<u>Response to Anonymous Referee #2</u>

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.