Product Description Document (PDD)

Temperature 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_08	I. Manso-Narvarte, L. Solabarrieta, A. Caballero, A. Anabitarte, J.A. Fernandes-Salvador	Document generation

1 Introduction

The effects of sea temperature on many environmental, ecological, and physical processes, convert this variable into one of the most crucial ones for a wide range of marine applications. Moreover, it is a robust indicator of climate change. Nowadays, there are different platforms (e.g., satellites, moorings, Argo floats) regularly monitoring the sea temperature. However, depending on the measurement platform, measurements are not always made with high resolution, are spatially isolated, or their acquisition depends on weather conditions (due to clouds when measuring from satellite sensors).

To address the above-mentioned scientific challenges, in-situ near-surface temperature data from ships can be used to feed public data repositories, to validate other types of data and sources and to assimilate into models.

This document describes a dataset built in the framework of the SusTunTech project from near-surface temperature data, collected onboard several vessels operating in the Northwestern Indian Ocean from March 2021 to March 2023.

2 Data provenance

The data were obtained by external temperature sensors (i.e. thermometers) placed on the vessels' hulls. These sensors are connected to acoustic current Doppler profilers (ADCP) and are mainly used for calculating the sound speed to correct current velocities measured by the ADCPs. The temperature data are transferred via a serial cable to an Ethernet converter and from there, it is sent to a Ratatosk server (inside the vessel) for the subsequent transmission in near real-time.

3 Dataset description

1 Hz resolution temperature data have been recorded from tuna vessels from March 2021 to March 2023, in the Northwestern Indian Ocean, providing a huge number of observations. The sensors are located at the vessels' hulls at around 7 m depth, thus providing temperature data

almost at the surface. To get rid of spurious observations, the data were quality controlled (QC) as explained in the following section.

4 Data QC, processing, and validation

Temperature data were QC based on the QARTOD manual of the Integrated Ocean Observing System (U.S. Integrated Ocean Observing System (IOOS), 2020), by 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. Bad values	Data with no temperature (i.e. not-a-number (NaN)) value were removed.
F2. 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.
F3. Location	Check for realistic locations: latitudes (longitudes) above or below 90° (180°) and -90° (-180°), respectively, were removed.
F4. Realistic values	Check for values that were not realistic: temperature values below or above 0°C and 40°C were removed.
F5. Spikes	Check for outliers found for values higher or lower than 0.15°C than the previous observation. Outliers were removed.
F6. Visual inspection	Visual inspection of the data series for ensuring that outliers were removed. Uncommon patterns that might remain in the data series were also checked.

After the QC, 20.63% of the data were removed. Note that few spikes were found and that most of the removed data corresponded to NaN values (Filter F1 in Table 1). Once the spurious data were removed, the remaining vessel data were compared against the SST_GLO_SST_L4_REP_OBSERVATIONS_010_011 product from Copernicus Marine Environment Monitoring Service (CMEMS) (https://doi.org/10.48670/moi-00168). This product provides daily reprocessed Level 4 sea surface temperature (SST) data measured from satellites on a 0.05° grid and it was produced by Met Office using the Operational and Ice Analysis (OSTIA) system (Donlon et al., 2012). Although this CMEMS product does not provide temporal high-resolution temperature data (compared to the 1 Hz data measured at the vessel), the comparisons make it possible to assess the quality of the vessel data in general terms.

In order to make the data comparable, vessel data were adapted by daily averaging temperature and position values. Then CMEMS data were interpolated to the vessels' positions and correlations, root mean square differences (RMSD) and scatterplots were computed. Two vessels were used for the comparisons (herein vessel 1 and vessel 2) with periods ranging from the beginning of the series of each vessel until 31 May 2022, which is the latest date when CMEMS data were available when the comparison was made.

The data series (Figure 1) show a good agreement between the CMEMS and vessels' data. The main seasonal patterns are detected in both comparisons.

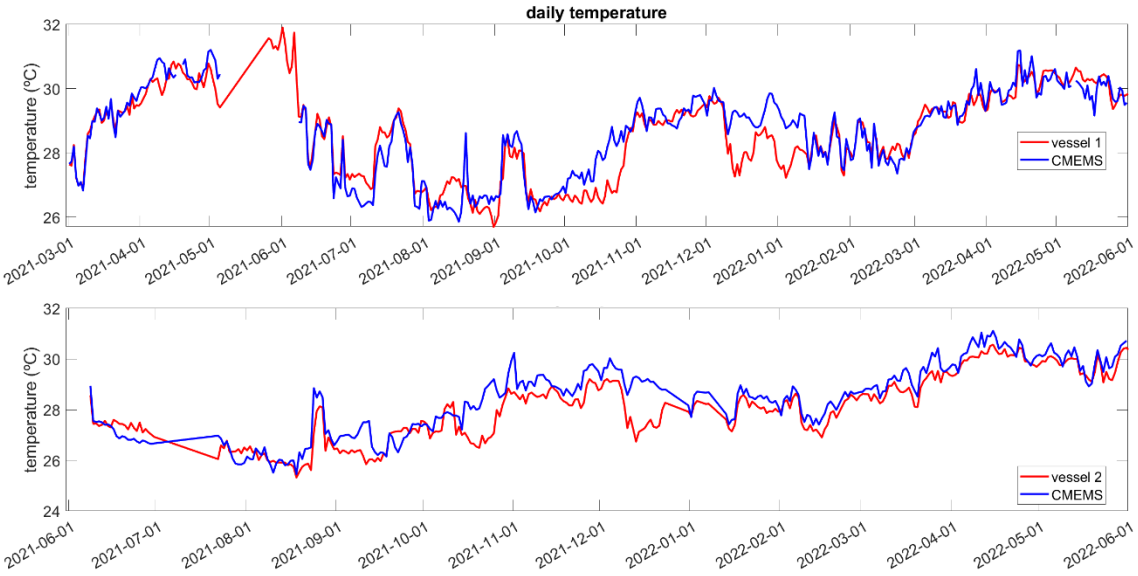

Figure 1. *Daily temperature (in °C) time series of vessels (red) and CMEMS (blue) for vessel 1 (top panel) and vessel 2 (bottom panel).*

The scatterplots of the time series (Figure 2) show that the clouds of points and the linear adjustments of the clouds are relatively well adjusted to the 1:1 isolines, thus representing the good agreement between CMEMS and vessel data. The slight overestimation of the temperature of CMEMS data is also noticeable. One of the possible sources of discrepancy between CMEMS and vessel data could be that the former corresponds to the surface whereas the latter to the near surface (at around 7 m depth), thus obtaining slightly higher temperatures for the CMEMS data. Another possible source is that the CMEMS data is an L4 product that has interpolated values when there were gaps in the SST satellite observations. In any case, the agreement is good, and the veracity of the vessel temperature data is demonstrated. It must be also considered that CMEMS data has its own error.

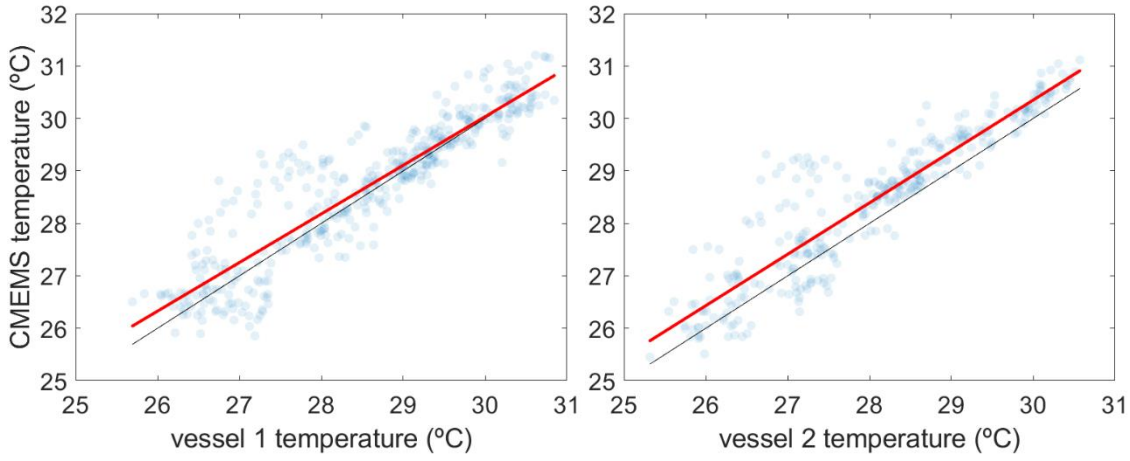

Figure 2. *Temperature comparisons between the vessel and CMEMS. For vessel 1 (left panel) and vessel 2 (right panel). The red line indicates the major axis regression model, and the black line indicates the 1:1 isoline.*

Moreover, the correlations show a good agreement (over 91 %) and the root mean square differences are not higher than 0.7°C (see Table 2). The slopes of the linear adjustment in Figure 2 are also very close to 1.

Table 2. *Correlation, RMSD and slope of the linear regression (shown in Figure 2) values of the comparisons between the temperatures of vessels 1 and 2 vs CMEMS.*

	Vessel 1 vs CMEMS	**Vessel 2 vs CMEMS**
Correlation (%)	91.01	91.94
RMSD (°C)	0.57	0.67
Slope	0.93	0.98

Since the temperature sensors are located at around 7 m under the water, this is the depth that is going to be considered for the datasets. For the final product, temperature measurements from all the vessels were temporally averaged and merged.

5 Product description

The temperature 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 temperature dataset.*

Spatial Coverage	Northwestern Indian Ocean
Spatial Resolution	Pointwise
Temporal Coverage	From March 2021 to March 2023
Temporal Resolution	0.5 hour
Variables	Temperature; 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 temperature dataset.*

Variable name	**Description**	**Units**
TIME	Date of the data	Days since 1950-01-01T00:00:00z
DEPH	Depth of the measurement	m
LATITUDE	Latitude of the data position	Degrees north
LONGITUDE	Longitude of the data position	Degrees east
TEMP	Sea water temperature	degree_Celsius

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 temperature 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	'Temperatures are obtained from the thermometers onboard the vessels. Raw data were quality controlled. The final product is an averaged set of temperatures at 7 m depth'
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	'Temperatures'
date_update	2023-03-09 23:30
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	'13.72'
geospatial_lat_min	'-5.60'
geospatial_lat_units	'degrees_north'
geospatial_lon_max	'69.47'
geospatial_lon_min	'44.78'
geospatial_lon_units	'degrees_east'
geospatial_vertical_max	'-7'
geospatial_vertical_min	'-7'
geospatial_vertical_units	'm'
history	2021-03-01 10:30 - 2023-03-09 23:30 data collected. 2023-07-12 23:34 netCDF file created using Matlab software
infoUrl	'https://www.sustuntech.eu/communicationmaterials/deliverables/'
institution	'AZTI (Spain)'
institution_edmo_code	'1623'
institution_references	'AZTI'
keywords	'TEMPERATURE, FISHERIES'
keywords_vocabulary	'GCMD Science Keywords'
last_update	2023-07-12 23:34
license	'This vessel mounted temperature 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 QC based on the QARTOD manual of the Integrated Ocean Observing System (IOOS, https://ioos.noaa.gov/ioos-in-action/temperature-salinity/) 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	‘Temperatures are obtained from the thermometers onboard the vessels. Raw data were quality controlled. The final product is an averaged set of temperatures at 7 m depth’
time_coverage_end	2023-03-09 23:30
time_coverage_resolution	‘0.5 hour’
time_coverage_start	2021-03-01 10:30
title	‘Vessels hourly temperatures at 7 m depth in the Northwestern Indian Ocean by SusTunTech project’
update_interval	‘void’
wmo_inst_type	‘ ‘
wmo_platform_code	‘ ‘

6 References

Donlon, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E., & Wimmer, W. (2012). The operational sea surface temperature and sea ice analysis (OSTIA) system. Remote Sensing of Environment, 116, 140-158.

U.S. Integrated Ocean Observing System (2020) Manual for Real-Time Quality Control of In-situ Temperature and Salinity Data Version 2.1: a Guide to Quality Control and Quality Assurance of In-situ Temperature and Salinity Observations. Silver Spring, MD, U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, Integrated Ocean Observing System, 50pp. DOI:10.25923/x02m-m555.