

Dear Editor and Reviewers,

We thank the Editor and both reviewers for their careful and constructive evaluation of our manuscript. We have carefully considered all comments and revised the manuscript accordingly. We believe that the revisions have substantially improved the clarity, reproducibility, and overall usefulness of the benchmark dataset.

The main revisions include clarification of data availability and update procedures; improved descriptions of the water-level and wave observations, processing methods, and uncertainty estimates; clarification of the scope and future use of the complementary HF radar and AirSWOT observations; and correction of several minor terminology and typographical issues.

In particular, in response to Reviewer #2, we have substantially revised the geodetic conversion procedure in Section 3.1.3. We now provide a step-by-step workflow for converting the in situ observations from the Canadian geodetic reference system to the vertical reference used by SWOT, including the CGVD28/HTv2.0 conversion, the NAD83(CSRS)--ITRF2014 transformation, the permanent deformation correction, and the use of the SWOT-provided mean-tide EGM2008 geoid. The workflow is summarized in a new table, and the corresponding conversion factors for each station are provided in Table A1, allowing users to reproduce the conversion and directly compare the observations with SWOT measurements.

Detailed point-by-point responses to all comments are provided below. Reviewer #1 comments and responses are presented in *italic blue*, while Reviewer #2 comments and responses are presented in *italic red*.

We again thank the Editor and Reviewers for their thoughtful comments and suggestions, which have helped us significantly improve the manuscript.

Sincerely,

Pascal Matte, on behalf of the co-authors

Anonymous Referee #1

General comments

The manuscript provides an overview of new dataset that was deployed specifically as part of the SWOT Cal/Val, but has endured beyond this period. The datasets and instruments are well explained, as are the quality control parameters.

Response: We thank the reviewer for the constructive feedback. Our responses to Referee #1 are provided below using italic blue text.

One thing that is unclear is whether any data available in real time to the user? Or is all data only available after instrument recovery (i.e., post-processed from the memory card).

Response: Most sensors were not equipped with telemetry and, consequently, the data are not available in real-time. Data stored on memory cards are retrieved during sensor maintenance, which is performed at least annually. The data then undergo post-processing and quality control before being made available through updates of the database.

Specific questions/issues

- Line 13: It would be good to define when this dataset will be ‘periodically’ available.
Response: We have clarified in the manuscript that updates are planned on an approximately annual basis. However, because the timing of updates depends on instrument maintenance schedules and the time required for data processing and quality control, the exact timing cannot be guaranteed. In the abstract, we added: “[...] will be updated approximately annually, as new observations are retrieved, processed, and quality controlled [...]”. In the code and data availability section, we added “Updates are targeted approximately annually, depending on data retrieval, processing, and quality-control activities”. We have also modified Figure 5 to add ‘X’ markers at the end of the timelines for stations that have been dismantled and will no longer be updated.
- Line 92: Consider replacing ‘right shore’ with a cardinal direction – i.e., ‘eastern shoreline’.
Response: We agree that “right shore (in the direction of upstream tidal propagation)” may be ambiguous to readers unfamiliar with the local tidal dynamics and terminology. In English, the conventional geographical terminology used for the St. Lawrence River is “North Shore” and “South Shore”, despite the predominantly diagonal orientation of the river. We have therefore replaced “right shore” with “North Shore” to avoid ambiguity.
- Figure 1: It might be helpful in Figure 1 to include a rough outline of which section is seasonally covered in ice.
Response: According to the sea ice-duration climatology presented by Galbraith et al. (2024, their Figure 2; <https://doi.org/10.1029/2023JC020784>), the St. Lawrence Estuary has a climatological ice duration ranging from 7 to 15 weeks over the area presented in Figure 1. Thus, delineating the “seasonally ice-covered” section in Figure 1 would essentially involve identifying the

entire estuary. Moreover, ice conditions may vary substantially from year to year. We therefore chose not to add a fixed ice-cover boundary to Figure 1.

- Line 150-165: It would be helpful to be a bit more consistent in your reporting of error. At lines 150-165 you begin by reporting ‘expected’ errors less than Xcm (Xsigma); but then you mention +/- X cm at line 164. Are the errors from the latter sensors not able to be characterised in the same way? *Response: In line 164, the ± 0.3 cm and ± 0.5 cm values reported for the Solinst Levelogger and Seametrics CT2X sensors, respectively, were manufacturer specifications rather than estimates of the total propagated uncertainty. To ensure consistency in the revised manuscript, we have removed this sentence and, instead, we now report an estimated upper bound of the uncertainty of 7 cm (1 sigma) for these observations. This value corresponds to the largest uncertainty among the five sensors and is therefore used as a common, conservative uncertainty estimate for this group of observations. The individual vertical accuracies reported by the RTK-GNSS solution are provided in Table A1. We also refer readers to Sect. 3.1.3 for a detailed description of the datum corrections and their associated uncertainties.*
- Line 250: This idea about truncation is a bit confusing – does this simply refer to areas that were not surveyed (i.e., did not have observations)? If so, then perhaps just clarify that other (comparable) sites outside of the study area may experience water level truncation as well? *Response: The truncation does not refer to areas that were not surveyed or to missing observations. Rather, it occurs when low water levels expose extensive intertidal flats in the vicinity of a sensor. Under these conditions, the sensor can measure the elevation of the exposed bed surface rather than the water level. Because the bed elevation remains essentially constant while the area is exposed, the resulting water level time series can appear truncated or flat for a period of time, until the rising tide inundates the area again. We have revised the manuscript to clarify this point.*
- Line 260-264: Define all depths for wave observations in text. *Response: We have added the water depth at each wave buoy location to the main text. These depths are also provided in Table A3, together with the buoy locations, operating dates, and GPS coordinates.*
- Line 275: Is this the standard sampling plan for AZMP wave buoys? Or customised to this site? Regardless, it should be explained why this sampling plan is used (i.e., to target specific wave frequencies that dominate this region?). *Response: Thank you for highlighting this point; the AZMP sampling protocol is uncommon for wave measurements and warrants explicit clarification. This sampling strategy is used across all sites to manage shared computing resources among onboard sensors, thereby limiting the data volume allocated for wave monitoring. Due to these constraints, a customized sampling scheme to reduce intermittency could not be requested. We have clarified this in the revised manuscript.*
- Line 305: It’s unclear how wave calculations are conducted over 30min when they appear to be collected in 10min bursts (with 5 min gaps). Please expand on this. *Response: We agree that this was not sufficiently clear in the original text. We have revised the manuscript to explicitly describe the sampling and processing scheme: ‘Bulk wave parameters were*

calculated from the average cross-spectral density of paired, consecutive 10-minute bursts recorded every 15 minutes, yielding an effective temporal resolution of 30 minutes.'

- Line 311: is 'provenance' wave direction? *Response: Wave "provenance" refers to the direction from which the waves originate, whereas wave "direction" refers to the direction toward which they travel (relative to the buoy). In the revised manuscript, we specified that the wave's provenance is the angle taken relative to its sources.*

Technical corrections

- Line 267: small typo – should be AZMP not AMZP. *Response: Corrected.*
- Line 274: perhaps add the specific wave buoys you're referring to here with these dimensions. *Response: To facilitate a more direct comparison between AZMP directional wave buoys and commercial standards, we categorized some recognizable brands (Sofar Spotter, Datawell Waverider series, NextSens CB series and Triaxys products) into compact (40-70 cm, 7-30 kg) and medium-duty (70-150 cm, 50-250 kg) size classes. The manuscript has been modified accordingly.*
- Line 276: small typo – 'a245 ' ? *Response: Corrected.*

Anonymous Referee #2

This paper provides a valuable and dense dataset of 18 water level measurements (barometrically-corrected), supplemented with 13 GNSS-Interferometric Reflectometry sensors, and 3 directional wave buoys for validation of SWOT's nadir and wide-swath altimetry products. Location is the St. Lawrence Estuary and Saguenay Fjord in Quebec. This is a significant effort - deploying and maintaining the instruments for this dataset.

Response: We thank the reviewer for the constructive feedback. Our responses to Referee #2 are provided below using italic red text.

Other data are mentioned such as - high-frequency radars deployed during SWOT Cal/Val period and the AirSWOT campaign conducted in August 2023 (Simard et al., 2023). Why aren't the radar measurements or the data from the AirSWOT campaign available?

Response: The HF radar and AirSWOT observations were acquired as part of the broader observing effort described in the manuscript, but they are not included in the present dataset because their processing and validation require dedicated workflows that are substantially different from those used for the in situ water-level and wave observations presented here.

For the HF radars, extracting wave information is particularly demanding because wave parameters are derived from the second-order components of the Doppler spectrum, which are substantially weaker than the first-order Bragg peaks and are therefore more sensitive to noise, range, sea state, and other factors. The resulting wave products require dedicated processing and validation against independent observations before they can be considered suitable for SWOT validation. The HF radar observations and derived products will therefore be documented separately.

Similarly, the AirSWOT observations require dedicated airborne interferometric processing and validation, including the characterization of measurement uncertainties and comparison with independent in situ observations.

The water-level and wave observations presented in this paper are therefore not only a benchmark dataset for SWOT validation, but also provide an independent reference dataset that can be used to process, validate, and assess the HF radar and AirSWOT observations. These complementary datasets will be documented separately once their respective processing and validation workflows are completed. We have clarified this distinction in the revised manuscript, in both the introduction and conclusion sections.

The description of the processing is appreciated as well as the multi-sensor intercomparisons. However, for a paper providing a benchmark dataset for SWOT, it should include actual SWOT comparisons with the measured data.

Response: We agree that comparison with actual SWOT observations is an important application of the dataset. Such an analysis is currently being conducted in a separate manuscript in

preparation, which will assess SWOT observations over the St. Lawrence Estuary as well as in other estuarine environments worldwide.

The primary objective of the present paper is to establish and comprehensively document the benchmark in situ dataset, including the instrumentation, data processing, quality control, uncertainty characterization, and multi-sensor intercomparisons. We believe that the actual comparison with SWOT products warrants a dedicated study because it involves additional methodological considerations specific to SWOT, including the treatment of layover, product-specific corrections, spatial and temporal sampling differences, and the evaluation of SWOT measurement uncertainties. Including this analysis in the present paper would substantially expand its scope beyond the documentation and characterization of the benchmark dataset.

We have nevertheless clarified throughout the manuscript that the dataset is specifically designed to support SWOT validation and that the in situ observations provide an independent reference for assessing SWOT products. The forthcoming SWOT-focused study will build directly on the benchmark dataset presented here.

Particularly, the issue of geoid transformation is the most ambiguous, and the paper mentions "care must be taken" and points to the SWOT handbook. A user wanting to actually do the SWOT comparison has to reconstruct the chain themselves. For a benchmark dataset I find it as a real omission and all the workflow should be provided (codes included) to directly compare the measured data with SWOT. This would make the dataset easy to apply and be used along with SWOT.

Response: We agree that the original description of the geodetic reference systems and datum conversion was not sufficiently explicit to allow users to readily reproduce the transformation from the in situ observations to the SWOT vertical reference. We have therefore substantially revised Section 3.1.3 to provide a step-by-step conversion workflow, also summarized in a new table (Table 1).

The revised text now describes how to (1) convert the harmonized CGVD28 heights to ellipsoidal heights using the HTv2.0 realization at epoch 1997.0, using the NRCan GPS-H tool; (2) transform the resulting NAD83(CSRs) epoch 1997.0 coordinates to ITRF2014 at the measurement epoch using the NRCan TRX tool; (3) account for the permanent deformation associated with the difference between the conventional tide-free GNSS reference and the convention used by SWOT, following Eq. (11.7) of the SWOT User Handbook; and (4) obtain water surface elevations relative to the SWOT mean-tide EGM2008 reference by subtracting the geoid height provided directly in the SWOT product (Version D or earlier). We also clarify that no additional mean-tide correction is required when the SWOT-provided geoid height is used. The conversion factors used for each step and station are now provided in Table A1, allowing users to apply the procedure directly to the benchmark dataset.

For consistency across the benchmark dataset, we use a fixed reference epoch of 11 July 2023 (epoch 2023.5233), corresponding to the end of the SWOT Calibration/Validation phase, for the NAD83(CSRs)--ITRF2014 transformation factors (Δh_{TRX}) provided in Table A1. We estimate that the resulting vertical error remains on the order of 1 cm over the first three years of the mission

(2023--2026), which is negligible compared with other sources of uncertainty associated with the in situ surveys. Users requiring higher accuracy for individual SWOT overpasses can nevertheless calculate conversion factors for the specific epoch of each overpass.

Although the transformations rely on the publicly available NRCan GPS-H and TRX tools rather than custom code, all intermediate conversion factors are provided in Table A1. Together with the step-by-step procedure described in Section 3.1.3, this allows users to reproduce the conversion and directly compare the in situ observations with SWOT measurements.

These additions provide a reproducible workflow for converting the in situ observations to the SWOT reference system while retaining the original Canadian geodetic reference information.

Check also AMZP" vs AZMP inconsistency.

Response: Corrected.