which compose the bulk of the dataset.

Table 4. *Variables contained in the wind dataset of vessel 1.*

Variable name	Description	Units
TIME	Date of the data	Seconds since 1970-01-01T00:00:00z
DEPH	Depth of the measurement	m
LATITUDE	Latitude of the data position	Degrees north
LONGITUDE	Longitude of the data position	Degrees east
WSPE	West-east wind component	m/s
WSPN	Meridional sea water velocity	m/s

Table 5 provides much more information and details about the dataset and enables a deep understanding of it contributing to making the dataset FAIR (Findable, Accessible, Interoperable, Reusable). The proposed main and global attribute metadata followed the structures and standard names of the 'NetCDF CF Metadata Convention Standard Name Table Version 1.6' (https://cfconventions.org/cf-conventions/v1.6.0/cf-conventions.html). They also followed the ISOs 19115 and 19139, and several non-standard attributes. The SeaDataNet common vocabularies were also used. The use of common standard metadata and vocabularies made the dataset more FAIR.

Table 5. *Global attributes contained in the metadata of the wind dataset.*

Global attribute	Value
acknowledgment	'These data were collected in the framework of the SusTunTech project, processed by AZTI and made freely available by Marine Instruments through the SusTunTech project and the programs that contribute to it'
area	'Northwestern Indian Ocean'
cdm_data_type	'Point'
citation	'These data were collected in the framework of the SusTunTech project, processed by AZTI and made freely available by Marine Instruments through the SusTunTech project and the programs that contribute to it'
comment	'Wind velocities were obtained from the anemometers onboard the vessels. First raw data were quality controlled, and then the data affected by the vessel structures were identified and removed. The final product is an averaged set of wind velocities at 10 m'
contact_email	'sustuntech_WP4@azti.es'
Conventions	' '
creator_email	'sustuntech_WP4@azti.es'
creator_name	'AZTI'
creator_type	'Institution'
creator_url	'https://www.azti.es/'
data_assembly_center	'AZTI'
data_language	'eng'
data_mode	'D'
data_type	'Winds'
date_update	2023-10-11 22:30
distribution_statement	'These data are public and free of charge. User assumes all risk for use of data. User must display citation in any publication or product using data. User must contact PI prior to any commercial use of data'
DOI	
format_version	'1.4'
geospatial_lat_max	'13.64'
geospatial_lat_min	'-5.62'
geospatial_lat_units	'degrees_north'
geospatial_lon_max	'69.55'
geospatial_lon_min	'44.77'
geospatial_lon_units	'degrees_east'
geospatial_vertical_max	'10'
geospatial_vertical_min	'10'
geospatial_vertical_units	'm'
history	2021-06-07 06:00 - 2023-10-11 22:30 data collected. 2023-11-15 14:19 netCDF file created using Matlab software
infoUrl	'https://www.sustuntech.eu/communicationmaterials/deliverables/'
institution	'AZTI (Spain)'
institution_edmo_code	'1623'
institution_references	'AZTI'
keywords	'WINDS, FISHERIES'
keywords_vocabulary	'GCMD Science Keywords'
last_update	2023-11-15 14:19
license	'This vessel mounted anemometer wind dataset is licensed under a Creative Commons Attribution 4.0 International License. You should have received a copy of the license along with this work. If not, see http://creativecommons.org/licenses/by/4.0/'

metadata_language	‘eng’
naming_authority	‘Marine Instruments, AZTI
NetCDF_format	‘netcdf4_classic’
NetCDF_version	‘netCDF-4 classic model’
pi_name	‘ ‘
platform_code	‘ ‘
platform_name	‘ ‘
publisher_email	‘sustuntech@globalmarine.es’
publisher_name	‘Marine Instruments’
publisher_type	‘Institution’
publisher_url	‘https://www.marineinstruments.es/es/’
qc_manual	‘Data were quality controlled (QC) based on the QARTOD manual of the Integrated Ocean Observing System (IOOS, https://doi.org/10.7289/V5FX77NH) selecting the QC filters that best suited the data and adapting them if needed‘
references	‘https://www.sustuntech.eu/, https://www.sustuntech.eu/communicationmaterials/deliverables/’
Standard_name_vocabulary	‘NetCDF Climate and Forecast Metadata Convention Standard Name Table Version 1.6; and https://www.seadatanet.org/Standards/Common-Vocabularies’
summary	‘Wind velocities are derived from the anemometers onboard the vessels. First raw data were quality controlled, and then the data affected by the vessel structures were identified and removed. The final product is an averaged set of wind velocities at 10 m’
time_coverage_end	2023-10-11 22:30
time_coverage_resolution	‘0.5 hour’
time_coverage_start	2021-06-07 06:00
title	‘Vessel anemometer hourly winds at 10 m in the Northwestern Indian Ocean by SusTunTech project’
update_interval	‘void’
wmo_inst_type	‘ ‘
wmo_platform_code	‘ ‘

6 References

U.S. Integrated Ocean Observing System (2017) Manual for Real-Time Quality Control of Wind Data: a Guide to Quality Control and Quality for Coastal and Oceanic Wind Observations. Silver Spring, MD, U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, Integrated Ocean Observing System, 47pp. DOI: http://doi.org/10.7289/V5FX77NH.