https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

Open Access Earth System Science Data Discussions

A multi-year dataset of integrated water vapor derived from shipborne GNSS observations collected aboard eight French research vessels during oceanographic campaigns (2015-2024)

Aurélie Panetier[1,2], Pierre Bosser[3], and Félix Mercier[3,4,5]

[1]Shom/DOPS/STM/ORCL, Brest, France
[2]Centre national d'études spatiales (CNES), Paris, France
[3]Lab-STICC/M[3] UMR 6285 CNRS, ENSTA, IP Paris, Brest, France
[4]ENSG, Marne-La-Vallée, France
[5]IGN, Saint-Mandé, France

Correspondence: Aurélie Panetier (panetier.aurelie@gmail.com)

Abstract. In the context of climate change and the growing need for improved observations of the atmospheric water cycle, measurements of atmospheric water vapor over the global ocean remain scarce compared with those available over land. This observational gap can be addressed using Integrated Water Vapor (IWV) derived from shipborne Global Navigation Satellite System (GNSS) observations, which provide a robust and well-established method for monitoring atmospheric moisture over the oceans.

This study presents a shipborne IWV dataset obtained from the processing of raw data collected by GNSS antennas installed on research vessels. The dataset benefits from substantial support and data access provided by Genavir, the operator of the French Oceanographic Fleet (FOF), and the Ifremer archive department SISMER. It is based on oceanographic campaigns conducted worldwide by eight vessels over a ten-year period (2015–2024), representing a total of 6,427 campaign days in both offshore and coastal regions.

After describing the methodology used to derive IWV from raw GNSS observations and to remove spurious measurements through a screening procedure, the dataset is evaluated through comparisons with the ERA5 reanalysis and satellite radiometer measurements from Remote Sensing Systems. These comparisons yield mean differences of $(0.3 \pm 2.0)\ \mathrm{kg\,m^{-2}}$ and $(-0.4 \pm 1.8)\ \mathrm{kg\,m^{-2}}$, respectively.

To further quantify the inherent uncertainty of the shipborne IWV retrieval, the dataset is cross-validated using instances where two vessels were within $50\ \mathrm{km}$ of each other. This comparison results in an estimated uncertainty of $0.96\ \mathrm{kg\,m^{-2}}$, demonstrating the suitability of the dataset for climate studies.

Local discrepancies identified in these comparisons are discussed, highlighting limitations in each dataset considered.

The GNSS-derived IWV dataset is intended to be updated annually to support long-term monitoring of atmospheric water vapor over the global ocean.

The IWV estimates are available at https://doi.org/10.25326/876 (Panetier and Bosser, 2026) through the AERIS data center (https://en.aeris-data.fr/, last access: 20 March 2026), and currently span the period from 2015 to the end of 2024.

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

Open Access Earth System Science Data Discussions

1 Introduction

The French Oceanographic Fleet (FOF) comprises a suite of research vessels used by the French scientific community for oceanographic campaigns (https://www.flotteoceanographique.fr/, last access: 5 March 2026). The fleet is managed by Genavir, a subsidiary of Ifremer, which operates the vessels and records the onboard scientific and navigation data during the campaigns. Shipborne GNSS raw observations are thus collected and managed by Genavir before being archived and distributed by the SISMER (Systèmes d'informations scientifiques de la Mer) data centre at Ifremer, ensuring data quality control and their long-term preservation and accessibility.

Integrated water vapor (IWV) is a critical variable for understanding atmospheric processes, energy fluxes, and the hydrological cycle, yet oceanic measurements remain scarce compared to terrestrial observations (Trenberth et al., 2005; Wang et al., 2020). Since the late 1990s, Global Navigation Satellite System (GNSS) has been commonly used for global monitoring of atmospheric water vapor. Indeed, IWV retrieval from GNSS observations is made possible through the estimation of tropospheric propagation delays affecting GNSS signals, which must be determined as part of the precise analysis of GNSS measurements (Bevis et al., 1992; Guerova et al., 2016). The tens of thousands of permanent stations deployed worldwide now constitute a dense network of sensors capable of retrieving IWV (Yuan et al., 2023). These data are used for various applications, including numerical weather prediction (Poli et al., 2007), numerical model assessment (Yuan et al., 2025), climatological studies (Bock et al., 2024), and dedicated field campaigns (Bock et al., 2021). Since the mid-2000s, this approach has progressively been extended to GNSS antennas installed on ships (Rocken et al., 2005; Wang et al., 2019; Bosser et al., 2021)

The technique is more sensitive in offshore environments because of the strong correlation between GNSS tropospheric delay and antenna height, which are estimated simultaneously at each epoch, unlike ground-based antennas where position is typically estimated daily. Early promising results demonstrated the feasibility of the method (Rocken et al., 2005), and it was progressively implemented during dedicated meteorological campaigns (Boniface et al., 2012). Intercomparisons with colocated or spatially distributed measurements, as well as with meteorological reanalysis products, have confirmed the quality of the retrieved IWV (Wu et al., 2022b). The operational and routine retrieval of these products is now feasible (Bosser et al., 2022), and their value for assimilation in numerical weather prediction systems has also been demonstrated (Ikuta et al., 2021). All these developments are particularly valuable as shipborne GNSS observations provide a unique source of atmospheric information over both coastal and open-ocean regions, where in situ observations of atmospheric water vapor remain sparse.

The dataset presented in this article provides IWV derived from shipborne GNSS observations archived by Ifremer since 2015. These opportunistic measurements, collected during FOF campaigns, represent a valuable source of water vapor observation over open-ocean regions. By processing, quality-controlling, and validating these observations, we provide a consistent long-term IWV record suitable for studies of marine atmospheric variability as well as for climate monitoring applications. The dataset is designed to be updated annually as new shipborne GNSS observations are archived, ensuring the progressive extension of the record and supporting long-term monitoring of atmospheric water vapor over the global ocean.

This article is organized as follows. Section 2 describes the shipborne GNSS dataset provided by Genavir, as well as the GNSS data processing, screening procedures, and IWV retrieval methodology. Section 3 presents statistical characteristics

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

Open Access Earth System Science Data Discussions

of the retrieved IWV values and their validation against external datasets, including IWV estimates from ERA5, the fifth reanalysis produced by the European Centre for Medium-Range Weather Forecasts, and satellite radiometer measurements provided by Remote Sensing Systems (RSS). It also includes internal comparisons taking advantage of situations where two vessels were navigating within 50 km to each other. In Section 4, we investigate situations exhibiting notable discrepancies between techniques and between vessels, which helps provide a more comprehensive characterization of the dataset. Finally, section 5 summarizes the main findings and the potential applications of the dataset.

2 Material and Methods

2.1 From Shipborne GNSS Observation to IWV

2.1.1 Spatial and Temporal Coverage of the GNSS Dataset

The dataset spans 2015–2024 and comprises observations collected aboard eight FOF research vessels. It includes all offshore vessels (L'Atalante, Marion Dufresne, Pourquoi Pas?, Thalassa), the semi-offshore vessels Alis and Antea, and two coastal vessels (L'Europe and Thalia). With the exception of the Marion Dufresne, GNSS raw data acquired onboard all vessels are recorded by Genavir and subsequently archived by Ifremer. The data is archived as hourly files in a manufacturer-specific binary format. On the Marion Dufresne, the observations are collected using a GNSS antenna specifically dedicated to atmospheric monitoring, deployed as part of the MAP-IO program (Tulet et al., 2024). This data is archived at an hourly sampling interval by the University of La Réunion. All the data was converted from binary format to the RINEX 3.03 format for the purposes of this study.

Campaign details are listed in supplementary material. Figure 1 shows the spatial coverage of all campaigns during this period. A zoom on Western Europe, where data density is high, is displayed in Appendix A. Coastal vessels operate near Marseille (L'Europe) and Brest (Thalia), semi-offshore vessels mainly in tropical areas, and offshore vessels worldwide, including Arctic (L'Atalante) and Antarctic (Marion Dufresne) missions. The Marion Dufresne has some specific operational characteristics. The vessel is based on Réunion Island in the Indian Ocean and carries out scientific missions in the Southern and Indian Oceans for approximately half of the year. During the remaining months, it performs logistical operations in support of the French Southern and Antarctic Lands (TAAF). Consequently, its navigation routes display recurring and relatively regular patterns during these logistical rotations.

Temporal availability is shown in Fig. 2, highlighting gaps due to campaign schedules, crew changes, weather, pandemics, and technical stops. Antenna changes over time are marked with red vertical bars. In 2015, only two vessels were subject to automatic data archiving. The number of vessels included in this system progressively increased over time, reaching seven vessels in 2024. Data acquisition on board the Alis ended in 2022 following the vessel's decommissioning.

Information on antennas and GNSS devices onboard each vessel is summarized in Table 1.

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

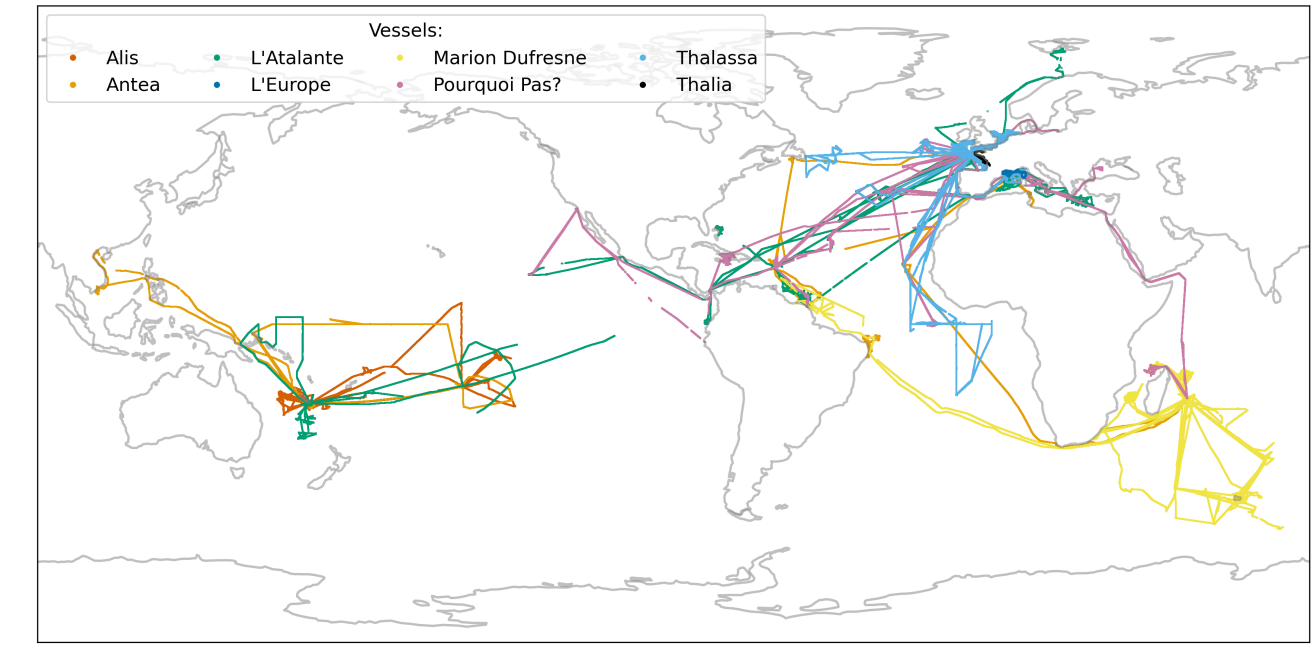


Figure 1. Map of the oceanographic campaigns conducted by the eight FOF research vessels between 2015 and 2024 using data provided by Ifremer. The vessels are distinguished by color as indicated in the legend. Because the data density is particularly high around Western Europe, a zoom of this region is shown in Fig. A1 in Appendix A.

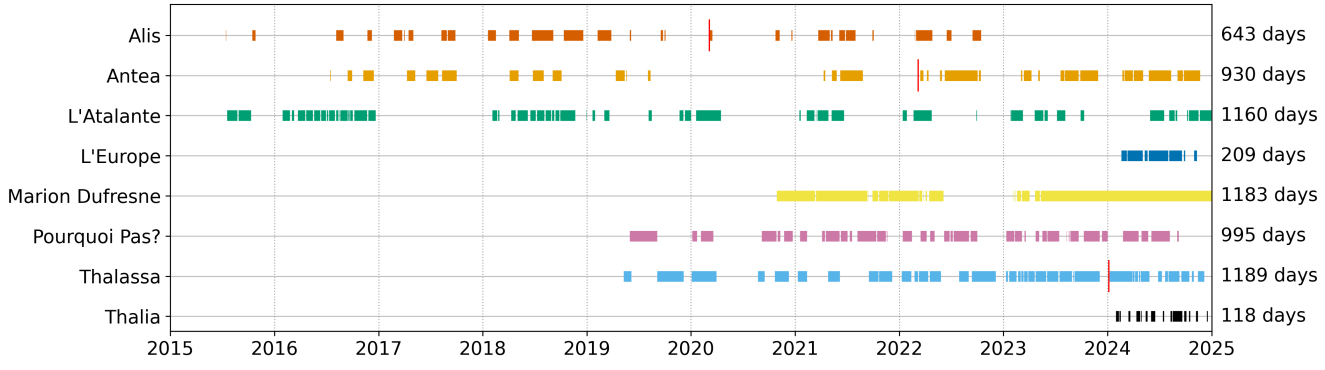


Figure 2. Availability of shipborne GNSS dataset for each vessel. Colors indicate periods with available data, while gaps correspond to missing observations. The larger red bars denote antenna replacements on the vessel mast, or antenna relocations in the case of the first two changes on Marion Dufresne. The total number of available days is indicated for each vessel on the right-hand side of the plot.

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

Open Access Earth System Science Data Discussions

Vessel	Type	Antenna	Receiver	Length	Height
Alis					
up to 2021	semi-offshore	AeroAntenna Technology AERAT1675_32	Ashtech ProFlex 800	29 m	13 m
from 2022		AeroAntenna Technology AERAT1675_182	Ashtech ProFlex 800		8 m
Antea					
up to 2021	semi-offshore	AeroAntenna Technology AERAT1675_182	Ashtech ProFlex 800	35 m	18 m
from 2022		Trimble TRM44830.00	Trimble BX992		19 m
L'Atalante	offshore	AeroAntenna Technology AERAT1675_32	Ashtech ProFlex 800	85 m	27 m
L'Europe	coastal	AeroAntenna Technology AERAT1675_32	Ashtech ProFlex 800	30 m	16 m
Marion Dufresne					
up to 03/2021	offshore	Trimble TRM55971_00	Trimble ALLOY	121 m	28 m
up to 11/2021					37 m
up to 04/2022					23 m
from 04/2022		Trimble TRM126000_00	Trimble ALLOY		23 m
Pourquoi Pas?	offshore	AeroAntenna Technology AERAT1675_182	Ashtech ProFlex 800	108 m	32 m
Thalassa					
up to 2023	offshore	AeroAntenna Technology AERAT1675_182	Ashtech ProFlex 800	74 m	29 m
from 2024		Trimble TRM44830.00	Trimble ALLOY		29 m
Thalia	coastal	AeroAntenna Technology AERAT1675_32	Ashtech ProFlex 800	25 m	8 m

Table 1. Characteristics of the GNSS navigation systems onboard each vessel. The columns display the name of the vessels, as well as the periods between antenna changes such as shown through the red bars in the timeline of Fig. 2; the type of research vessel, the antenna and receiver models, the lenght of the vessel and the medium height of the GNSS antenna above mean sea level for the whole period.

This table provides, for each vessel, the type of ship and its total length, and for each period between changes marked in red in Fig. 2, the antenna and receiver used, as well as the approximate antenna height above sea level. In general, the

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

GNSS antennas are installed on the crow's nest of each vessel, providing an unobstructed view of the GNSS constellations. On the Marion Dufresne, however, the antenna was relocated to the aft deck in 2021 after interference was observed following the deployment of nearby communication antennas (Wi-Fi and Iridium). As a result, the antenna now operates with a more restricted view of the sky on this vessel. The earliest installations on FOF vessels relied on Ashtech ProFlex 800 receivers, which output data in the manufacturer-specific ATOM format. These receivers have progressively been replaced by Trimble receivers producing observation files in the T02 or T04 formats. Except for the Marion Dufresne up to April 2022, all the ships are equipped with GNSS antennas designed for use in a marine environment.

2.1.2 GNSS Processing

GNSS data formatted in RINEX 3.03 were analyzed using Precise Point Positioning (PPP) with ambiguity resolution with the GipsyX 2.2 software, developed by the Jet Propulsion Laboratory (JPL) (Bertiger et al., 2020). Final JPL products for orbits, clocks, and widelane phase biases were used. The configuration of the receivers installed on the vessels currently permits only GPS observations to be recorded; therefore, only the GPS constellation was considered in the analysis. To minimize discontinuities at midnight, the processing time span was extended to 30 hours centered on noon. Kinematic processing employed a 3° cutoff angle and elevation-weighted observations proportional to the square root of the sine of elevation, such as recommanded in Panetier et al. (2023) for shipborne GNSS antennas. To avoid errors resulting from antennas not being oriented toward true north, antenna phase center variation (PCV) correction maps were averaged over azimuth.

Positions are estimated every 30 s without constraints between consecutive epochs. Zenith hydrostatic and wet tropospheric delays (ZHD and ZWD) are initially estimated using the VMF3 grids (Landskron and Böhm, 2017), and their elevation-dependent mapping is modeled with the VMF1 grids (Boehm et al., 2006), as GipsyX 2.2 does not support VMF3 mapping functions. ZWD corrections and horizontal gradients are estimated at a 30 s interval as stochastic parameters, with a random walk of 5 $\mathrm{mm\,h^{-0.5}}$ and 0.5 $\mathrm{mm\,h^{-0.5}}$, respectively.

The analysis is performed in two steps. First, an initial solution is computed at a 300 s temporal resolution to obtain a coarse trajectory of the antenna. This preliminary trajectory is then used to determine ZHD and ZWD along its path using the VMF3 grids. Finally, the coarse trajectory together with the ZHD and ZWD estimates are used in the final analysis performed at a 30 s temporal resolution. The zenith total delay (ZTD) is finally computed as the sum of the a priori ZHD and ZWD and the ZWD correction estimated during the analysis.

2.1.3 Data Screening

After analysis, the estimated ZTD was screened to remove spurious and outlying estimates, following principles in Bosser et al. (2022). This step is standard practice in the construction of climatological datasets and is crucial to guarantee their quality. For this purpose, the estimated ZTD, the associated formal errors, and the formal errors of the estimated antenna height are examined. Antenna height provide a valuable indicator of estimation quality at a given epoch. Indeed, ZTD and height are strongly correlated during the estimation, and height is more sensitive to poor-quality observations because successive positions are unconstrained. The screening is then performed in successive steps:

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

1. Height formal error range check: epochs where formal error exceeds 1 m are removed
2. ZTD range check: values outside 2–3 m are removed
3. ZTD formal error check: values exceeding 8 mm are removed
4. ZTD outlier check: values differing from the median ZTD of the corresponding vessel by more than 0.5 m are removed
5. ZTD formal error outlier check: values exceeding the median formal error plus three interquartile ranges, computed over the full time series of each vessel, are removed
6. Time variation check: successive ZTD estimates showing variations greater than $30\ \mathrm{mm\,h^{-1}}$ are removed

The threshold values are determined from the analysis of the distribution of the parameters used, for each vessel. While the screening process can remove a significant portion of the data and is not perfect, it remains efficient and fast for detecting spurious retrievals. The ZTD solutions are archived, enabling the dataset to be updated if the screening methodology is improved in the future.

An overall 3.8 % of the dataset is thus rejected from this screening procedure. Figure 3 summarizes the proportion of retained and rejected observations for each vessel after the screening procedure. With the exception of Thalia, final data rejection rates

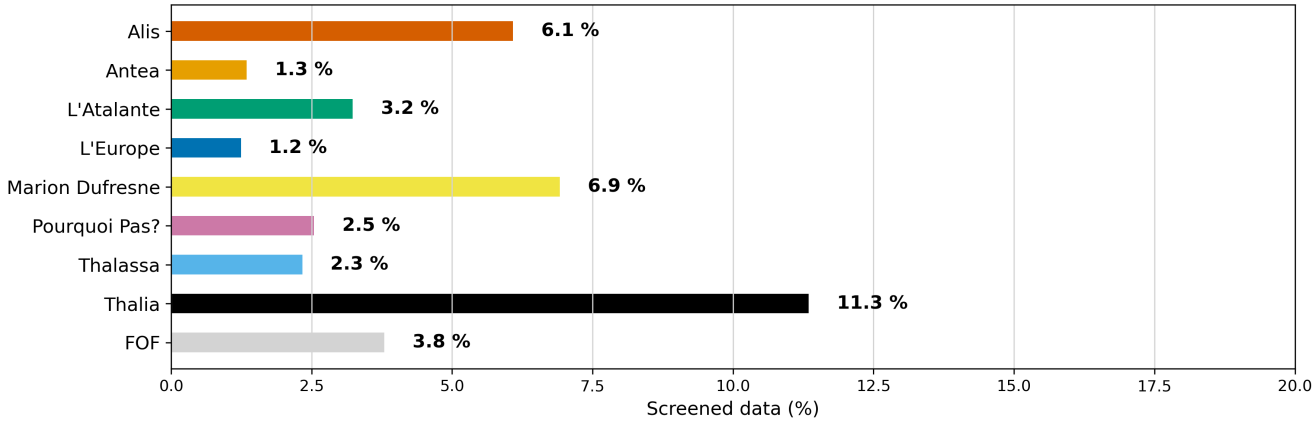


Figure 3. Summary of the screening results for each vessel. The last line shows the percentage of screened data for the overall FOF dataset. Bars show the proportion of rejected observations after the screening procedure. The percentages are written in front of each bar.

remain below 10 %, with the lowest values observed for the coastal vessel L'Europe (1.2 %) and the semi-offshore vessel Antea (1.3 %). The coastal vessel Thalia exhibits a higher rejection rate of 11.3 %, while Alis and Marion Dufresne both exceed 6 %, indicating comparatively lower data quality for these vessels. Marion Dufresne is known to present discrepancies, as its antenna has been repeatedly affected by nearby transmitters, and its placement on the aft deck since early 2022 is not optimal (Bosser and Tulet, 2024). For Alis and Thalia, GNSS data had not been used prior to this study. Consequently, the history of their

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

Open Access Earth System Science Data Discussions

antenna installations is unknown. Antenna placement is not always optimal and varies between vessels. Several factors may influence data quality, including vessel size, antenna characteristics, and surrounding obstructions or signal masking, which are often more pronounced on smaller vessels where antenna placement options are limited (see Table 1). For instance on Thalia in 2022, the antenna was installed on the upper bridge beneath an instrumented mast, which may cause interference and signal masking. These screening rates are somewhat higher than those typically reported for fixed terrestrial antennas (Bosser et al., 2021), likely reflecting more challenging acquisition conditions and the inherently greater complexity of GNSS processing for mobile platforms.

After screening, the dataset contains 17,775,961 observations distributed among the eight vessels as shown in Fig. 4. This distribution aligns with the data availability diagram presented in Fig. 2. The largest datasets are associated with the Marion Dufresne, Thalassa, and L'Atalante. For L'Atalante, this mainly reflects the long period of data availability (from 2015 onward), whereas for Marion Dufresne and Thalassa, it is largely due to a lower frequency of data interruptions since the beginning of their respective acquisition periods. Data have been collected only since early 2024 on Thalia and L'Europe, resulting in the smallest available datasets among the fleet.

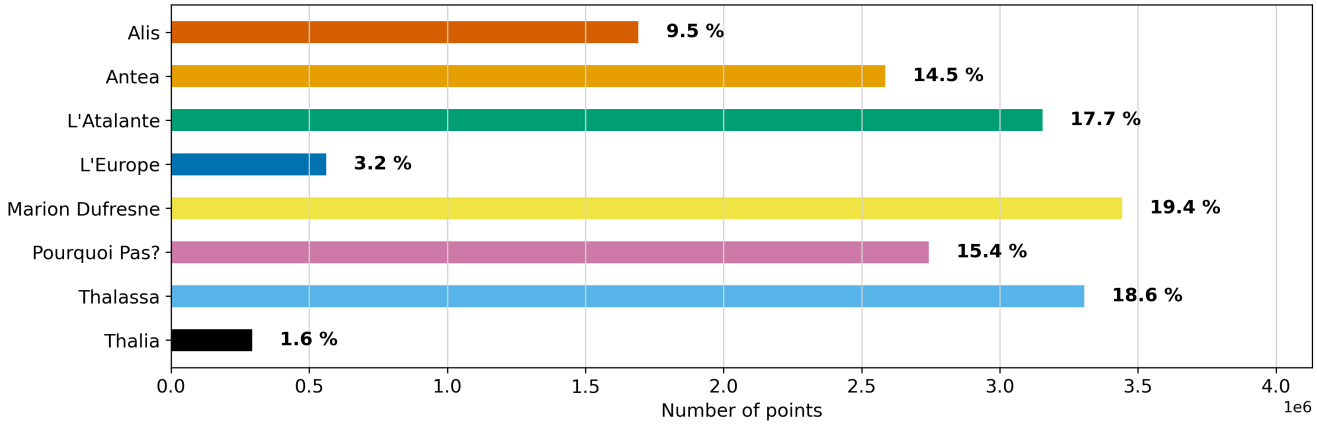


Figure 4. Distribution of data among research vessels after screening.

2.1.4 IWV Retrieval

IWV was retrieved from the ZWD estimated during the GNSS analysis. In this process, ZWD was obtained by subtracting the ZHD from the ZTD:

$$ZWD = ZTD - ZHD \tag{1}$$

The ZHD was computed using the Saastamoinen (1972) formula, updated by Bosser et al. (2007), with atmospheric pressure at the antenna level. Here, the atmospheric pressure was extracted from the ERA5 mean sea level pressure field (Hersbach et al., 2018). ZHD was initially extrapolated to the antenna height using the formula proposed by Steigenberger et al. (2009), such as

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

described in Bosser et al. (2021):

$$ZHD(h_{ANT}) = ZHD(h_{MSL}) - 10^{-6}\, k_1\, \frac{p_{MSL}}{T_{MSL}}\, \frac{g_{MSL}}{g_{atm}}\, (h_{ANT} - h_{MSL}) \tag{2}$$

where h_{MSL}, p_{MSL}, and T_{MSL} denote the mean sea surface height, pressure and temperature respectively, $g_{MSL} = 9.8026\ \mathrm{m\,s^{-2}}$ and $g_{ATN} = 9.7840\ \mathrm{m\,s^{-2}}$ denote the gravity at mean sea level and the approximated gravity of the atmosphere mass center [Saastamoinen (1972)], and $k_1 = 0.776452\ \mathrm{K\,Pa^{-1}}$ is a refractivity constant (Thayer, 1974; Bock et al., 2021). The EGM2008 geoid model is used to compute the height of the antenna above mean sea level.

ZHD was finally interpolated spatio-temporally to the GNSS antenna location and observation times.

Then, the IWV was derived from ZWD as:

$$IWV = \frac{10^6}{(\frac{k_3}{T_m} + k_2')\, R_v}\, ZWD \tag{3}$$

where T_m denotes the weighted mean temperature of the troposphere (Bevis et al., 1992). It was computed from the VMF1 grids provided by TU Wien and interpolated both spatially and temporally to the GNSS antenna locations and observation epochs (Boehm et al., 2006). We used the updated refractivity constants (Thayer, 1974) as specified in Bock et al. (2021): $k_2' = k_2 - k_1 \frac{R_d}{R_v} = 0.229\ \mathrm{K\,Pa^{-1}}$ with $k_2 = 71.2\ \mathrm{K\,Pa^{-1}}$, and $k_3 = 3752.0\ \mathrm{K^2\,Pa^{-1}}$; $R_d = 287.05\ \mathrm{J\,kg^{-1}\,K^{-1}}$ and $R_v = 461.5\ \mathrm{J\,kg^{-1}\,K^{-1}}$ are the specific gas constants of dry air and water vapor respectively.

Finally, IWV was extrapolated to the mean sea level to facilitate its use and comparison with other measurement techniques, as well as to ensure its global applicability, independent of the vessel. It was then vertically extrapolated to mean sea surface using vertical correction as described in Bock et al. (2005):

$$\Delta IWV = -4 \cdot 10^{-4}\, IWV\, \Delta h \tag{4}$$

2.2 Comparison Datasets

2.2.1 ERA5

ERA5 (Hersbach et al., 2018) provides a global gridded reanalysis of IWV on an hourly basis, at mean sea level. This dataset is available under the variable name "Total Column Water Vapor" (TCWV) and is provided in the form of a 0.25-degree-grid with a 1-hour time resolution.

To enable a direct comparison with the shipborne GNSS-derived IWV at mean sea level, the ERA5 IWV fields were linearly interpolated to the GNSS antenna horizontal location using the four nearest grid points (Bosser et al., 2021). From the shipborne IWV time series, only the measurements closest in time to each hourly ERA5 value were retained. This procedure produces an hourly comparison time series between shipborne IWV and ERA5 IWV at the exact vessel locations, enabling an accurate assessment of the dataset's consistency and representativeness.

Since the launch of the ERA5 reanalysis, its performance has been frequently highlighted, and it is commonly used for evaluating water vapor measurements and for intercomparison studies. (Chen et al., 2021) reports an overall root mean square (RMS) difference of $1.6\ \mathrm{kg\,m^{-2}}$ between worldwide GNSS reference station IWV and ERA5 in 2019. However, several studies

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

using shipborne GNSS have reported values up to twice as large at sea: larger deviations may be related to a potentially poorer estimation of GNSS-derived IWV from shipborne antennas, as well as limitations of the reanalysis over open ocean, where fewer observations are available for assimilation (Bosser et al., 2021; Gong et al., 2022; Männel et al., 2021; Wu et al., 2022b).

2.2.2 Satellite Radiometers

Remote Sensing Systems[1] (RSS) provides satellite radiometer observations from numerous satellites. The radiometer datasets available throughout the whole period from 2015 to 2024 are the Special Sensor Microwave Imager Sounder (SSMIS) instruments F16, F17 and F18 (Wentz et al., 2012), as well as the Global Precipitation Measurement (GPM) Microwave Imager (GMI) (Wentz et al., 2015) and the Advanced Microwave Scanning Radiometer 2 (AMSR-2) (Wentz et al., 2021), with a version 8.2 released in May 2021 for the latter that we use for the purpose of this study. Windsat satellite radiometer (Wentz et al., 2013) and the Special Sensor Microwave/Imager (SSM/I) instrument F15 (Wentz et al., 2012) are available at the beginning of the time series up to 2017 and 2018 respectively. We use all of these datasets available on RSS to qualify the shipborne-derived IWV by comparison.

The RSS products are provided on a 0.25-degree-grid for the morning and another for the afternoon, for each instrument. The observables on the grid are independently dated to the time of satellite pass. As far as each node of the grid represents the satellite measurement located in a box centered on this node, we then compare the vessel-derived IWV to the satellite products that are located within $0.25°$ of the vessel location and within 15 s to the vessel measurement.

The radiometer comparison dataset provides sparse data, especially for the coastal vessels that sail in areas where the radiometer cannot sense because of land contamination. However for the other vessels, several comparisons a day are available when they are not located in a harbor.

The IWV measurements provided by such microwave radiometers over Oceans are considered highly reliable (Mears et al., 2018). Previous studies (Wu et al., 2022b; Shoji et al., 2023) report RMS differences of approximately $2.2\ \mathrm{kg\,m^{-2}}$ between shipborne GNSS-derived and satellite microwave radiometer-derived IWV.

2.2.3 Vessel Crossing Comparison

During their respective cruises, vessels occasionally pass in close proximity to one another. When two vessels are within 50 km of each other, their IWV time series are directly compared to assess the retrieval accuracy, under the assumption that atmospheric conditions are effectively identical within this distance (Bosser et al., 2021). This intercomparison provides an independent evaluation of the precision and consistency of the shipborne GNSS-derived IWV dataset under nearly identical environmental conditions.

[1] www.remss.com, last access 5 March 2026

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

Open Access Earth System Science Data Discussions

3 Results

3.1 Comparison to ERA5

The confrontation of the vessel IWV to the ERA5 IWV led to a comparison dataset containing 143'092 points overall the whole fleet. Globally, the vessel IWV exhibits a small wet bias of $0.3\ \mathrm{kg\,m^{-2}}$ compared to ERA5, as shown in the histogram on Fig. 5.

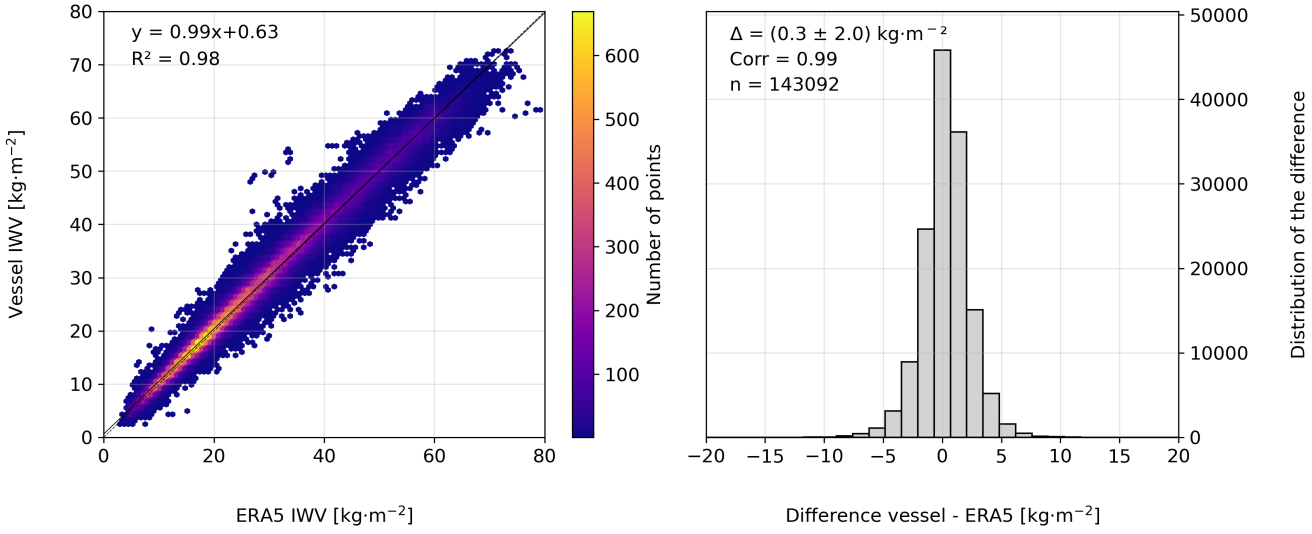


Figure 5. Scatter plot (left) and histogram of the differences (right) between IWV from shipborne GNSS antenna and ERA5. The density of points on the scatter plot is showed through the color of the points. The dashed black line on the scatter plot refers to the 1:1 line while the plain black line shows the linear regression.

The correlation is close to 1, indicating that both techniques consistently capture atmospheric temporal variations on an hourly timescale. The standard deviation of $2\ \mathrm{kg\,m^{-2}}$ is consistent with similar studies considering shipborne GNSS IWV retrieval (Boniface et al., 2012; Wang et al., 2019; Bosser et al., 2021; Männel et al., 2021; Wu et al., 2022b). The magnitude of the differences between GNSS-derived IWV and ERA5 is broadly consistent with uncertainty levels generally considered acceptable for climatological applications (Offiler, 2010), suggesting that GNSS observations may be suitable for characterizing water vapor variability in a climate context. The scatter plot on Fig. 5 shows strong agreement between the two datasets, with a regression slope of 0.99. RMS of residuals is limited to about $2.0\ \mathrm{kg\,m^{-2}}$, with a few outliers, which will be discussed later in Section 4. The distribution density is higher around $20\ \mathrm{kg\,m^{-2}}$ as a result of increased campaign density in temperate latitudes.

Figure 6 presents comparisons for each vessel: it is insightful to analyze the comparison for individual vessels, which differ in size, research purpose, research location, and antenna characteristics as presented in Table 1 and Fig. 1.

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

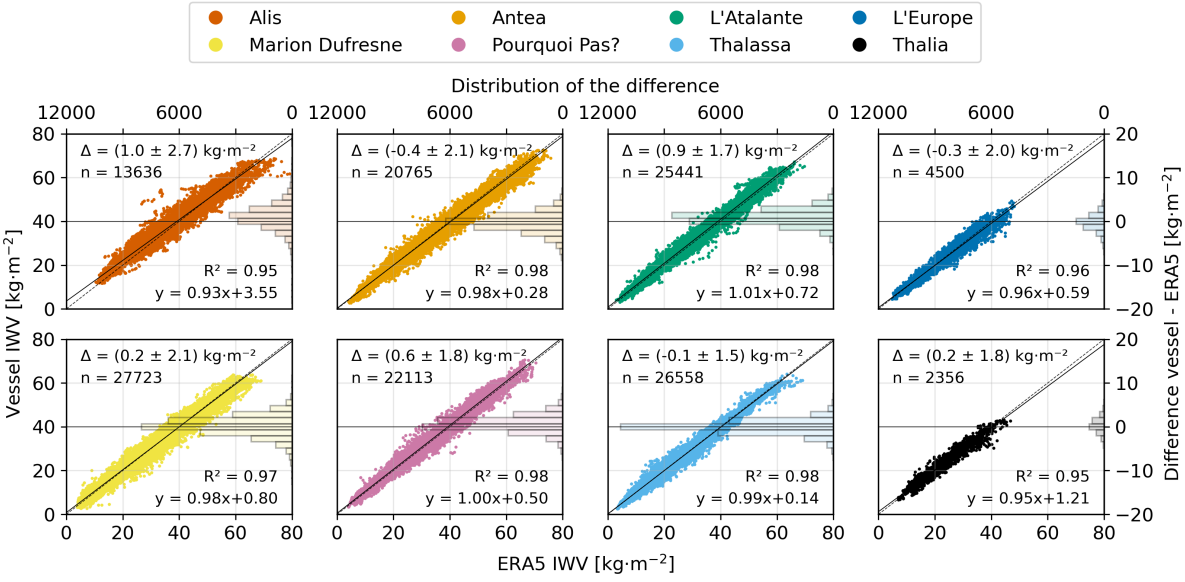


Figure 6. Scatter plot over histogram of differences between IWV from shipborne GNSS antenna and ERA5, per vessel.

From Fig. 6, we can notice that the large outliers observed in Fig. 5 are mainly associated with IWV measurements from Alis. Smaller irregularities are also visible for L'Atalante and Antea, although they remain less pronounced. Coastal vessels such as Thalia and L'Europe exhibit narrower IWV distributions, generally below $50\ \mathrm{kg\,m^{-2}}$, consistent with their regional operations along the French coasts. Wet and dry biases vary between vessels: Antea, L'Europe, and Thalassa show small dry biases, whereas Alis exhibits a larger wet bias and increased standard deviation. This is likely related to the lower data quality of Alis, as highlighted in Fig. 3, as well as its operations in the tropical Pacific. The latter reason will be further discussed in Section 4.

3.2 Comparison to RSS dataset

The vessel-to-radiometer comparison yields 21,860 points, almost seven times fewer than the ERA5 comparison due to satellite observation scarcity. Figure 7 presents the scatter plot and histogram of differences. The overall difference is $(-0.4 \pm 1.8)\ \mathrm{kg\,m^{-2}}$, with a correlation of 0.99, comparable to ERA5. The regression slope is slightly lower (0.96 vs 0.99), indicating that high IWV values are associated with an increasing dry bias. This effect, which shifts the distribution towards drier values, is discussed in Section 4.

The standard deviation of the differences is smaller than for ERA5 comparison (1.8 vs $2.0\ \mathrm{kg\,m^{-2}}$), likely because radiometer measurements are spatially and temporally collocated with the GNSS observations, in contrast to the spatial smoothing inherent to ERA5. Outliers, mainly occurring at high IWV, do not coincide with those identified in the ERA5 comparison and will also

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

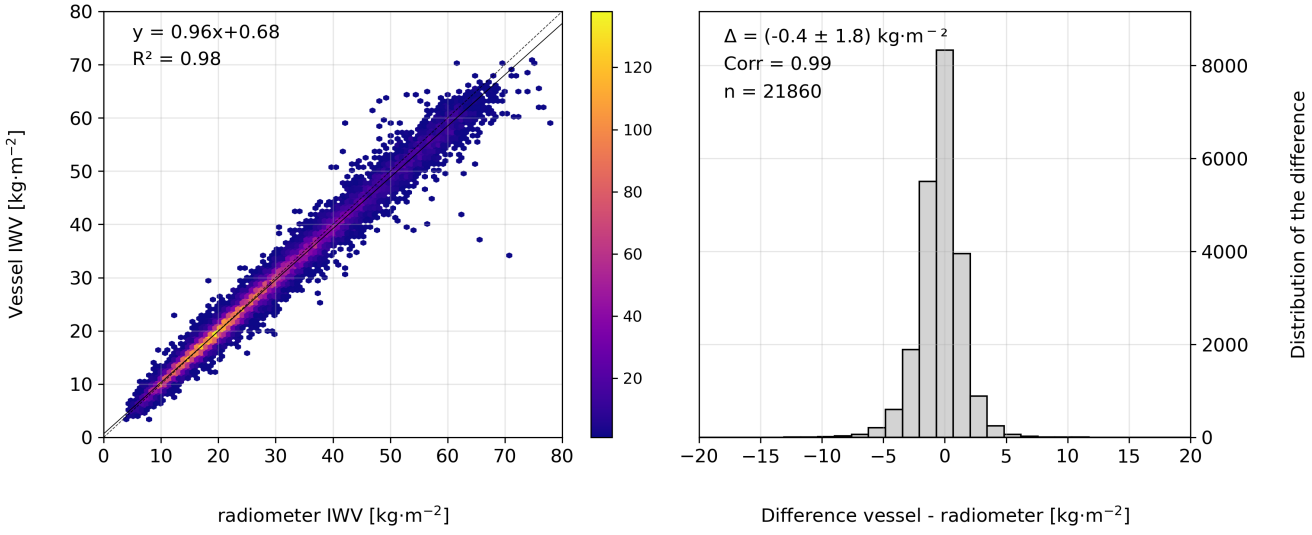


Figure 7. Scatter plot (left) and histogram of the differences (right) between IWV from shipborne GNSS antenna and RSS. The density of points is shown by color. The dashed black line represents the 1:1 line, and the solid black line shows the linear regression.

be addressed in Section 4. This suggests that the discrepancies in RSS comparison are unlikely to originate from the shipborne GNSS measurements, which were previously validated against ERA5 for these periods.

Figure 8 shows scatter plots and histograms per vessel. Coastal vessels L'Europe and Thalia provide relatively few comparison points due to land contamination affecting radiometer measurements within 50 km of the coast, as described in Section 2. This limitation also affects semi-offshore vessels such as Alis and Antea. Alis and L'Atalante exhibit overall wet biases, consistent with those observed in the ERA5 comparison, whereas Antea shows a stronger dry bias than in ERA5, indicating a slight shift in bias between the two comparisons. The standard deviations are slightly smaller than for ERA5, suggesting a better agreement with the radiometer measurements. Outliers are observed across nearly all vessels, in contrast to the ERA5 comparison where they were predominantly associated with Alis.

3.3 Vessel crossing comparisons

The consistency of IWV retrievals across the different vessels can be assessed when two vessels are located within 50 km of each other. Such inter-vessel comparisons provide an independent evaluation of the stability of the GNSS processing and the overall homogeneity of the dataset.

A total of 27 crossings were identified, of which 19 lasted longer than 4 hours. The corresponding time series are shown in Fig. 9 and indicate an overall good agreement between the vessels. The IWV absolute differences (displayed in grey) generally range between 0 and 2 $\mathrm{kg\,m^{-2}}$. The most noticeable disagreement occurs during the comparison between Marion Dufresne

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

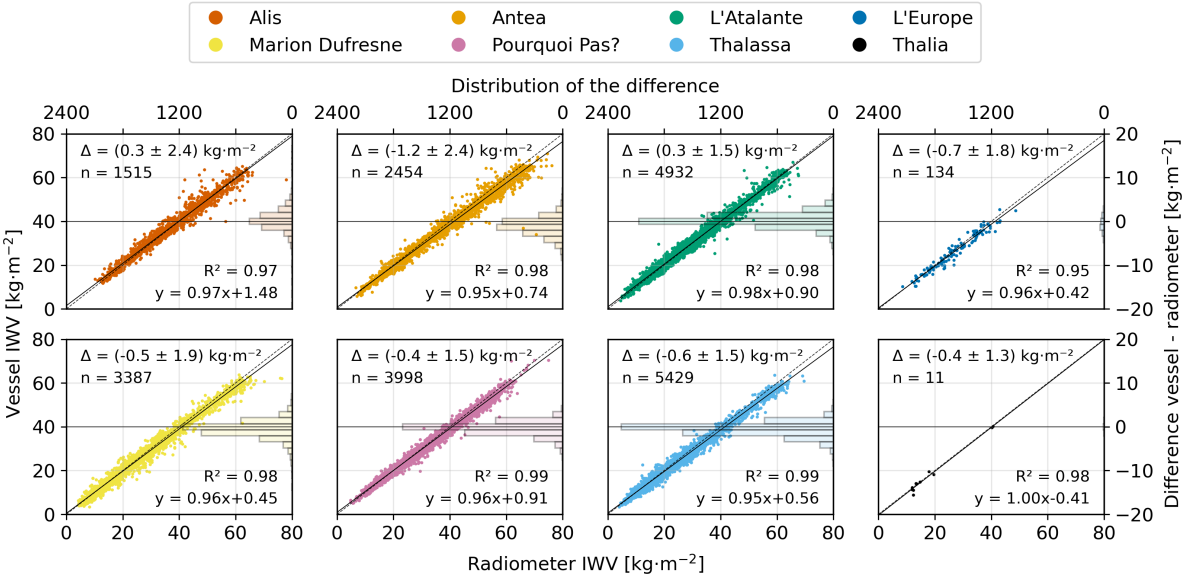


Figure 8. Scatter plot over histogram of differences between IWV from shipborne GNSS antenna and satellite radiometers, per vessel.

(yellow) and Pourquoi Pas? (pink) on 9–10 June 2021, with differences reaching up to 2 $\mathrm{kg\,m^{-2}}$, despite both vessels being co-located in the same harbor for approximately 24 h. This discrepancy is discussed in Section 4.

This analysis allows us to assess the self-consistency of the shipborne GNSS IWV retrieval. Across the 13 668 collocated measurements over the 27 vessel crossings, the mean RMS difference of the IWV is 0.96 $\mathrm{kg\,m^{-2}}$, indicating a good overall agreement between independent vessels.

4 Discussion

The comparisons presented in Section 3 reveal generally good agreement between vessel-derived IWV and the reference datasets, but also highlight a number of notable outliers. In this section, we investigate these discrepancies and explore possible explanations related to measurement conditions, atmospheric variability, and retrieval characteristics.

4.1 Outliers in the comparison to ERA5

4.1.1 Vessel-related measurement issues

Although the overall agreement between vessel-derived IWV and ERA5 is good, a number of outliers remain. In this subsection, we examine the most prominent cases in order to better understand their origin.

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

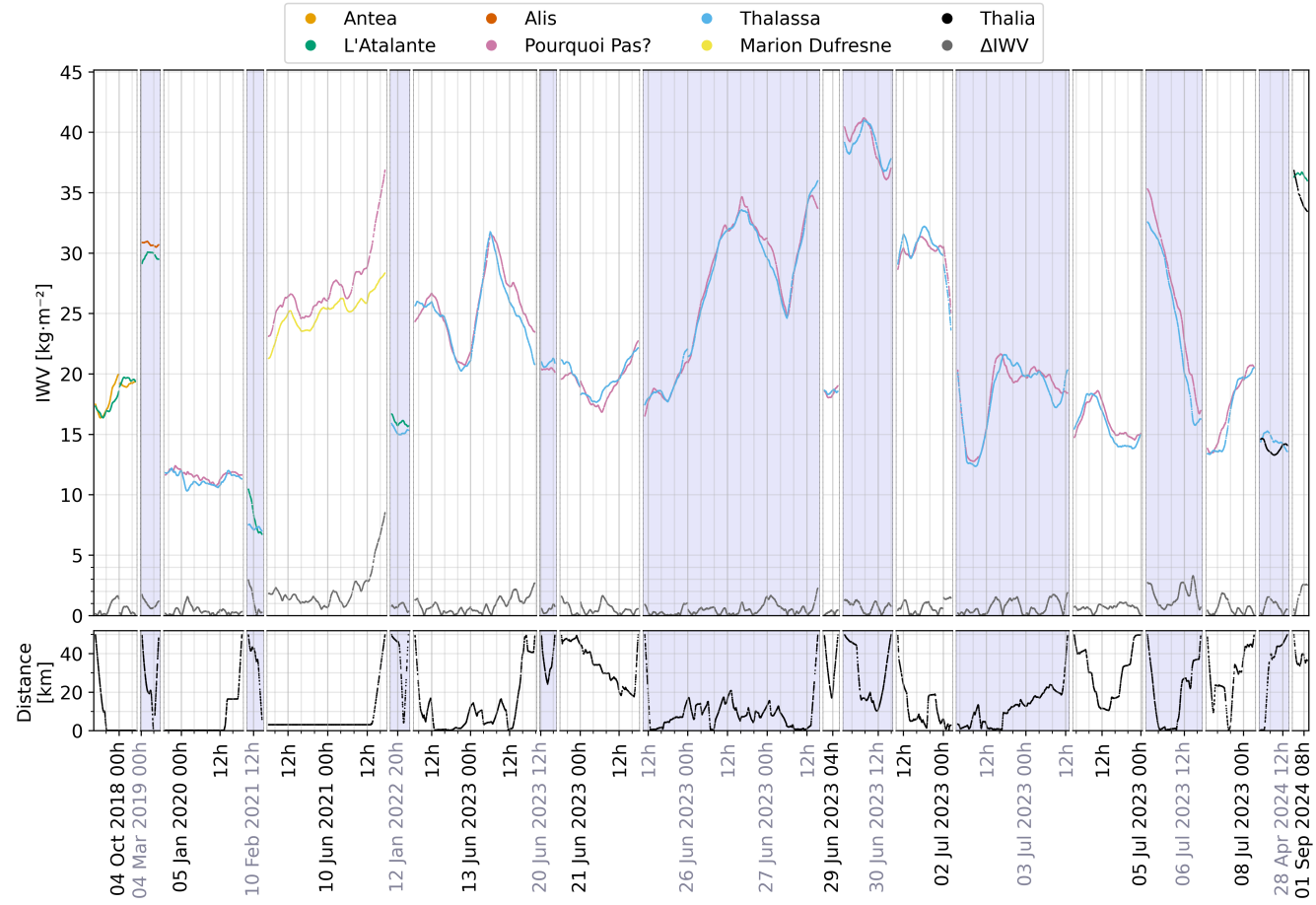


Figure 9. IWV time series for two crossing vessels (colored lines) and their difference (grey) when the vessels are within 50 km of each other. The lower panel shows the distance between the vessels during the comparison periods. Only the time series longer than 4 hours are represented here.

The two clusters of outliers that can be seen in the scatter plot in the left of Fig. 5 are due to ERA5 comparison of Alis-retrieved IWV.

The first cluster, located in the upper-left part of the scatter plot where ERA5 IWV values are much lower than the vessel-retrieved values by around 20 $\mathrm{kg\,m^{-2}}$, corresponds to measurements on 28 June 2021, when the vessel was sailing along the North coast of New Caledonia.

Both IWV vessel and ERA5 time series are shown together in Fig. 10 for the day of interest and the adjacent days. On the days before and after 28 June, both time series are relatively comparable. However, on 28 June, a jump of around 20 $\mathrm{kg\,m^{-2}}$ appears in the vessel-retrieved IWV, which is not captured by ERA5. While several hours of observations were filtered out around midnight during the vessel data screening, such a sudden increase in IWV within a few hours is unlikely. Although a

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

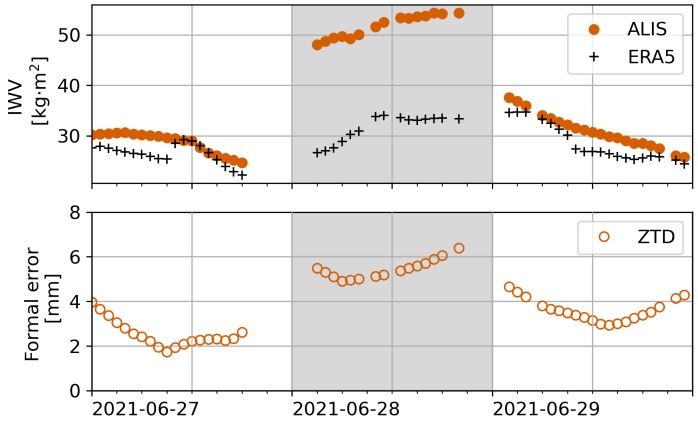


Figure 10. Investigation of the outlier cluster in the comparison of Alis-retrieved IWV with ERA5 on 28 June 2021 (shaded background) near New Caledonia. In the upper panel, orange markers denote the Alis measurements, while black markers depict the ERA5 interpolated values at the vessel location. In the lower panel, orange circles denote the ZTD formal errors from the GipsyX processing.

heavy rain event could in principle explain such a jump, meteorological archives from Noumea report no significant event on this day. Furthermore, June corresponds to the dry season in New Caledonia, as evidenced by the low IWV values on adjacent days, making strong precipitation events unlikely.

The formal error of the vessel ZTD, shown in the lower panel, is higher on this day, which may have contributed to the observed jump in the Alis time series. The screening ceiling of $8\ \mathrm{m}$ for the ZTD standard deviation was insufficient to remove this discrepancy, but lowering it to $4\ \mathrm{m}$ would have excluded roughly 20 % of the FOF dataset, which was deemed excessive.

4.1.2 Reanalysis spatial smoothing effect limitations

The second cluster of outliers appears in the upper-right part of the scatter plot of Fig. 5, where the vessel-retrieved IWV is much lower than ERA5, reaching nearly $80\ \mathrm{kg\,m^{-2}}$. This event occurred on 28 February 2022, during the cyclonic season (January–March), while the vessel was docked in Noumea harbor.

Figure 11 shows that the event occurs on the first day of a research campaign, when the antenna had just started acquiring data. On 28 February, ERA5 time series exhibits two sudden jumps: one at the start of the acquisition and another spanning several hours around midnight. In contrast, the vessel-retrieved IWV remains continuous and consistent. Unlike the first outlier cluster, these anomalies appear to be linked to the local variability in ERA5 rather than the vessel measurement. The interpolation of ERA5 values to the vessel position relies on the four nearest grid points, which in this case include both land and ocean points in an island region. This leads to pronounced instabilities in the interpolated ERA5 time series, as illustrated by the hourly IWV maps in Fig. 12. The surrounding ocean shows highly variable values, while the atmosphere above the island remains relatively stable except during the sudden jumps. The phenomenon described above is corroborated by comparison with the IWV retrieved from nearby GNSS ground station NRMD00NCL, which displays variations similar to those of Alis.

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

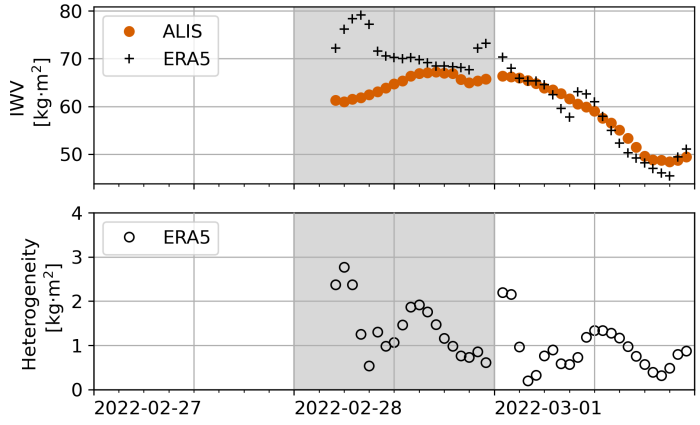


Figure 11. Investigation of the outlier cluster in the comparison of Alis-retrieved IWV with ERA5 on 28 February 2022 (shaded background) in Noumea harbor, New Caledonia. Orange markers denote the Alis measurements, while black markers represent ERA5 interpolated values at the vessel location. In the lower panel, black circles denote the heterogeneity among the four closest ERA5 grid points used for the interpolation.

Local meteorological archives indicate that 28 February 2022 was marked by difficult weather conditions in Noumea[2], with mainly cloudy skies, high humidity (around 85 %), and light intermittent showers (around 2 mm in total). Although conditions were not extreme, the unsettled weather could have contributed to local atmospheric variability that is challenging to capture with ERA5, particularly when interpolating between land and ocean grid points on a small mountainous island. This highlights the difficulty of capturing fine-scale variability in ERA5, especially when interpolating between ocean and land grid points in small insular areas.

Overall, this case illustrates the limitations encountered in both GNSS and ERA5 IWV retrievals in island and coastal regions over the open ocean. It emphasizes the impact of local variability, fine-scale meteorological conditions, and grid interpolation on ERA5, particularly in small insular areas such as Noumea harbor.

4.1.3 Latitude dependency in the ERA5 comparison

In Section 3.1, Alis shows weaker agreement with ERA5. In addition to its lower data quality pointed out in Fig. 3 and 6, this vessel operated exclusively in the tropical Pacific Ocean during 2015–2022, while the others are also sailing in mid-latitudes. Figure 13 presents boxplots of the FOF dataset comparison to ERA5 for three latitude bands: 0–30°, 30–60°, and above 60° (the latter containing very limited data from a single visit by L'Atalante). While no distinction is made between hemispheres here, the differences between tropical and mid-latitude statistics are highlighted. In particular, the standard deviation is approximately 50 % larger at low latitudes than at mid-latitudes. The GNSS–ERA5 differences also exhibit a marked latitudinal dependence, with larger deviations in tropical and polar regions compared to mid-latitudes. In this dataset, ERA5 is consistently drier than

[2]https://www.historique-meteo.net/oceanie/nouvelle-caledonie/noumea/2022/02/28/, last access 6 March 2026

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

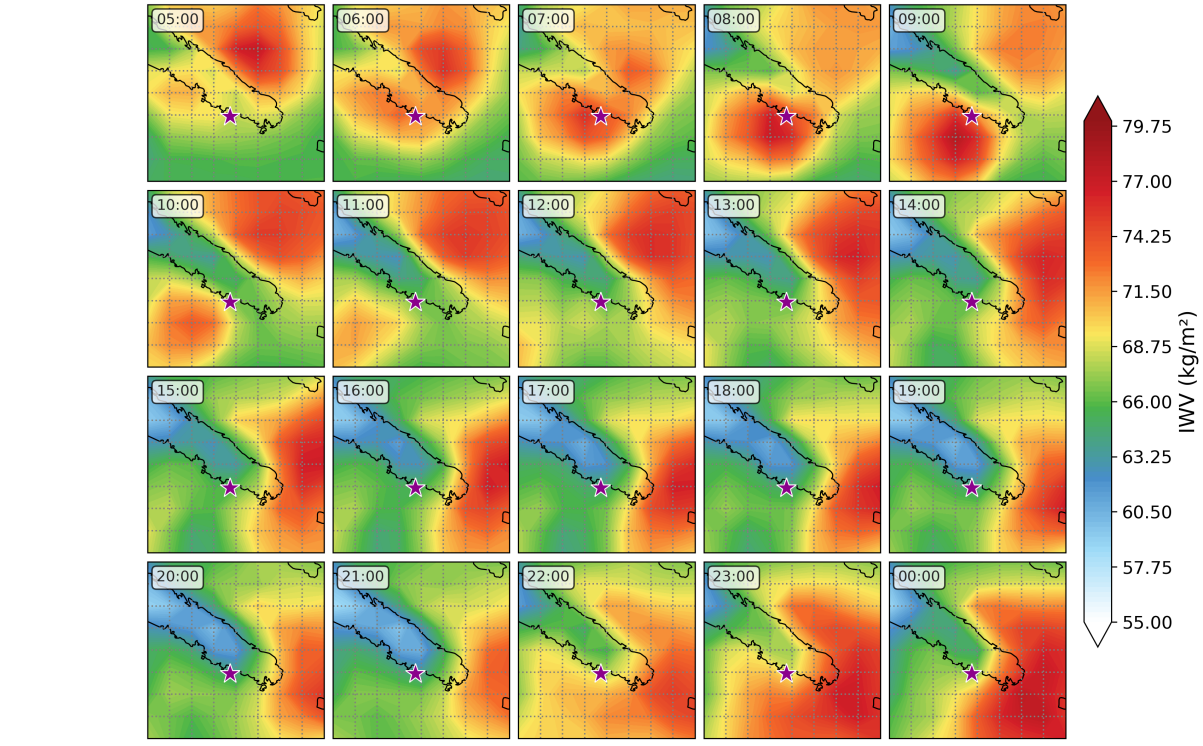


Figure 12. Hourly ERA5 IWV around New Caledonia on 28 February 2022. The maps show a small convective cell moving from east to west across the ocean surrounding the island. ERA5 interpolated values at the vessel location in Noumea harbor remain relatively moderate (60 $\mathrm{kg\,m^{-2}}$) most of the time, except during sudden jumps highlighted in the time series. The red star denotes the location of Alis in Noumea harbor during the whole day. The ERA5 grid is shown through the grey lines.

the GNSS-retrieved IWV, with a mean difference of about $0.52\ \mathrm{kg \cdot m^{-2}}$ in the tropics and $0.12\ \mathrm{kg \cdot m^{-2}}$ at mid-latitudes. Although the absolute sign of the bias differs from previous comparisons against ground-based GNSS stations (Bosser and Bock, 2021), where a dry bias was reported in the tropics and a wet bias at mid-latitudes, the pattern of drier reanalysis biases at low latitudes than at mid-latitudes is consistent between the two studies. A detailed vessel-by-vessel latitude-dependent analysis is also provided in Appendix C, which confirms that this increase in differences is consistently observed across all vessels. Despite this, Alis still exhibits the largest standard deviation among all vessels at low latitudes ($2.7\ \mathrm{kg\,m^{-2}}$), which may also reflect its comparatively lower data quality identified during the screening process. However, although it also suffered high rejection rate during the screening process, Marion Dufresne does not show such comparatively higher standard deviation at low latitudes ($2.2\ \mathrm{kg\,m^{-2}}$).

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

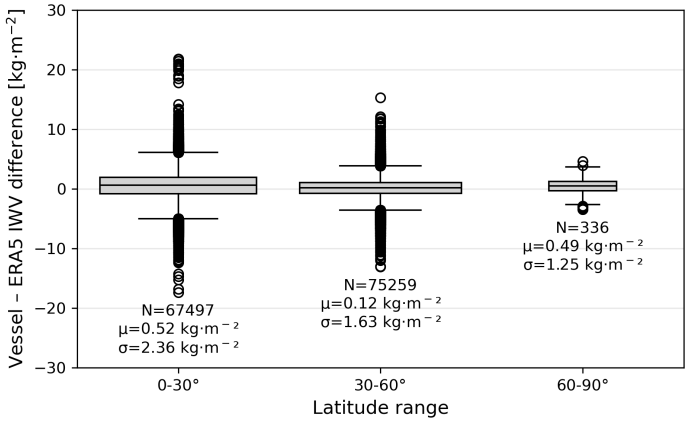


Figure 13. Differences between GNSS-derived IWV and ERA5 reanalysis as a function of latitude. Boxes represent the interquartile range (IQR, 25th–75th percentiles), with the horizontal line indicating the median. Whiskers extend to the most extreme values within 1.5 × IQR, and circles denote outliers. Box widths are scaled according to the sample size (0.3 + 0.5 × n), such that wider boxes correspond to larger numbers of observations.

4.2 Cloud cover dependency in the comparison to RSS

Analysis of the largest outliers indicates that they predominantly occur in comparisons with AMSR-2 and F15. Among the 14 largest outliers, nearly half involve AMSR-2, which is generally considered reliable. Comparing the vessel differences to ERA5 and the radiometer differences to ERA5 for these events suggests that ERA5 agrees more closely with the vessel IWV than AMSR-2 does (not shown). Based on this observation, we studied the vessel–RSS differences separately for each radiometer.

We examined how vessel–radiometer IWV differences depend on several environmental factors, including ERA5 significant wave height, radiometer precipitation, and cloud cover. Only cloud cover showed a consistent relationship with the differences, while wave height and precipitation explained little of the largest outliers. We therefore focus on the link between IWV differences and cloud cover.

Figure 14 shows the vessel–radiometer IWV differences as a function of cloud cover, expressed as "total cloud liquid water" in the RSS dataset, for each radiometer. Consistent with the global bias shown previously in Fig. 5, the vessel exhibits a dry bias of $-0.4\ \mathrm{kg\,m^{-2}}$ relative to three radiometers (F17, F18, F35). Smaller dry biases are observed for comparisons with AMSR-2 and WindSat, whereas F15 shows a pronounced wet bias of $2.4\ \mathrm{kg\,m^{-2}}$. The F16 comparison reveals a dry bias more than twice the global average of $-0.4\ \mathrm{kg\,m^{-2}}$, likely due to retrieval degradation of F16 following 2009[3], which affects its water vapor products.

Beyond the overall mean biases discussed above, distinct cloud-dependent patterns emerge for comparisons with F15, WindSat, and AMSR-2. Satellite F15 presents a dry bias with respect to shipborne GNSS for nearly all comparison points. This dry

[3]https://www.remss.com/support/known-issues, last access 17 March 2026

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

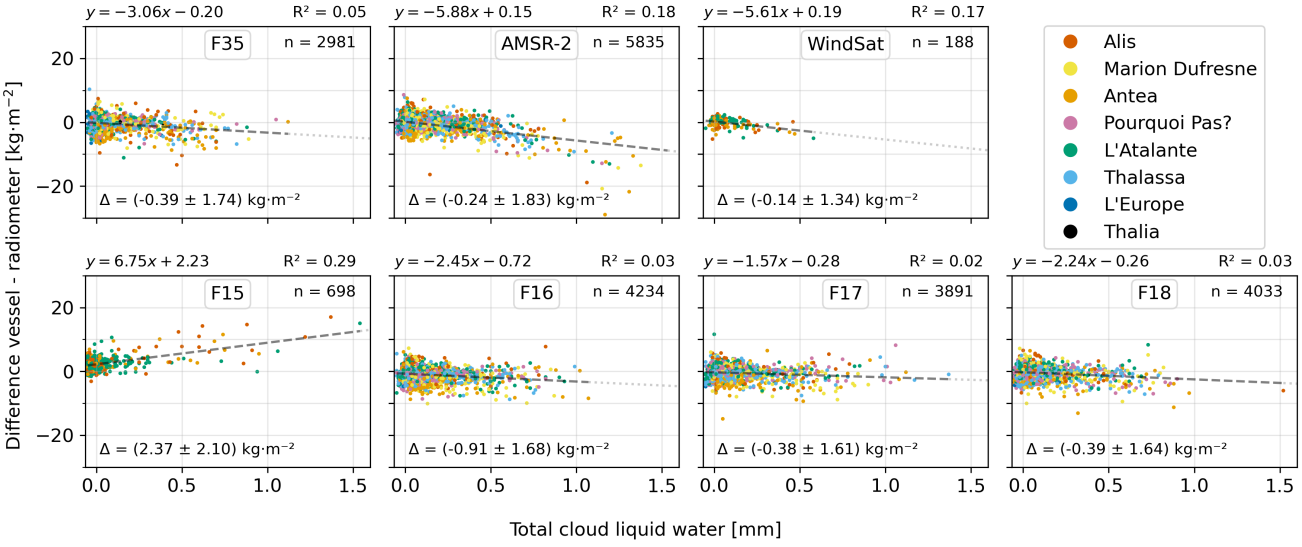


Figure 14. Vessel–radiometer IWV differences relative to the RSS dataset, shown as a function of total cloud liquid water measured by the respective radiometer, plotted separately for each instrument. The dashed line indicates the linear regression, with the corresponding equation, coefficient of determination (R^2), and number of points (n) displayed at the top of each panel.

bias increases with cloud cover, consistent with the known limitations of F15 for climate studies post-2006[4]. In contrast, AMSR-2 and WindSat exhibit a stronger wet tendency with increasing cloud cover when compared to shipborne GNSS. Comparisons with other radiometers, including F16, F17, F18, and F35, reveal smaller cloud-dependent wet biases with the vessel. No similar cloud-related patterns were associated with comparisons involving F15 up to 2015, thus before the period covered by our study (Mears et al., 2015). Such cloud-dependent patterns have been observed with the AMSR-E satellite, the predecessor of AMSR-2, confirming that cloud effects can introduce systematic biases in microwave satellite retrievals of IWV, which must be considered when interpreting satellite-based IWV data.

A linear regression confirms an overall tendency for the radiometer to report higher IWV with increasing cloud cover, except for F15, where the trend is reversed. The slope of the cloud dependence is most pronounced for AMSR-2, F15, and Windsat, nearly double that of the other radiometers. The coefficient of determination (R^2) for these three radiometers indicates that cloud fraction explains a substantial portion of the vessel–radiometer IWV differences (0.17–0.29), compared to less than 0.05 for the others.

These results underscore that cloud cover can systematically influence vessel–radiometer IWV differences, revealing biases that are radiometer-dependent. In particular, AMSR-2, while designed for atmospheric water vapor retrievals and generally validated against radiosondes and GPS (around 3 $\mathrm{kg\,m^{-2}}$ uncertainty, (JAXA)), exhibits a clear cloud-dependent bias in our

[4]https://www.remss.com/support/known-issues, last access 17 March 2026

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

dataset. This behavior is consistent with documented challenges in microwave retrievals under cloudy or unstable atmospheric conditions, due to the difficulty of separating water vapor and cloud liquid water signals (Ohara et al., 2023).

These findings highlight the value of large, independent in-situ datasets, such as GNSS measurements from ships, to identify and quantify subtle satellite retrieval biases under realistic atmospheric conditions.

4.3 Discrepancy in the IWV comparison between Marion Dufresne and Pourquoi Pas?

In contrast to the other inter-vessel comparisons presented in Fig. 9, the IWV time series from Marion Dufresne and Pourquoi Pas? show a noticeably larger disagreement. This is particularly unexpected because both vessels were located in the same harbor in Le Port, La Réunion, during this period, implying that they were exposed to nearly identical atmospheric conditions. Under such circumstances, very similar IWV estimates would normally be expected.

This situation therefore provides an interesting case study to investigate potential sources of discrepancy in vessel-based GNSS retrievals, including local environmental effects, platform-specific characteristics, or processing-related factors.

First, the geographical setting of the harbor is noteworthy because of the nearby mountainous terrain. Such topography could potentially lead to a misestimation of the zenith hydrostatic delay (ZHD) derived from ERA5 if one of the four surrounding ERA5 grid points used for the correction of the GNSS-derived zenith total delay (ZTD) were located over elevated terrain for only one of the vessels. However, both vessels fall within the same ERA5 grid cell and exhibit high ZHD homogeneity, suggesting that this effect is unlikely to explain the observed discrepancy.

Second, the ZTD values estimated through the PPP processing of the two time series during this period appear to differ significantly, even though the formal errors on the estimated positions and ZTD remain within their expected ranges. This suggests that the discrepancy originates from the GNSS-estimated ZTD themselves. Since shipborne GNSS-based IWV retrievals are typically expected to reach an accuracy of about $2\ \mathrm{kg\,m^{-2}}$ (Wang et al., 2019), this situation may represent a case where the error approaches the upper bound of the expected uncertainty. Over the period when both vessels were docked, the mean IWV difference reaches $1.61\ \mathrm{kg\,m^{-2}}$, with maximum differences up to $3.36\ \mathrm{kg\,m^{-2}}$.

Finally, both IWV time series were compared with ERA5 in Fig. 15 in order to assess the quality of the vessel-derived IWV estimates through an external reference. Marion Dufresne appears to follow the ERA5 values more closely, whereas Pourquoi Pas? systematically shows higher IWV values. This suggests that the ZTD estimation for Pourquoi Pas? may be less reliable during this specific period.

Overall, this case highlights that occasional discrepancies may arise in vessel-based GNSS IWV retrievals, even under nearly identical atmospheric conditions. However, this situation remains an outlier within the dataset, as most inter-vessel comparisons exhibit substantially better agreement even when involving Pourquoi Pas?, supporting the robustness of the retrieval approach.

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

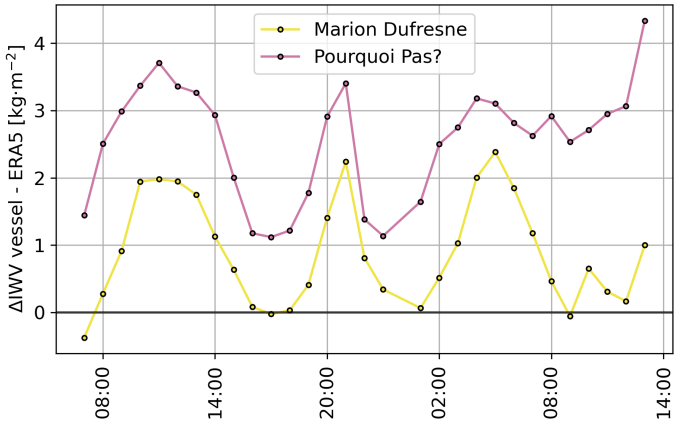


Figure 15. Time series of the differences between vessel-derived IWV and ERA5 for Marion Dufresne (yellow) and Pourquoi Pas? (pink) during 9–10 June 2021, when both vessels were docked in Le Port, La Réunion.

5 Conclusions

This work presents a comprehensive GNSS-based IWV dataset derived from PPP processing of antennas onboard eight research vessels since 2015. The dataset, covering all the world's oceans, represents a substantial record of IWV over maritime regions and is being provided as a valuable resource for the scientific community.

Comparisons with external datasets, including ERA5 reanalysis and satellite radiometers from RSS, generally show good agreement with mean differences to the reference of $(0.3 \pm 2.0)\ \mathrm{kg\,m^{-2}}$ and $(-0.4 \pm 1.8)\ \mathrm{kg\,m^{-2}}$ respectively. However, some isolated discrepancies highlight limitations in shipborne GNSS, reanalysis or satellite products, especially in regions with complex local variability, such as coastal and insular areas. Notably, satellite radiometer comparisons reveal cloud-dependent biases for certain instruments, particularly F15, AMSR-2 and WindSat, demonstrating the potential of vessel-based GNSS IWV to detect and quantify such effects. Inter-vessel comparisons further confirm the internal consistency of the dataset. Across 13,668 collocated measurements, the mean RMS IWV difference is $0.96\ \mathrm{kg\,m^{-2}}$, illustrating the reliability of GNSS-derived IWV across different platforms. Outlier cases, such as the comparison between Marion Dufresne and Pourquoi Pas? in Le Port, La Réunion, illustrate that occasional discrepancies can occur in GNSS ZTD estimation, leading to misestimated IWV.

Overall, this dataset provides a unique, high-quality record of atmospheric water vapor over the oceans. By making it available to the research community through the Aeris Data Center, it can support validation of satellite products, meteorological and climate reanalyses. It could facilitate investigations into cloud- and weather-dependent biases and lead to a better understanding of atmospheric water vapor variability across diverse maritime environments. Beyond these applications, the dataset also offers opportunities for cross-disciplinary studies, long-term monitoring efforts, and the development of new retrieval or assimilation methodologies, thereby contributing to more robust and reliable ocean–atmosphere observations.

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

This dataset is intended to be regularly updated as new data are acquired by vessels of FOF, ensuring its continuity and progressive enrichment over time. It may also be extended to additional platforms, including both scientific and non-scientific vessels, in order to broaden its spatial and temporal coverage. In this context, the emergence of low-cost GNSS devices, whose performance for atmospheric water vapor retrieval has been demonstrated (Wu et al., 2022a; Bosser et al., 2024), represents a promising opportunity to further enhance ocean observation, particularly through the development of participatory (citizen science) research initiatives, which could contribute to increasing data availability over poorly sampled oceanic regions.

6 Data availability

ERA5 data (Hersbach et al., 2018) was downloaded from the Copernicus Climate Change Service (2023). The results contain modified Copernicus Climate Change Service information 2020. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains.

SSM/I, SSMIS, GMI, WindSat and AMSR-2 data are produced by Remote Sensing Systems and sponsored by the NASA Earth Science MEaSUREs DISCOVER Project and the NASA Earth Science Physical Oceanography Program. RSS data from SSM/I, SSMIS, GMI, WindSat and AMSR-2 data are available at www.remss.com.

The full campaign list is provided as supplementary material. More information on individual campaigns is available on https://campagnes.flotteoceanographique.fr/campaign (last access: 20 March 2026).

The screened IWV estimates are available for download at the AERIS website: https://doi.org//10.25326/876 (Panetier and Bosser, 2026).

Author contributions. AP performed data curation, developed the GNSS post-processing chain, conducted the analyses, and wrote the original draft and figures. PB developed GNSS processing chain tools with GipsyX, managed the database, and extracted ERA5 data. FM contributed to updating the processing chain and handled radiometer data. AP and FM performed the validation against external datasets, and AP carried out the inter-vessel comparisons. AP and PB contributed to interpretation and manuscript revision. Supervision was provided by PB and AP.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. AP was supported by CNES through a postdoctoral fellowship. This work builds upon AP's PhD research. The authors thank SISMER (Ifremer), and in particular Cécile Pertuisot, for providing access to the shipborne GNSS data. ERA5 data were obtained from the Copernicus Climate Data Store, and radiometer data were provided by Remote Sensing Systems (RSS). The authors also acknowledge Shom for hosting AP during her postdoctorate. Python code used to standardize and generate figures was developed with the assistance of Large Language Models. Large Language Models were also used to improve the English. All code and scientific content were

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

reviewed, verified, and finalized by the authors. This work has been supported by the Programme National de Télédétection Spatiale (PNTS, https://programmes.insu.cnrs.fr/pnts/description-du-programme/), grant n° PNTS-2026-01.

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

References

Bertiger, W., Bar-Sever, Y., Dorsey, A., Haines, B., Harvey, N., Hemberger, D., Heflin, M., Lu, W., Miller, M., Moore, A. W., Murphy, D., Ries, P., Romans, L., Sibois, A., Sibthorpe, A., Szilagyi, B., Vallisneri, M., and Willis, P.: GipsyX/RTGx, a new tool set for space geodetic operations and research, Advances in Space Research, 66, 469–489, https://doi.org/10.1016/j.asr.2020.04.015, 2020.

Bevis, M., Bussinger, S., Herring, T. A., Rocken, C., Anthes, R. A., and Ware, R. H.: GPS Meteorology: Remote Sensing of Atmospheric Water Vapor Using the Global Positioning System, Journal of Geophysical Research: Atmospheres, 97, 15 787–15 801, https://doi.org/10.1029/92JD01517, 1992.

Bock, O., Keil, C., Richard, E., Flamant, C., and Bouin, M.-N.: Validation of precipitable water from ECMWF model analyses with GPS and radiosonde data during the MAP SOP, Quarterly Journal of the Royal Meteorological Society, 131, 3013–3036, https://doi.org/10.1256/qj.05.27, publisher: Wiley, 2005.

Bock, O., Bosser, P., Flamant, C., Doerflinger, E., Jansen, F., Fages, R., Bony, S., and Schnitt, S.: IWV observations in the Caribbean Arc from a network of ground-based GNSS receivers during EUREC^{4}A, Earth System Science Data, https://doi.org/10.5194/essd-2021-50, 2021.

Bock, O., Mears, C. A., Ho, S. P., and Shao, X.: Total Column Water vapor, in: State of the climate in 2023, edited by Dunn, R., pp. 57–59, Bulletin of the American Meteorological Societ, https://doi.org/10.1175/BAMS-D-24-0116.1, 2024.

Boehm, J., Werl, B., and Schuh, H.: Troposphere mapping functions for GPS and very long baseline interferometry from European Centre for Medium-Range Weather Forecasts operational analysis data, Journal of Geophysical Research, 11, 2406–+, https://doi.org/10.1029/2005JB003629, 2006.

Boniface, K., C, C., Chery, J., Ducrocq, V., Rocken, C., Doerflinger, E., and Collard, P.: Potential of shipborne GPS atmospheric delay data for prediction of Mediterranean intense weather events, Atmospheric Science Letters, 13, 250–256, https://doi.org/10.1002/asl.391, 2012.

Bosser, P. and Bock, O.: IWV retrieval from ground GNSS receivers during NAWDEX, Advances in Geosciences, 55, 13–22, https://doi.org/10.5194/adgeo-55-13-2021, 2021.

Bosser, P. and Tulet, P.: Near real-time water vapour monitoring with shipborne GNSS for numerical weather prediction, EGU General Assembly 2024, https://doi.org/10.5194/egusphere-egu24-8726, poster, 2024.

Bosser, P., Bock, O., J, P., and Thom, C.: An improved mean gravity model for GPS hydrostatic delay calibration, IEEE Geoscience and Remote Sensing Letters, 4, 3–7, https://doi.org/10.1109/LGRS.2006.881725, 2007.

Bosser, P., Bock, O., Flamant, C., Bony, S., and Speich, S.: Integrated water vapour content retrievals from ship-borne GNSS receivers during EUREC^{4}A, Earth System Science Data, 13, 1499–1517, https://doi.org/10.5194/essd-13-1499-2021, 2021.

Bosser, P., Van Baelen, J., and Bousquet, O.: Routine Measurement of Water Vapour Using GNSS in the Framework of the MAP-IO Project, Atmosphere, 13, 903, https://doi.org/10.3390/atmos13060903, publisher: MDPI AG, 2022.

Bosser, P., Ancelin, J., Métois, M., Rolland, L., and Vidal, M.: Evaluation of tropospheric estimates from CentipedeRTK, a collaborative network of low-cost GNSS stations, GPS Solutions, 28, https://doi.org/10.1007/s10291-024-01699-3, 2024.

Chen, B., Yu, W., Wang, W., Zhang, Z., and Dai, W.: A Global Assessment of Precipitable Water Vapor Derived From GNSS Zenith Tropospheric Delays With ERA5, NCEP FNL, and NCEP GFS Products, Earth and Space Science, 8, e2021EA001 796, https://doi.org/10.1029/2021EA001796, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2021EA001796, 2021.

Copernicus Climate Change Service: ERA5 hourly data on single levels from 1940 to present, https://doi.org/10.24381/cds.adbb2d47, accessed: May 2025, 2023.

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

Gong, Y., Liu, Z., and Foster, J. H.: Evaluating the Accuracy of Satellite-Based Microwave Radiometer PWV Products Using Shipborne GNSS Observations Across the Pacific Ocean, IEEE Transactions on Geoscience and Remote Sensing, 60, 1–10, https://doi.org/10.1109/TGRS.2021.3129001, 2022.

Guerova, G., Jones, J., Douša, J., Dick, G., de Haan, S., Pottiaux, E., Bock, O., Pacione, R., Elgered, G., Vedel, H., and Bender, M.: Review of the state of the art and future prospects of the ground-based GNSS meteorology in Europe, Atmospheric Measurement Techniques, 9, 5385–5406, https://doi.org/10.5194/amt-9-5385-2016, 2016.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N.: ERA5 hourly data on single levels from 1940 to present, https://doi.org/10.24381/cds.adbb2d47, accessed: May 2025, 2018.

Ikuta, Y., Seko, H., and Shoji, Y.: Assimilation of shipborne precipitable water vapour by Global Navigation Satellite Systems for extreme precipitation events, Quarterly Journal of the Royal Meteorological Society, 148, 57–75, https://doi.org/10.1002/qj.4192, publisher: Wiley, 2021.

(JAXA), J. A. E. A.: GCOM-W1/AMSR2 Level-2 Total Precipitable Water (TPW) Product: Algorithm and Validation, Tech. rep., Earth Observation Research Center (EORC), JAXA, https://www.eorc.jaxa.jp/AMSR/satellite/pdf.gcom-w/AMSR2_L2.pdf, validation shows RMS error of about 2.9 kg m^{-2} and bias near 0.09 kg m^{-2} (radiosondes, GPS validation)., 2013.

Landskron, D. and Böhm, J.: VMF3/GPT3: refined discrete and empirical troposphere mapping functions, Journal of Geodesy, 92, 349–360, https://doi.org/10.1007/s00190-017-1066-2, publisher: Springer Science and Business Media LLC, 2017.

Mears, C. A., Wang, J., Smith, D., and Wentz, F. J.: Intercomparison of total precipitable water measurements made by satellite-borne microwave radiometers and ground-based GPS instruments, Journal of Geophysical Research: Atmospheres, 120, 2492–2504, https://doi.org/10.1002/2014JD022694, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2014JD022694, 2015.

Mears, C. A., Smith, D. K., Ricciardulli, L., Wang, J., Huelsing, H., and Wentz, F. J.: Construction and Uncertainty Estimation of a Satellite-Derived Total Precipitable Water Data Record Over the World's Oceans, Earth and Space Science, 5, 197–210, https://doi.org/10.1002/2018EA000363, 2018.

Männel, B., Zus, F., Dick, G., Glaser, S., Semmling, M., Balidakis, K., Wickert, J., Maturilli, M., Dahlke, S., and Schuh, H.: GNSS-based water vapor estimation and validation during the MOSAiC expedition, Atmospheric Measurement Techniques, 14, 5127–5138, https://doi.org/10.5194/amt-14-5127-2021, 2021.

Offiler, D.: EIG EUMETNET GNSS Vater Vapour Program : Products requirements document version 1.0, Tech. rep., EUMETNET, 2010.

Ohara, K., Kubota, T., Kachi, M., and Kazumori, M.: Comparison of Long-Term Total Precipitable Water Products by the Advanced Microwave Scanning Radiometer 2 (AMSR2), Journal of the Meteorological Society of Japan. Ser. II, 101, 289–308, https://doi.org/10.2151/jmsj.2023-018, 2023.

Panetier, A. and Bosser, P.: Integrated Water Vapor contents from GNSS observations collected onboard research vessels of the French Oceanographic Fleet, https://doi.org/10.25326/876, aERIS Data Centre [data set], 2026.

Panetier, A., Bosser, P., and Khenchaf, A.: Sensitivity of Shipborne GNSS Estimates to Processing Modeling Based on Simulated Dataset, Sensors, 23, https://doi.org/10.3390/s23146605, 2023.

Poli, P., Moll, P., Rabier, F., Desroziers, G., Chapnik, B., Berre, L., Healy, S. B., Andersson, E., and El Guelai, F.-Z.: Forecast impact studies of zenith total delay data from European near real-time GPS stations in Météo-France 4DVAR, Journal of Geophysical Research, 112, D06 114, https://doi.org/10.1029/2006JD007430, 2007.

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

Rocken, C., Johnson, J., Van Hove, T., and Iwabuchi, T.: Atmospheric water vapor and geoid measurements in the open ocean with GPS, Geophysical Research Letters, 32, L12 813, https://doi.org/10.1029/2005GL022573, 2005.

Saastamoinen, J.: Atmospheric Correction for the Troposphere and Stratosphere in Radio Ranging Satellites, in: The Use of Artificial Satellites for Geodesy, edited by AGU, pp. 247–251, Wiley Online Library, https://doi.org/10.1029/GM015p0247, 1972.

Shoji, Y., Miura, J., Tsubaki, S., Higashi, Y., Hibino, S., Kojima, A., Nakamura, T., and Shutta, K.: A Study on Analysis Setting Optimization of Ship-Based GNSS Measurements for Maritime Precipitable Water Vapor Monitoring, Journal of the Meteorological Society of Japan. Ser. II, 101, 323–346, https://doi.org/10.2151/jmsj.2023-020, 2023.

Steigenberger, P., Boehm, J., and Tesmer, V.: Comparison of GMF/GPT with VMF1/ECMWF and implications for atmospheric loading, Journal of Geodesy, 83, 943–951, https://doi.org/10.1007/s00190-009-0311-8, 2009.

Thayer, G.: An improved equation for the radio refractive index of air, Radio Science, 9, 803–807, https://doi.org/10.1029/RS009i010p00803, 1974.

Trenberth, K. E., Fasullo, J., and Smith, L.: Trends and variability in column-integrated atmospheric water vapor, Climate Dynamics, 24, 741–758, https://doi.org/10.1007/s00382-005-0017-4, 2005.

Tulet, P., Van Baelen, J., Bosser, P., Brioude, J., Colomb, A., Goloub, P., Pazmino, A., Portafaix, T., Ramonet, M., Sellegri, K., Thyssen, M., Gest, L., Marquestaut, N., Mékiès, D., Metzger, J.-M., Athier, G., Blarel, L., Delmotte, M., Desprairies, G., Dournaux, M., Dubois, G., Duflot, V., Lamy, K., Gardes, L., Guillemot, J.-F., Gros, V., Kolasinski, J., Lopez, M., Magand, O., Noury, E., Nunes-Pinharanda, M., Payen, G., Pianezze, J., Picard, D., Picard, O., Prunier, S., Rigaud-Louise, F., Sicard, M., and Torres, B.: MAP-IO: an atmospheric and marine observatory program on board *Marion Dufresne* over the Southern Ocean, Earth System Science Data, 16, 3821–3849, https://doi.org/10.5194/essd-16-3821-2024, 2024.

Wang, J., Wu, Z., Semmling, M., Zus, F., Gerland, S., Ramatschi, M., Ge, M., Wickert, J., and Schuh, H.: Retrieving Precipitable Water Vapor From Shipborne Multi-GNSS Observations, Geophysical Research Letters, 46, https://doi.org/10.1029/2019GL082136, 2019.

Wang, S., Xu, T., Nie, W., Jiang, C., Yang, Y., Fang, Z., Li, M., and Zhang, Z.: Evaluation of Precipitable Water Vapor from Five Reanalysis Products with Ground-Based GNSS Observations, Remote Sensing, 12, 1817, https://doi.org/10.3390/rs12111817, publisher: MDPI AG, 2020.

Wentz, F. J., Hilburn, K. A., and Smith, D. K.: Remote Sensing Systems DMSP SSM/I and SSMIS Daily Environmental Suite on 0.25° Grid, Version 7, https://www.remss.com/missions/ssmi, accessed: September 2025, 2012.

Wentz, F. J., Ricciardulli, L., Gentemann, C., Meissner, T., Hilburn, K. A., and Scott, J.: Remote Sensing Systems Coriolis WindSat Daily Environmental Suite on 0.25° Grid, Version 7.0.1, https://www.remss.com/missions/windsat, accessed: September 2025, 2013.

Wentz, F. J., Meissner, T., Scott, J., and Hilburn, K. A.: Remote Sensing Systems GPM GMI Daily Environmental Suite on 0.25° Grid, Version 8.2, https://www.remss.com/missions/gmi, accessed: September 2025, 2015.

Wentz, F. J., Meissner, T., Gentemann, C., Hilburn, K. A., and Scott, J.: RSS GCOM-W1 AMSR2 Daily Environmental Suite on 0.25 deg grid, Version 8.2, https://doi.org/10.56236/RSS-bq, 2021.

Wu, Z., Lu, C., Lyu, H., Han, X., Zheng, Y., Liu, Y., Liu, Y., and Jin, K.: Sensing Real-time Water Vapor over Oceans with Low-cost GNSS Receivers, IEEE Transactions on Geoscience and Remote Sensing, pp. 1–1, https://doi.org/10.1109/tgrs.2022.3213427, publisher: Institute of Electrical and Electronics Engineers (IEEE), 2022a.

Wu, Z., Lu, C., Zheng, Y., Liu, Y., Liu, Y., Xu, W., Jin, K., and Tang, Q.: Evaluation of Shipborne GNSS Precipitable Water Vapor Over Global Oceans From 2014 to 2018, IEEE Transactions on Geoscience and Remote Sensing, 60, 1–15, https://doi.org/10.1109/tgrs.2022.3142745, publisher: Institute of Electrical and Electronics Engineers (IEEE), 2022b.

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

Yuan, P., Blewitt, G., Kreemer, C., Hammond, W. C., Argus, D., Yin, X., Van Malderen, R., Mayer, M., Jiang, W., Awange, J., and Kutterer, H.: An enhanced integrated water vapour dataset from more than 10\hspace0.167em000 global ground-based GPS stations in 2020, Earth System Science Data, 15, 723–743, https://doi.org/10.5194/essd-15-723-2023, 2023.

Yuan, P., Blewitt, G., Kreemer, C., Jiang, W., Liu, T., He, L., Shan, Q., Balidakis, K., Schuh, H., Wickert, J., and Deng, Z.: A Global Assessment of Diurnal Discontinuities in ERA5 Tropospheric Zenith Total Delays Using 10 Years of GNSS Data, Geophysical Research Letters, 52, https://doi.org/10.1029/2024gl113140, 2025.

Appendix A: FOF vessels trajectories in western Europe

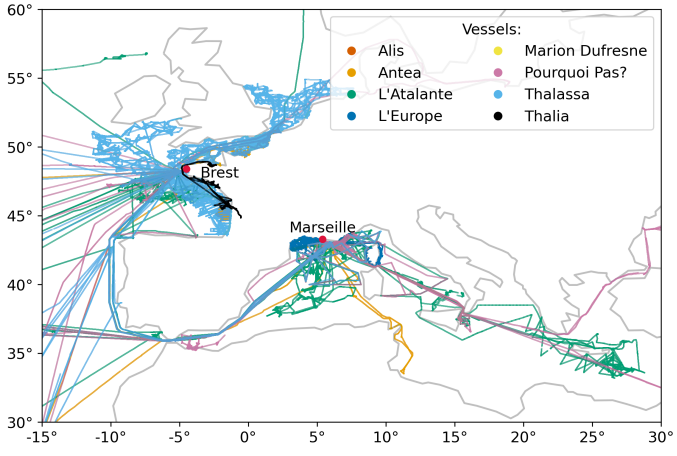


Figure A1. FOF trajectory map around Western Europe in 2015 - 2024.

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

Open Access Earth System Science Data Discussions

Appendix B: FOF GNSS products screening distribution

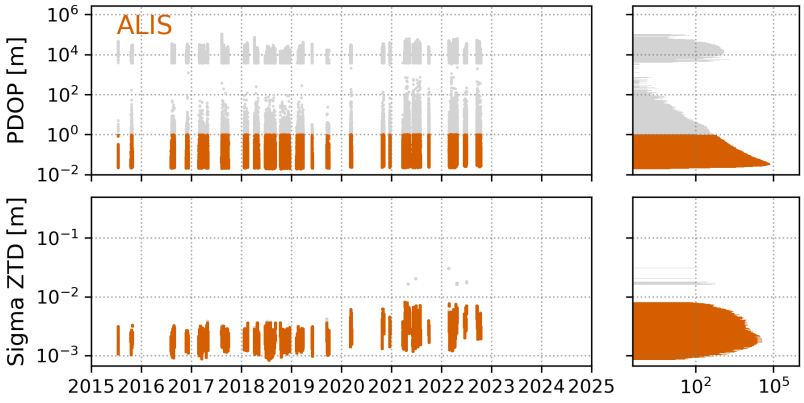


Figure B1. Alis time series and distribution of formal errors on position (top) and ZTD (bottom) and corresponding histograms. The colored plot corresponds to data kept after screening, while grey plot corresponds to filtered out data.

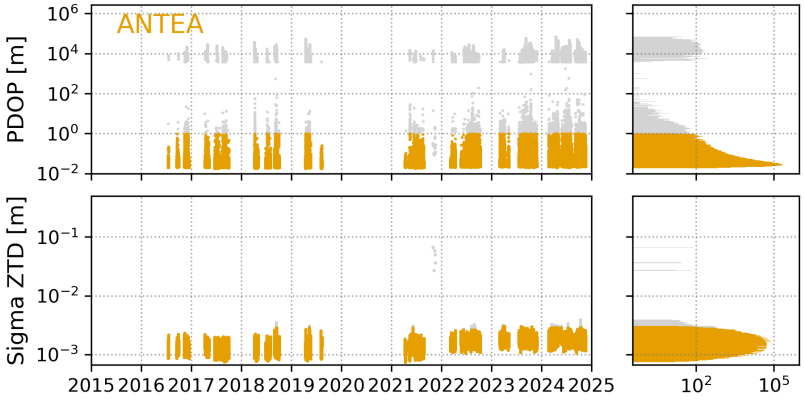


Figure B2. Antea time series and distribution of formal errors on position (top) and ZTD (bottom) and corresponding histograms. The colored plot corresponds to data kept after screening, while grey plot corresponds to filtered out data.

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

Open Access Earth System Science Data Discussions

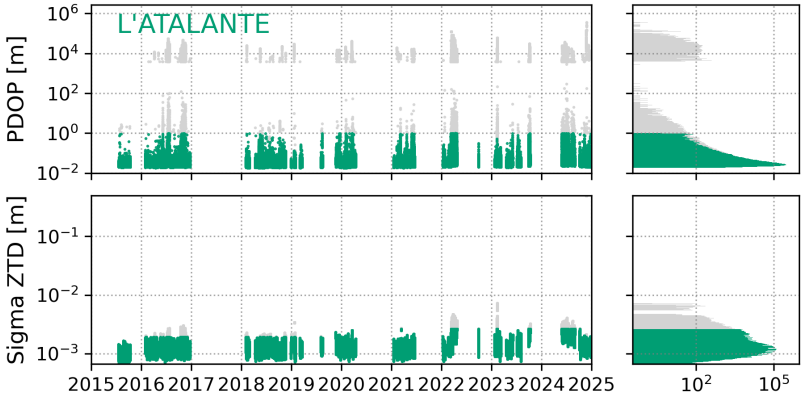


Figure B3. L'Atalante time series and distribution of formal errors on position (top) and ZTD (bottom) and corresponding histograms. The colored plot corresponds to data kept after screening, while grey plot corresponds to filtered out data.

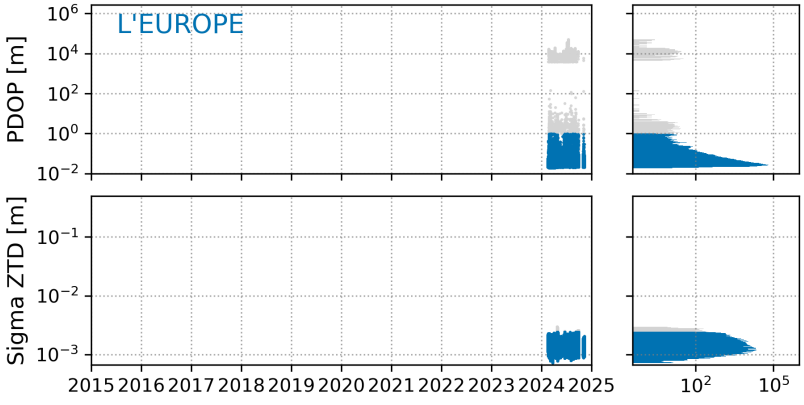


Figure B4. L'Europe time series and distribution of formal errors on position (top) and ZTD (bottom) and corresponding histograms. The colored plot corresponds to data kept after screening, while grey plot corresponds to filtered out data.

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

Open Access Earth System Science Data Discussions

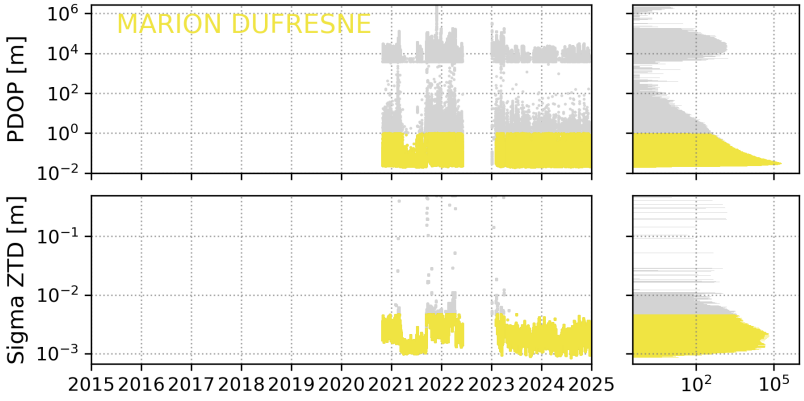


Figure B5. Marion Dufresne time series and distribution of formal errors on position (top) and ZTD (bottom) and corresponding histograms. The colored plot corresponds to data kept after screening, while grey plot corresponds to filtered out data.

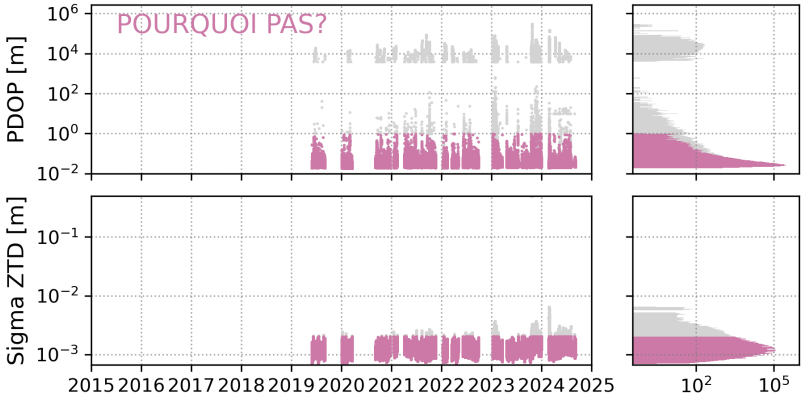


Figure B6. Pourquoi Pas? time series and distribution of formal errors on position (top) and ZTD (bottom) and corresponding histograms. The colored plot corresponds to data kept after screening, while grey plot corresponds to filtered out data.

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026

Open Access Earth System Science Data Discussions

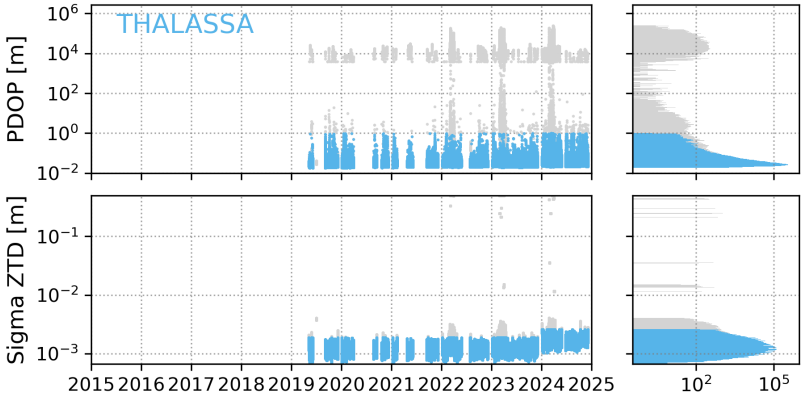


Figure B7. Thalassa time series and distribution of formal errors on position (top) and ZTD (bottom) and corresponding histograms. The colored plot corresponds to data kept after screening, while grey plot corresponds to filtered out data.

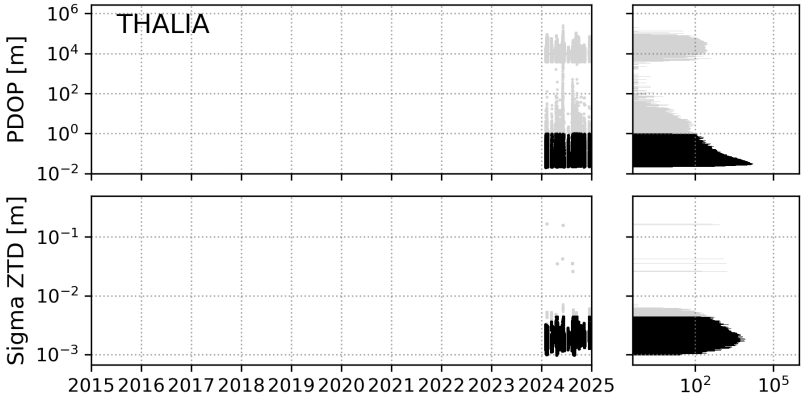
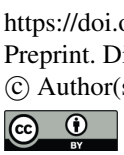


Figure B8. Thalia time series and distribution of formal errors on position (top) and ZTD (bottom) and corresponding histograms. The colored plot corresponds to data kept after screening, while grey plot corresponds to filtered out data.

https://doi.org/10.5194/essd-2026-223
Preprint. Discussion started: 20 April 2026



Open Access Earth System Science Data Discussions

Appendix C: Latitude impact in the comparison to ERA5 for each vessel

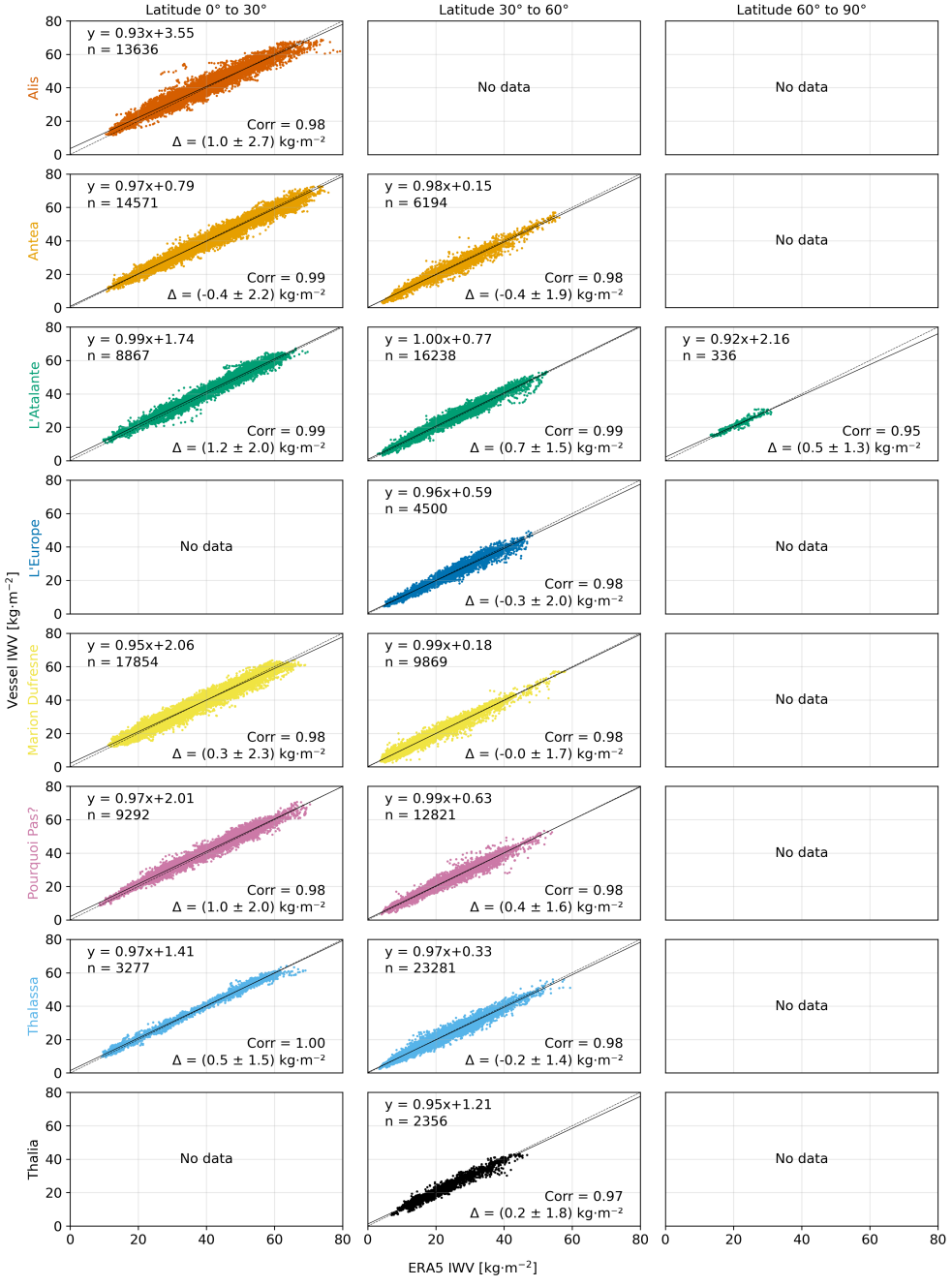


Figure C1. Scatter plots of ERA5 vs GNSS IWV for individual ships, separated by latitude ranges (0 to 30°, 30 to 60°, 60 to 90°).