The Significance of the Long-Wavelength Correction for Studies of Baroclinic Tides with SWOT

Edward D Zaron[1]

[1]Oregon State University

April 08, 2024

Abstract

The long-wavelength correction (LWC) of SWOT data is intended to reduce errors related to the stability of the SWOT antenna and its attitude in orbit, especially those errors related to the roll of the satellite.The algorithms used to compute the LWC utilize SWOT KaRIn sea surface-height (SSH) measurements and additional data, and the LWC may abosrb geophysical SSH into the correction. Different LWC algorithms are used on the L2 and L3 SWOT products, which are analyzed here during the 1-day repeat (Cal/Val) mission phase lasting approximately 100 days. During this mission phase the SSH anomaly (SSHA) computed using the L3 LWC is much more realistic than the L2 LWC, shown here by comparing spatial statistics of the L2 and L3 products.The L3 LWC algorithm is nonlinear insofar as it depends on second-order statistics of the SSHA and multi-satellite SSHA differences, making it difficult to quantify the extent to which it could absorb baroclinic tidal signals. To overcome this difficulty, a proxy L3 LWC algorithm is developed which mimics the L3 LWC but is strictly linear in the SSHA. The proxy LWC is applied to the predicted internal tide available on the products, and it is found to absorb roughly 5% to 10% of the variance of the internal tide; although, this figure varies strongly depending on the magnitude and orientation of the tidal waves.

The Significance of the Long-Wavelength Correction for Studies of Baroclinic Tides with SWOT

Edward D. Zaron[1]

[1]College of Earth, Ocean, and Atmospheric Science, Oregon State University

Key Points:

- The long-wavelength correction has the potential to remove geophysical signals of interest
- The correction is implemented as a nonlinear data-driven filter, making it difficult to characterize its response using linear techniques
- A linear approximation of the filter is developed here and used to characterize the response to baroclinic tides in the open ocean

Corresponding author: Edward D. Zaron, `Edward.D.Zaron@oregonstate.edu`

Abstract

The long-wavelength correction (LWC) of SWOT data is intended to reduce errors related to the stability of the SWOT antenna and its attitude in orbit, especially those errors related to the roll of the satellite. The algorithms used to compute the LWC utilize SWOT KaRIn sea surface-height (SSH) measurements and additional data, and the LWC may abosrb geophysical SSH into the correction. Different LWC algorithms are used on the L2 and L3 SWOT products, which are analyzed here during the 1-day repeat (Cal/Val) mission phase lasting approximately 100 days. During this mission phase the SSH anomaly (SSHA) computed using the L3 LWC is much more realistic than the L2 LWC, shown here by comparing spatial statistics of the L2 and L3 products. The L3 LWC algorithm is nonlinear insofar as it depends on second-order statistics of the SSHA and multi-satellite SSHA differences, making it difficult to quantify the extent to which it could absorb baroclinic tidal signals. To overcome this difficulty, a proxy L3 LWC algorithm is developed which mimics the L3 LWC but is strictly linear in the SSHA. The proxy LWC is applied to the predicted internal tide available on the products, and it is found to absorb roughly 5% to 10% of the variance of the internal tide; although, this figure varies strongly depending on the magnitude and orientation of the tidal waves.

Plain Language Summary

In order to make the SWOT data useful for studies of sea level associated with horizontal length scales of roughly 100 km and larger, it is necessary to compute a correction which aligns the measured SSH with independent data from other satellite missions. The correction is implemented with a type of data-driven spatial low-pass filter; however, the data-driven character of this filter means that it is nonlinear and its response cannot be characterized using standard technqiues of linear filter analysis. In this paper a linear approximation of the long-wavelength correction is developed which is amenable to standard linear analysis techniques. Using this linear approximation, the extent to which the long-wavelength correction may absorb signals of interest arising from the baroclinic tides is quantified. It is found that the filter response is generally below 10% of the signal variance, which could lead to small but non-negligible errors in studies of tides using data from the Cal/Val mission phase.

1 Introduction

The SWOT KaRIn instrument is an imaging interferometer(Fjørtoft et al., 2010). Because its orbit altitude is very much larger than the KaRIn antenna baseline, SWOT sea surface-height (SSH) measurements are sensitive to uncontrolled perturbations in the attitude of the instrument. The SSH measurements are also influenced by electrical path delays, thermal effects, and mechanical stability of the KaRIn antenna. Because these factors typically evolve slowly during the orbit of the satellite, they are associated with long-wavelength errors in the SSH. The SWOT mission objectives for observation of oceanic SSH place strict requirements on the error budget for observations at scales from 150 km to 4 km(Esteban-Fernandez et al., 2010; Morrow et al., 2019), and these are largely independent of the long-wavelength errors. Nonetheless, the long-wavelength errors may be significant to observations of certain geophysical phenomena, such as ocean tides, and so corrections have been developed to reduce the long-wavelength errors over the ocean.

Table 1. Tidal alias periods (days) for SWOT during the one-day repeat orbit phase

Mission repeat [days]	Q_1	O_1	P_1	S_1	K_1	N_2	M_2	S_2	K_2	M_4	MS_4
0.99349	9	13	108	153	262	9	12	76	131	6	11

This manuscript analyzes the long-wavelength correction (LWC) provided by the presently available SWOT ocean products during the 1-day repeat (Cal/Val) orbit phase. The 1-day repeat phase, which lasted from 2023-03-30 to 2023-07-10, provided nearly 100 days of data. The duration of this orbit phase is sufficient to estimate several of the larger tides using harmonic analysis (Table 1), and the daily repeats may contain useful information about the non-phase-locked or modulated tides which are not available from other data sources. Thus, in spite of the relatively limited geographic coverage (Figure 1), there is interest in understanding the usefulness of these data for tidal studies.

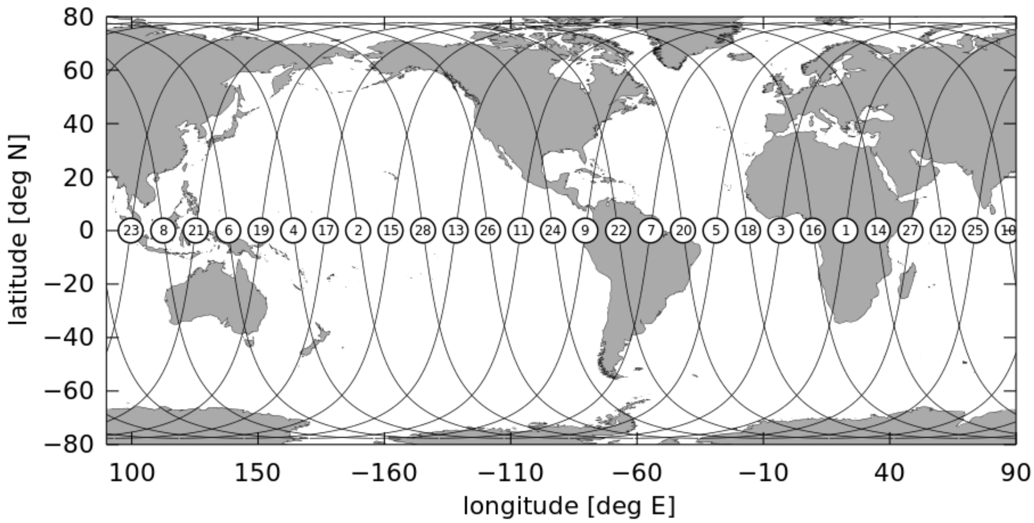

Figure 1. SWOT ground tracks and pass numbers during the 1-day-repeat orbit phase, from 2023-03-30 to 2023-07-10.

2 SWOT Data Products and Long-Wavelength Corrections (LWC)

Two SWOT data products are used here, namely, (1) the Level 2 Low Rate Sea Surface Height Data Product, version 1.1 (Beta Pre-Validated Product, personal communication, Shailen Desai, 2023-10-19)[1] and (2) the NRT SWOT KaRIn & nadir Global Ocean swath SSALTO/DUACS Sea Surface Height L3 product, version 0.3, CRM:0069553 (personal communication, AVISO, 2024-01-10)[2]. These shall be referred to as the L2 and L3 products, respectively.

The L2 and L3 both products provide data on a fixed geographic grid at a resolution of 2 km, the Low-Resolution grid. The L2 product provides two versions of the SSH anomaly (SSHA). The first version (`ssha_karin`; henceforth, names written in `this font` refer to the names of variables in the SWOT data products) utilizes data from the on-board radiometer and dual-frequency nadir altimeter to compute a suite of geophysical corrections which are applied to the interferometric data to compute SSHA; this will be referred to as L2a SSHA, below. The second version (`ssh_karin_2`) utilizes model-based corrections for the wet path delay and the sea state bias correction to reduce dependence on non-KaRIn SWOT data; this will be referred to as L2b SSHA, below. The L2b product contains a slightly larger quantity of valid SSHA data compared to the L2a product. Both the L2a and L2b products contain an identical suite of corrections for the barotropic

[1] `doi.org/10.24400/527896/A01-2023.015` Accessed 2024-02-05.
[2] `doi.org/10.24400/527896/A01-2023.017` Accessed 2024-02-05.

ocean, load, solid earth, and pole tides; dynamic atmosphere correction; dry troposphere correction; ionospheric path delay; mean sea surface; geoid; baroclinic tides (`internal_tide_hret`), and LWC (`height_cor_xover`). A detailed description of the L2 products is found in the Product Description[3].

The L3 products are derived from the L2 products but employ a different algorithm to compute the LWC (`calibration`), described below. As with the L2 product, there are two versions of the SSHA provided, a default version (`ssha`) and a "noiseless" version (`ssha_noiseless`) which uses a nonlinear data-adaptive filter developed with machine-learning techniques to reduce the small-scale noise of the SSHA product. These two versions shall be referred to as the L3a and L3b SSHA products.

The LWC is a mitigation for the "roll error" which arises from imperfect knowledge of the orientation of the KaRIn antenna relative to the ocean surface. The relative positions of the KaRIn antennas must be known with micrometer position to achieve the centimeter precision of the SWOT SSHA measurement (Dibarboure & Ubelmann, 2014). In practice, this level of precision is not available from the on-orbit information, and there are additional systematic errors associated with the phase-screen and thermo-mechanical design, all of which the LWC is intended to correct. The design goal of the LWC is to remove the time-invariant and slowly evolving signals associated with these errors while leaving unaffected the signals from oceanic variability. In practice, the LWC error is modelled as the sum of bias, linear, and quadratic components in the across-track direction, separately within the left and right sides of each KaRIn swath, and the coefficients describing these components are permitted to vary over time scales of 3 minutes and longer. Because this error evolution timescale corresponds to an along-track length scale of roughly 1000 km, it is referred to as long-wavelength error.

While the L2 and L3 LWC employ the same basic formulation, they rely on different approaches and use different data to estimate the correction (Dibarboure et al., 2022). The L2 algorithm is designed to be used in near-real time, and utilizes only SWOT data. The coefficients for the along- and across-track errors are constrained by minimizing the weighted difference between corrected SSHA in the crossovers between ascending and descending swaths. Away from crossovers, the coefficients are determined by smoothly interpolating them in the along-track direction with a 1000-km Gaussian kernel, constrained by the crossover values.

The L3 algorithm uses more external data than the L2 algorithm. It is designed to be used in non-real-time mode when high-quality information from the constellation of nadir satellite altimeter missions is available. The idea is roughly the same, though, that coefficients for the mean, linear, and quadratic components of the LWC are estimated by optimizing agreement between the corrected SWOT SSHA and other missions where the tracks cross. The various data and smoothness constraints are weighted according to the residual differences between the different sources and the time-offsets between their observations.

One simple assessment of the L2 and L3 LWCs is provided by the spatial variogram of the SSHA.The variogram is computed by averaging over the data from the 1-day-repeat orbit period as,

$$\Gamma^2(\Delta x, \Delta y, x_0) = \langle (\eta(x_0 + \Delta x, y + \Delta y) - \eta(x_0, y))^2 \rangle \tag{1}$$

where $\eta(x, y)$ represents the SSHA at across-track pixel x and along-track pixel y, Δx and Δy represent across- and along-track separation, the angle brackets represent averaging over all orbit cycles in the 1-day-repeat, and x_0 represents a reference pixel in

[3] https://archive.podaac.earthdata.nasa.gov/podaac-ops-cumulus-docs/web-misc/swot_mission_docs/pdd/D-56407_SWOT_Product_Description_L2_LR_SSH_20220902_RevA.pdf Accessed 2024-02-05.

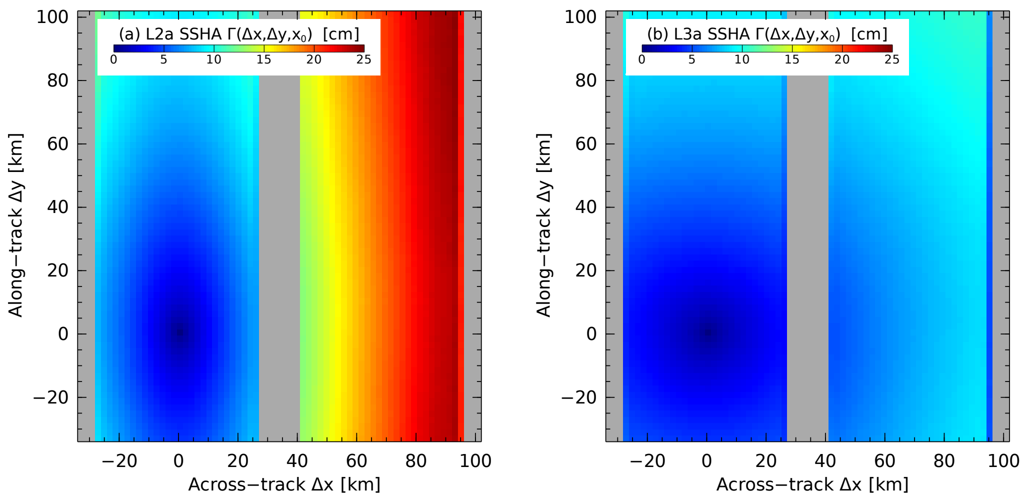

Figure 2. The square root of the spatial variogram ($\Gamma(\Delta x, \Delta y, x_0)$, equation 1) is shown for the (a) L2a and (b) L3a SSHA products. Note that the global minimum of $\Gamma(\Delta x, \Delta y, x_0)$ occurs at $\Delta x = \Delta y = 0$ at the x_0 pixel in the middle of the left swath.

the middle of the left-side swath. The averaging to compute the variogram is restricted to use data from pixels with the best quality flags. Furthermore, averages have been computed over the four latitude ranges, $(-72°, -45°)$, $(-45°, 0°)$, $(0°, 45°)$, $(45°, 72°)$, to look for differences related to the winter/summer hemispheres; however, only the results for the $(0°, 45°)$ range are shown here as no significant differences between the regions were noted.

A comparison of the L2a and L3a SSHA variograms is shown in Figure 2, and it reveals the large-scale structure which is a consequence of the different LWC algorithms (note that Γ is shown, units of cm). The magnitude of the L3a variogram is smaller than that of the L2a variogram, which is consistent with a more accurate LWC for the L3 product and consistent with the larger amount of data involved in computing the L3 LWC. The L2a variogram is notably anisotropic and it also exhibits a jump from one side of the swath to the other, both of which are unrealistic features. Both products also exhibit unrealistic variogram structures in the outermost and innermost pixels on each side; this is apparently related to how SSHA is interpolated onto the geographically fixed grid (Clement Ubelmann, personal communication).

The isotropy of the L3 variogram is quite remarkable, considering that the LWC is modeled with separate parameterizations in the along- and across-track directions. The differences between the L2 and L3 products and their "a" and "b" versions are revealed in cross-sections through the Γ function in Figure 3. The L2b product contains slightly more variance than the L2a product, visible at the largest Δx and Δy as larger values for the red versus the black lines in Figure 3a. In other words, the model-based corrections are slightly noisier than the radiometer-based corrections, as would be expected. In contrast, the L3a and L3b products only differ significantly at scales smaller than about 20 km, and the along- and across-track partitioning of the variance is very nearly equal (Figure 3b). The features noted are consistent with the intended performance of the L3 LWC and small-scale de-noising algorithms.

From the variogram it is not possible to assess whether the LWC is absorbing any ocean signals. The comparisons simply reveal that the L3 product exhibits a plausible rendition of the spatial covariance of ocean surface topography and the L2 product does

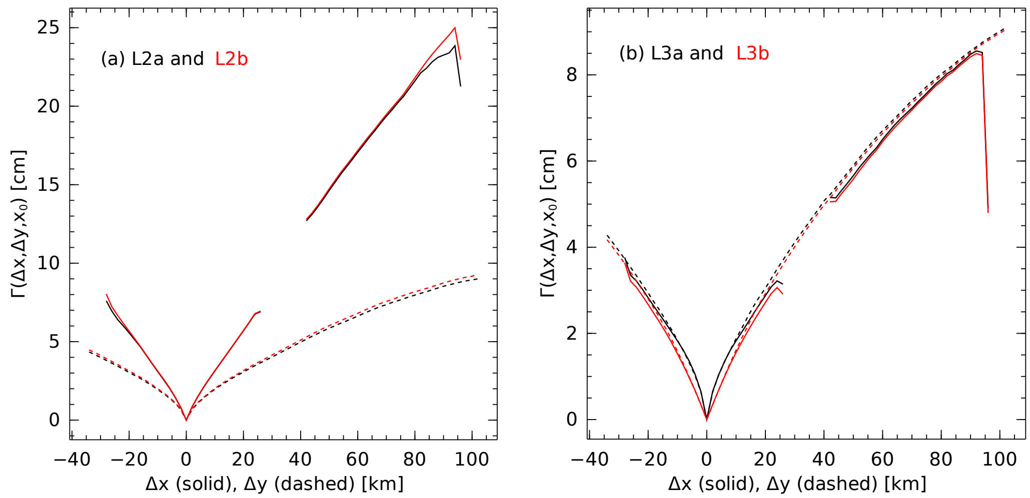

Figure 3. Cross-sections through $\Gamma(\Delta x, \Delta y, x_0)$ in the across-track (Δx, solid lines) and along-track (Δy, dashed lines) reveal the anisotropy of of the L2 products. The panels compare (a) the L2a and L2b products (black and red, respectively) and (b) the L3a and L3b products. Note the very different vertical scales used in each panel.

not. Recall that only data from the 1-day-repeat orbit phase are used here. Due to the sparse crossovers during this orbit phase, the findings shown in Figures 2a and 3a are not unexpected.

Details of the above L2 and L3 algorithms are provided in Dibarboure et al. (2022); however, the algorithms are implemented using weighted least-squares, with the weights derived from covariances estimated using crossover residuals. The data-driven character of the algorithms introduces nonlinearity; therefore, the LWC is not a simple linear filter acting on the measured data. This aspect of the LWC makes it impossible to assess whether it could distort or filter out signals of real geophysical processes, such as the baroclinic tides.

3 A Proxy LWC for Assessing the L3 LWC Using L2 Data

This section describes the development of a linear approximation to the L3 LWC which will be called the "proxy LWC."

The notation used in this section expresses the observed SSHA, $\eta(x, y)$—a function of cross-swath coordinate x and along-swath coordinate y—using the boldface vector notation, $\boldsymbol{\eta}$. The notation connotes that $\boldsymbol{\eta} \in R^N$ is regarded as a 1-dimensional vector containing the N values of $\eta(x, y)$ at each valid pixel for a given orbit cycle. The presence of missing data, which could be caused by environmental effects or instrument anomalies, causes the number of valid pixels, N, to vary in time. Linear operators which act on $\boldsymbol{\eta}$ will be written in matrix notation as bold capital letters, e.g., $\mathbf{U}$, where $\mathbf{U} \in R^{M \times N}$. The rank M will generally be unstated but implied by the context or dimension of the vectors comprising $\mathbf{U}$. When $\mathbf{U}$ consists of M ortho-normalized basis vectors on R^N, the notation, $\mathbf{U}^T\mathbf{U}\boldsymbol{\eta}$, shall be used to denote the projection of $\boldsymbol{\eta}$ onto this M-dimensional subspace.

The proxy LWC is a linear operator acting on the uncorrected SSHA, $\boldsymbol{\eta}_0$. It consists of two stages which are applied to 5000-km segments of SWOT data.

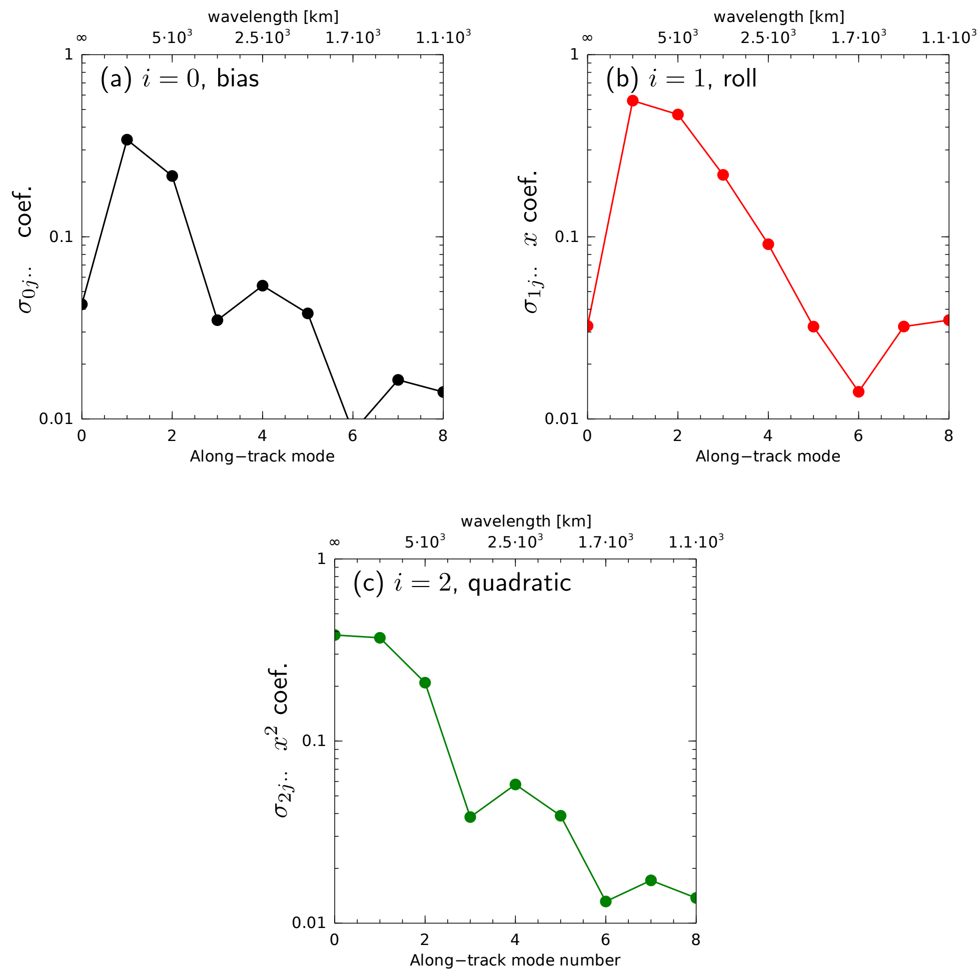

Figure 4. Proxy LWC weight coefficients, σ_m (equation (7)), for a pass #15 across the central Pacific. The coefficients of the (a) bias ($i = 0$), (b) roll ($i = 1$), and (c) quadratic ($i = 2$) terms, (5), are summarized by the mean of the left- and right-swaths (l) and sine and cosine terms (s) as a function of along-track mode (j) on the x-axis. The leading along-track modes ($j = 0$) are small for $i = 0$ and $i = 1$ because these terms are not orthogonal to the bilinear interpolant subtracted in the first stage of the proxy LWC. Overall, the along-track modes are considerably damped compared to the unweighted projection.

The first stage is designed to mimic the alignment of the observed SSHA with the very long wavelength SSHA observed by the nadir altimeter constellation. Because the internal tides are not spatially coherent across such planetary scales, this first step hardly alters the observed internal tides, but it is needed to isolate the SSHA to which a scale-dependent filtering is subsequently applied. This component of the proxy LWC is represented with a least-squares projection onto a bilinear interpolant,

$$\boldsymbol{\eta}_1 = \boldsymbol{\eta}_0 - \mathbf{V}^T\mathbf{V}\boldsymbol{\eta}_0 \tag{2}$$

where $\mathbf{V}$ consists of orthonormalized basis elements spanning the set,

$$\mathbf{v}_0 = 1, \quad \mathbf{v}_1 = x, \quad \mathbf{v}_2 = y, \quad \mathbf{v}_3 = xy, \tag{3}$$

on all valid pixels in the left- and right swaths. $\boldsymbol{\eta}_1$ denotes the modified SSHA which is used in the second stage of the proxy LWC.

The second stage of the proxy LWC is a weighted least-squares projection,

$$\boldsymbol{\eta} = \boldsymbol{\eta}_1 - \mathbf{U}^T\boldsymbol{\Sigma}\mathbf{U}\boldsymbol{\eta}_1, \tag{4}$$

where $\mathbf{U}$ consists of M orthonormalized basis functions which span the set,

$$u_{ijls}(x,y) = \begin{cases} x^i(H(x-x_{mid})-l)\cos(k_j y), & \text{for} \quad s=0 \\ x^i(H(x-x_{mid})-l)\sin(k_j y), & \text{for} \quad s=1. \end{cases} \tag{5}$$

The basis functions are defined in terms of four indexes: $i \in \{0,1,2\}$ indexes the structure of the across-track polynomial; $j \in \{0,\ldots,N\}$ indexes the along-track wavenumber, whether cosine or sine, $s \in \{0,1\}$; and $l \in \{0,1\}$ indexes the left and right sides of the swath, where $H(x)$ is the Heaviside function. Taking $L \approx 5000$ km as the length of the 45° pass segment under consideration, the along-track wavenumbers are discretized as, $k_j = j\pi/L$ for $j = 0,\ldots,P$ where $P = 2L/(1000 \text{ km})$. In total there are $M = 6(2P-1)$ spatial basis vectors (rows of $\mathbf{U}$), denoted $\mathbf{u}_m$ where m is a multi-index, $\{ijls\}$. $\boldsymbol{\Sigma}$ is a diagonal $M \times M$ matrix of weight coefficients.

Let $\delta\boldsymbol{\eta}$ denote the observed L3 LWC with large-scale mean removed as in (2), and let $\delta\hat{\boldsymbol{\eta}} = \mathbf{U}^T\boldsymbol{\Sigma}\mathbf{U}\boldsymbol{\eta}_1$ denote the proxy LWC. The weight matrix Σ is chosen to optimize agreement between the proxy LWC and the actual L3 LWC by minimizing,

$$J(\boldsymbol{\Sigma}) = \langle(\delta\boldsymbol{\eta} - \delta\hat{\boldsymbol{\eta}})^T(\delta\boldsymbol{\eta} - \delta\hat{\boldsymbol{\eta}})\rangle, \tag{6}$$

where the angle-brackets denote the time average. In other words, the proxy LWC is a minimum-variance estimator of the actual L3 LWC. Let σ_m denote the entry of $\boldsymbol{\Sigma}$ corresponding to $\mathbf{u}_m$, then the minimizer of (6) is given by,

$$\sigma_m = \frac{\langle(\mathbf{u}_m^T\boldsymbol{\eta}_1)(\mathbf{u}_m^T\delta\boldsymbol{\eta})\rangle}{\langle(\mathbf{u}_m^T\boldsymbol{\eta}_1)^2\rangle}. \tag{7}$$

The weight coefficients σ_m have been determined for all 28 SWOT orbit passes in orbit segments of 45° arclength, $\{(-45°, 0°), (-22.5°, 22.5°), (0°, 45°)\}$. This subset of pass segments is sufficient to evaluate the impact of the correction on baroclinic tidal signals. In general, the optimal weight coefficients exhibit scale-dependence compared to simple unweighted projection onto the basis (the unweighted projection would correspond to $\sigma_m = 1$), as shown in Figure 4. The weight coefficients vary between passes due to the geographically-dependent partition of observed SSHA variance between sea level variability and long-wavelength errors (Figure 5).

4 Results: Performace of the Proxy LWC

Using the proxy LWC it is possible to assess the approximate linear response of the L3 LWC to geophysical signals of interest such as the internal tides.

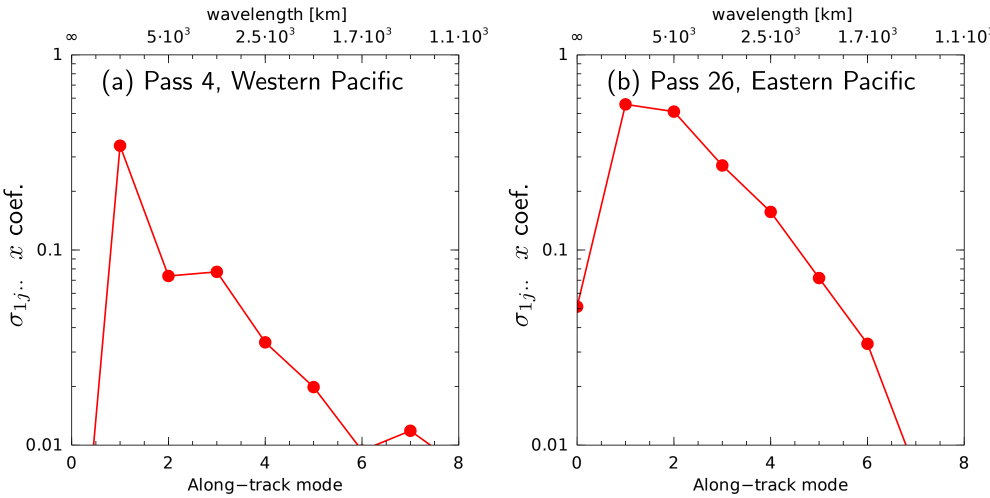

Figure 5. Mean proxy LWC weights for the roll coefficients ($\sigma_{1j\cdot\cdot}$) for two pass segments $(0°, 45°)$ from (a) the Western Pacific and (b) the Eastern Pacific illustrate how the different levels of mesoscale SSHA variability in these regions influence the optimal weight coefficients.

The overall performance of the proxy LWC is shown in terms of the spatial variogram in Figure 6 when the proxy LWC is applied to L2 data. Because the L2 and L3 data are both available for the same time period on the same reference grid, in practice there is little reason to use the proxy LWC on the L2 data; however, the L2 and L3 products differ in regard to how valid data have been flagged, so there may be some specialized applications where the proxy LWC can be usefully applied to the L2 data. In any case, the variogram of the L2 data with the proxy LWC is nearly isotropic and it closely approximates the L3 variogram (Figure 6).

The linear amplitude response of the proxy LWC is evaluated in Figure 7 for pass 15 in the Central Pacific $(0° - 45°)$, a representative example. To compute the response curves, the proxy LWC is applied to an idealized test wave propagating along the azimuthal direction indicated (x-axis) with a wavelength of (a) 150 km or (b) 300-km. The three curves show the impact of the unweighted projection (blue), the proxy LWC (red), and the larger-scale bilinear interpolant (black). Pass 15 is oriented at approximately 78° azimuth, and waves propagating in this direction are filtered the least. Overall, the proxy LWC (and large-scale bilinear interpolant) absorb an insignificant amount of variance from the test waves, except in a narrow range of propagation directions perpendicular to the pass. The simplied unweighted projection absorbs about 10 times more variance than the weighted proxy LWC, except within a band perpendicular to the pass where its performance is much worse.

The linear amplitude response curves for the proxy LWC in Figure 7 indicate that the LWC ought to absorb a negligible fraction of variance of typical baroclinic tidal waves (semidiurnal waves have a wavelength of about 150 km throughout the tropics and mid-latitudes). The amplitude response function does not capture phase errors and may not be the best metric for evaluating the impact of the LWC on the complex wavefield of the tides. A potentially more relevant assessment is provided by appling the proxy LWC to the predicted baroclinic tides, where the latter are provided by the High Resolution Empirical Tide model (HRET, Zaron (2019)), as provided by the SWOT products.

Figure 8a shows a snapshot of the tidal prediction for a representative example, pass 28 $(-45°, 0°)$, from a region with relatively large internal tides. Note the extreme dis-

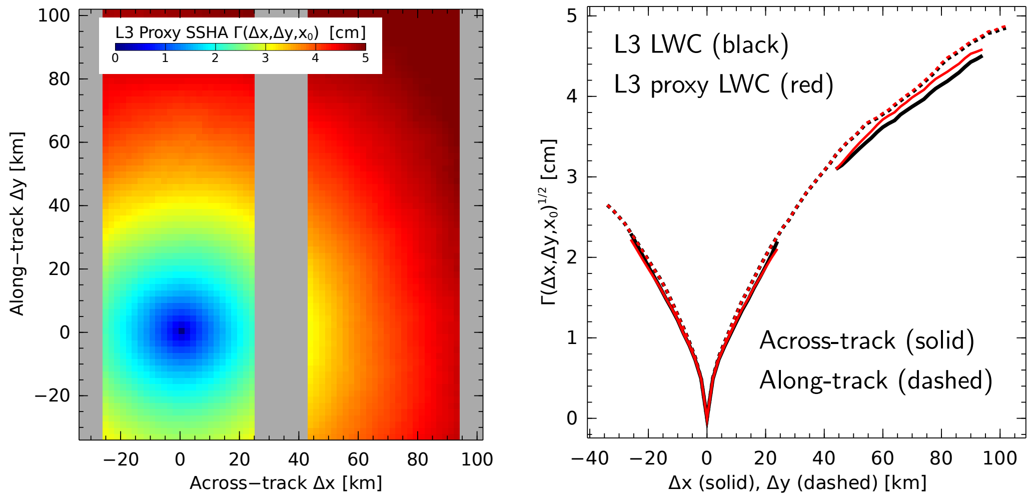

Figure 6. Left panel: The two-dimensional variogram (equation (1)) for the L2 data computed using the proxy LWC is nearly isotropic (cf., Figure 2a). Right panel: Slices through the variogram show that statistical structure of the proxy-LWC-corrected L2 data closely agrees with that of the (nonlinear) LWC-corrected L3 data. The data in this figure are taken from pass 15 in the Central Pacific ($0° - 45°$).

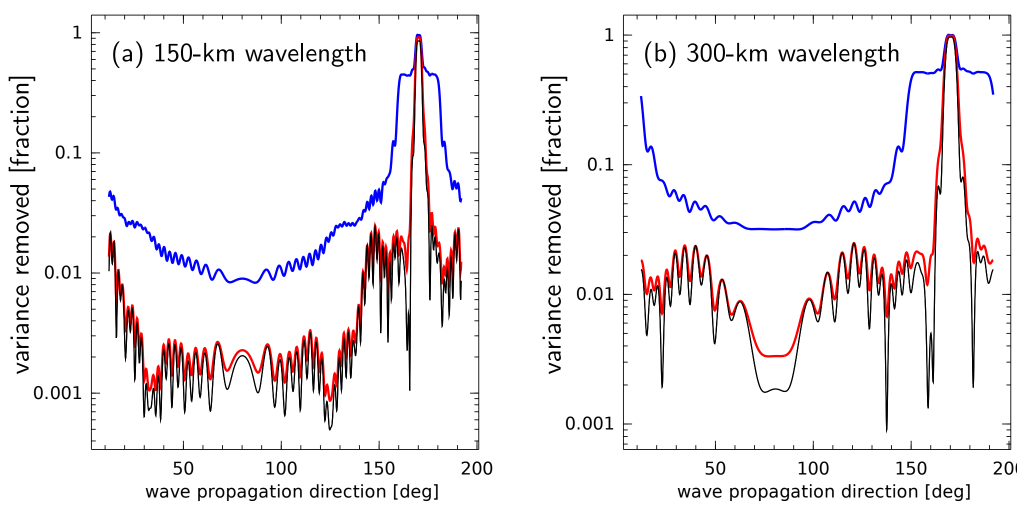

Figure 7. The amplitude response of the proxy LWC (red line) compared with an unweighted projection ($\sigma_m = 1$, blue line), and the large-scale bilinear interpolation (the first stage of the proxy LWC, black line). The response is evaluated with test waves of (a) 150-km wavelength and (b) 300-km wavelength, representative of the wavelengths of mode-1 semidiurnal and diurnal internal tides throughout the tropics and subtropics. The propagation direction of the test waves is indicated on the x-axis; note that the amplitude response is minimal for waves propagating in the same direction as the pass, 78°, and maximal for waves propagating perpendicular to the pass.

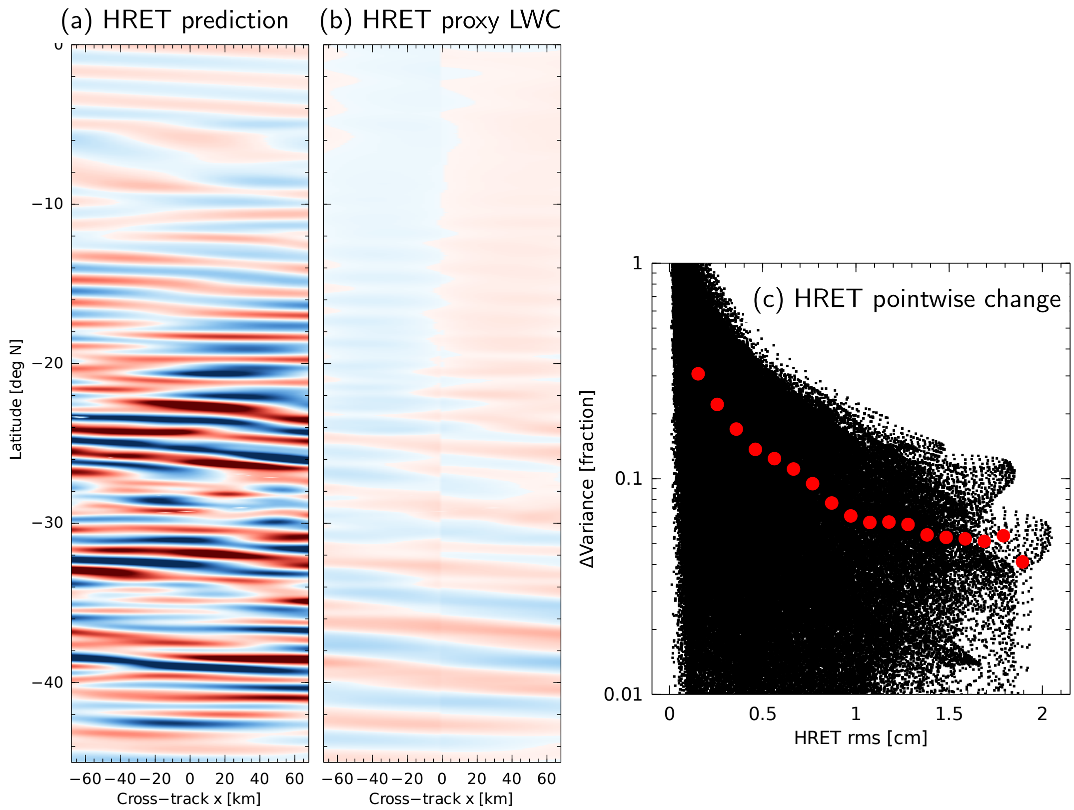

Figure 8. Impact of the proxy LWC on the HRET tide prediction (pass 28, $-45°, 0°$; a representative example). (a) A snapshot of the predicted internal tide SSHA; ± 1.5 cm colorscale range. (b) The proxy LWC from the HRET field given in panel (a); same ± 1.5 cm colorscale range. (c) The fractional variance of the proxy LWC applied to HRET at every pixel of the swath, averaged over all available orbit cycles. When the variance change is binned over ranges of HRET rms amplitude (red dots), the variance change exceeds 10% for pixels with amplitude less than about 0.8 cm. For larger amplitude tides, the error is approximately 5% to 10%.

tortion of the across- and along-track aspect ratio in the Figure, a factor of 50. The impact of the proxy LWC correction is visually small compared to the predicted tide (Figure 8b). In Figure 8c the variance absorbed by the proxy LWC has been computed pointwise over the pass segment and plotted as a function of the internal tide amplitude, averaging over all available orbit cycles. It is apparent that the phase errors caused by the LWC are not completely negligible and lead to changes of 10% or more for pixels with r.m.s. tidal amplitude less than about 0.8 cm. Pixels with larger tidal signals are influenced less, but errors on the scale of 5% to 10% may be anticipated for centimeter-amplitude baroclinic tidal waves in corrected L3 SWOT data.

5 Conclusions

A linear approximation of the nonlinear, data-driven, L3 LWC has been developed for SWOT data during the one-day repeat orbit phase of the mission. This correction, the so-called "proxy LWC", was implemented as a minimum-variance estimator of the L3 LWC. It is constructed from a set of spatial basis functions describing a plausible subset of possible long-wavelength errors. The proxy LWC was developed soley to characterize, approximately, the linear response of the L3 LWC and to quantify the extent to which the L3 LWC might remove baroclinic ocean tidal signals from the SSHA. The proxy LWC could be applied to other oceanographic fields or SWOT SSHA corrections; or it could be used as a LWC for the L2 data in specialized applications where the L3 product is, for whatever reason, not satisfactory.

The proxy LWC was shown to lead to corrected SSHA which closely mimics the spatial covariance properties of the L3 product. Of course, in the L3 product the large-scale SSHA is aligned with data from the existing nadir altimeter constellation, while the proxy-LWC-corrected L2 data would not be aligned with large-scale sea level. Furthermore, the L3 LWC contains a sophisticated model for the temporal structure of the LWC, a feature which is absent from the proxy LWC. Users of the proxy LWC should keep in mind these limitations if they attempt to use it with L2 data.

The proxy LWC was applied to both idealized and realistic waveforms representative of the low-mode baroclinic tidal SSHA. When measured in terms of a spatial filter response function, the proxy LWC generally absorbs a small fraction of variance from such waveforms. In particular, for most propagation directions, the proxy LWC absorbs about a factor of 10 less signal than would be absorbed by a simpler, unweighted, projection onto the LWC basis functions. For realistic tidal SSHA, the proxy LWC can lead to typical errors of 5% to 10% for waves with amplitude of a centimeter or larger. Overall, it is concluded that the L3 LWC should have a small, mostly insignificant, impact on baroclinic tidal signals,

Appendix A Interpretation of $\boldsymbol{\sigma_m}$ as weight coefficients

Regard the SSHA observations, $\boldsymbol{\eta}_0$, as the sum of a LWC, $\hat{\boldsymbol{\eta}}$, and an oceanographic signal, $\boldsymbol{\eta}_0 - \hat{\boldsymbol{\eta}}$. Assuming Gaussian statistics and Bayesian priors for their variances, $\mathbf{E}_{LWC}$ and $\mathbf{R}$, respectively, the maximum likelihood estimator of $\hat{\boldsymbol{\eta}}$ is the minimizer of,

$$J(\hat{\boldsymbol{\eta}}) = \hat{\boldsymbol{\eta}}^T \mathbf{E}_{LWC}^{-1} \hat{\boldsymbol{\eta}} + (\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0)^T \mathbf{R}^{-1} (\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0). \tag{A1}$$

If it is assumed that $\hat{\boldsymbol{\eta}}$ can be represented as a linear combination of orthonormalized basis vectors, $\hat{\boldsymbol{\eta}} = \mathbf{U}^T \boldsymbol{\alpha}$, then J can be expressed in terms of $\boldsymbol{\alpha}$ as

$$J(\boldsymbol{\alpha}) = \boldsymbol{\alpha}^T \mathbf{E}^{-1} \boldsymbol{\alpha} + (\mathbf{U}^T \boldsymbol{\alpha} - \boldsymbol{\eta}_0)^T \mathbf{R}^{-1} (\mathbf{U}^T \boldsymbol{\alpha} - \boldsymbol{\eta}_0), \tag{A2}$$

where $\mathbf{E} = \mathbf{U} \mathbf{E}_{LWC} \mathbf{U}^T$.

The optimal estimate of $\boldsymbol{\alpha}$ satistfies,

$$\frac{1}{2} \frac{\partial J}{\partial \boldsymbol{\alpha}} = \mathbf{E}^{-1} \boldsymbol{\alpha} + \mathbf{U} \mathbf{R}^{-1} (\mathbf{U}^T \boldsymbol{\alpha} - \boldsymbol{\eta}_0) = \mathbf{0}, \tag{A3}$$

or

$$(\mathbf{E}^{-1} + \mathbf{U} \mathbf{R}^{-1} \mathbf{U}^T) \boldsymbol{\alpha} = \mathbf{U} \mathbf{R}^{-1} \boldsymbol{\eta}_0. \tag{A4}$$

Suppose the oceanographic variability is uncorrelated and homogeneous, $\mathbf{R} = r\mathbf{I}$,

$$(\mathbf{E}^{-1} + r^{-1} \mathbf{I}) \boldsymbol{\alpha} = r^{-1} \mathbf{U} \boldsymbol{\eta}_0, \tag{A5}$$

and assume $\boldsymbol{E}$ is diagonal with entries, e_m, then the optimizing coeficients of $\boldsymbol{\alpha}$ may be expressed simply as

$$\alpha_m = \frac{r^{-1}}{e_m^{-1} + r^{-1}} u_m^T \boldsymbol{\eta}_0 = \frac{e_m}{r + e_m} u_m^T \boldsymbol{\eta}_0. \tag{A6}$$

Finally, the coefficient vector $\boldsymbol{\alpha}$ can be elimated in favor of $\boldsymbol{\Sigma}$, and the LWC can be written as,

$$\hat{\boldsymbol{\eta}} = \mathbf{U}^T \boldsymbol{\Sigma} \mathbf{U} \boldsymbol{\eta}_0, \tag{A7}$$

where the weights, σ_m, used in the text are given by,

$$\sigma_m = \frac{e_m}{r + e_m}. \tag{A8}$$

The weighting coefficients can be understood in relation to Bayesian priors for the variance of the LWC basis functions (e_m) and the oceanographic component of the SSHA variance (r). The expectation is that σ_m will tend to be smaller in regions of larger oceanic SSHA variance. Indeed, the example of shown in Figure 5 agrees; the σ_m coefficients are smaller in the Western Pacific, a region of enhanced mesoscale eddy activity due to its proximity to the western boundary current.

Open Research Section

The Level-2 SWOT data are publicly available from the NASA Jet Propulsion Laboratory PO.DAAC through the Earthdata website (`https://search.earthdata.nasa.gov`) using the PO.DAAC data-subscriber tool (`https://github.com/podaac/data-subscriber`). The Level-3 SWOT data are available from the AVISO website, with support from CNES, at `https://www.aviso.altimetry.fr/en/data/products/sea-surface-height-products/global/swot-l3-ocean-products.html`. NetCDF-format files containing the weighting coefficients, σ_m, defining the proxy LWC (2)-(7) are available from the Zenodo website `doi:10.5281/zenodo.10914546`.

Acknowledgments

The research reported in this manuscript was supported by NASA award 80NSSC21K0346 to Oregon State University.

References

Dibarboure, G., & Ubelmann, C. (2014). Investigating the performance of four empirical cross-calibration methods for the proposed SWOT mission. *Remote Sens.*, *6*(6), 4831–4869. doi: 10.3390/rs6064831

Dibarboure, G., Ubelmann, C., Flamant, B., Briol, F., Peral, E., Bracher, G., ... Picot, N. (2022). Data-driven calibration algorithm and pre-launch performance simulations for the SWOT mission. *Remote Sens.*, *14*(23), 6070. doi: 10.3390/rs14236070

Esteban-Fernandez, D., Rodriguez, E., Fu, L.-L., Alsdorf, D., & Vaze, P. (2010). The Surface Water and Ocean Topography Mission: centimetric spaceborne radar interferometry. In R. Meynart, S. P. Neeck, & H. Shimoda (Eds.), *Sensors, Systems, and Next-Generation Satellites XIV* (Vol. 7826, p. 782615). SPIE. doi: 10.1117/12.868535

Fjørtoft, R., Gaudin, J.-M., Pourthie, N., Lion, C., Mallet, A., Souyris, J.-C., ... Arnaud, A. (2010). KaRIn - the Ka-band radar interferometer on SWOT: Measurement principle, processing and data specificities. In *2010 IEEE International Geoscience and Remote Sensing Symposium* (p. 4823-4826). doi: 10.1109/IGARSS.2010.5650601

Morrow, R., Fu, L.-L., Ardhuin, F., Benkiran, M., Chapron, B., Cosme, E., ... Zaron, E. D. (2019). Global observations of fine-scale ocean surface topography with the Surface Water & Ocean Topography (SWOT) mission. *Front. Mar. Sci.*, *6*, 232.

Zaron, E. D. (2019). Baroclinic tidal sea level from exact-repeat mission altimetry. *J. Phys. Oceanogr.*, *49*(1), 193–210.

The Significance of the Long-Wavelength Correction for Studies of Baroclinic Tides with SWOT

Edward D. Zaron[1]

[1]College of Earth, Ocean, and Atmospheric Science, Oregon State University

Key Points:

- The long-wavelength correction has the potential to remove geophysical signals of interest
- The correction is implemented as a nonlinear data-driven filter, making it difficult to characterize its response using linear techniques
- A linear approximation of the filter is developed here and used to characterize the response to baroclinic tides in the open ocean

Corresponding author: Edward D. Zaron, Edward.D.Zaron@oregonstate.edu

Abstract

The long-wavelength correction (LWC) of SWOT data is intended to reduce errors related to the stability of the SWOT antenna and its attitude in orbit, especially those errors related to the roll of the satellite. The algorithms used to compute the LWC utilize SWOT KaRIn sea surface-height (SSH) measurements and additional data, and the LWC may abosrb geophysical SSH into the correction. Different LWC algorithms are used on the L2 and L3 SWOT products, which are analyzed here during the 1-day repeat (Cal/Val) mission phase lasting approximately 100 days. During this mission phase the SSH anomaly (SSHA) computed using the L3 LWC is much more realistic than the L2 LWC, shown here by comparing spatial statistics of the L2 and L3 products. The L3 LWC algorithm is nonlinear insofar as it depends on second-order statistics of the SSHA and multi-satellite SSHA differences, making it difficult to quantify the extent to which it could absorb baroclinic tidal signals. To overcome this difficulty, a proxy L3 LWC algorithm is developed which mimics the L3 LWC but is strictly linear in the SSHA. The proxy LWC is applied to the predicted internal tide available on the products, and it is found to absorb roughly 5% to 10% of the variance of the internal tide; although, this figure varies strongly depending on the magnitude and orientation of the tidal waves.

Plain Language Summary

In order to make the SWOT data useful for studies of sea level associated with horizontal length scales of roughly 100 km and larger, it is necessary to compute a correction which aligns the measured SSH with independent data from other satellite missions. The correction is implemented with a type of data-driven spatial low-pass filter; however, the data-driven character of this filter means that it is nonlinear and its response cannot be characterized using standard technqiues of linear filter analysis. In this paper a linear approximation of the long-wavelength correction is developed which is amenable to standard linear analysis techniques. Using this linear approximation, the extent to which the long-wavelength correction may absorb signals of interest arising from the baroclinic tides is quantified. It is found that the filter response is generally below 10% of the signal variance, which could lead to small but non-negligible errors in studies of tides using data from the Cal/Val mission phase.

1 Introduction

The SWOT KaRIn instrument is an imaging interferometer(Fjørtoft et al., 2010). Because its orbit altitude is very much larger than the KaRIn antenna baseline, SWOT sea surface-height (SSH) measurements are sensitive to uncontrolled perturbations in the attitude of the instrument. The SSH measurements are also influenced by electrical path delays, thermal effects, and mechanical stability of the KaRIn antenna. Because these factors typically evolve slowly during the orbit of the satellite, they are associated with long-wavelength errors in the SSH. The SWOT mission objectives for observation of oceanic SSH place strict requirements on the error budget for observations at scales from 150 km to 4 km(Esteban-Fernandez et al., 2010; Morrow et al., 2019), and these are largely independent of the long-wavelength errors. Nonetheless, the long-wavelength errors may be significant to observations of certain geophysical phenomena, such as ocean tides, and so corrections have been developed to reduce the long-wavelength errors over the ocean.

Table 1. Tidal alias periods (days) for SWOT during the one-day repeat orbit phase

Mission repeat [days]	Q_1	O_1	P_1	S_1	K_1	N_2	M_2	S_2	K_2	M_4	MS_4
0.99349	9	13	108	153	262	9	12	76	131	6	11

This manuscript analyzes the long-wavelength correction (LWC) provided by the presently available SWOT ocean products during the 1-day repeat (Cal/Val) orbit phase. The 1-day repeat phase, which lasted from 2023-03-30 to 2023-07-10, provided nearly 100 days of data. The duration of this orbit phase is sufficient to estimate several of the larger tides using harmonic analysis (Table 1), and the daily repeats may contain useful information about the non-phase-locked or modulated tides which are not available from other data sources. Thus, in spite of the relatively limited geographic coverage (Figure 1), there is interest in understanding the usefulness of these data for tidal studies.

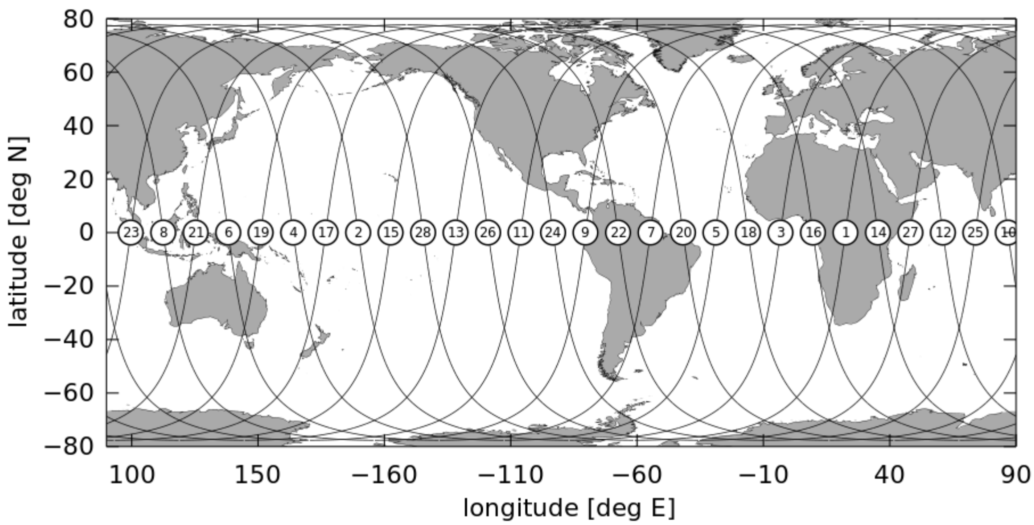

Figure 1. SWOT ground tracks and pass numbers during the 1-day-repeat orbit phase, from 2023-03-30 to 2023-07-10.

2 SWOT Data Products and Long-Wavelength Corrections (LWC)

Two SWOT data products are used here, namely, (1) the Level 2 Low Rate Sea Surface Height Data Product, version 1.1 (Beta Pre-Validated Product, personal communication, Shailen Desai, 2023-10-19)[1] and (2) the NRT SWOT KaRIn & nadir Global Ocean swath SSALTO/DUACS Sea Surface Height L3 product, version 0.3, CRM:0069553 (personal communication, AVISO, 2024-01-10)[2]. These shall be referred to as the L2 and L3 products, respectively.

The L2 and L3 both products provide data on a fixed geographic grid at a resolution of 2 km, the Low-Resolution grid. The L2 product provides two versions of the SSH anomaly (SSHA). The first version (`ssha_karin`; henceforth, names written in `this font` refer to the names of variables in the SWOT data products) utilizes data from the on-board radiometer and dual-frequency nadir altimeter to compute a suite of geophysical corrections which are applied to the interferometric data to compute SSHA; this will be referred to as L2a SSHA, below. The second version (`ssh_karin_2`) utilizes model-based corrections for the wet path delay and the sea state bias correction to reduce dependence on non-KaRIn SWOT data; this will be referred to as L2b SSHA, below. The L2b product contains a slightly larger quantity of valid SSHA data compared to the L2a product. Both the L2a and L2b products contain an identical suite of corrections for the barotropic

[1] `doi.org/10.24400/527896/A01-2023.015` Accessed 2024-02-05.
[2] `doi.org/10.24400/527896/A01-2023.017` Accessed 2024-02-05.

ocean, load, solid earth, and pole tides; dynamic atmosphere correction; dry troposphere correction; ionospheric path delay; mean sea surface; geoid; baroclinic tides (`internal_tide_hret`), and LWC (`height_cor_xover`). A detailed description of the L2 products is found in the Product Description[3].

The L3 products are derived from the L2 products but employ a different algorithm to compute the LWC (`calibration`), described below. As with the L2 product, there are two versions of the SSHA provided, a default version (`ssha`) and a "noiseless" version (`ssha_noiseless`) which uses a nonlinear data-adaptive filter developed with machine-learning techniques to reduce the small-scale noise of the SSHA product. These two versions shall be referred to as the L3a and L3b SSHA products.

The LWC is a mitigation for the "roll error" which arises from imperfect knowledge of the orientation of the KaRIn antenna relative to the ocean surface. The relative positions of the KaRIn antennas must be known with micrometer position to achieve the centimeter precision of the SWOT SSHA measurement (Dibarboure & Ubelmann, 2014). In practice, this level of precision is not available from the on-orbit information, and there are additional systematic errors associated with the phase-screen and thermo-mechanical design, all of which the LWC is intended to correct. The design goal of the LWC is to remove the time-invariant and slowly evolving signals associated with these errors while leaving unaffected the signals from oceanic variability. In practice, the LWC error is modelled as the sum of bias, linear, and quadratic components in the across-track direction, separately within the left and right sides of each KaRIn swath, and the coefficients describing these components are permitted to vary over time scales of 3 minutes and longer. Because this error evolution timescale corresponds to an along-track length scale of roughly 1000 km, it is referred to as long-wavelength error.

While the L2 and L3 LWC employ the same basic formulation, they rely on different approaches and use different data to estimate the correction (Dibarboure et al., 2022). The L2 algorithm is designed to be used in near-real time, and utilizes only SWOT data. The coefficients for the along- and across-track errors are constrained by minimizing the weighted difference between corrected SSHA in the crossovers between ascending and descending swaths. Away from crossovers, the coefficients are determined by smoothly interpolating them in the along-track direction with a 1000-km Gaussian kernel, constrained by the crossover values.

The L3 algorithm uses more external data than the L2 algorithm. It is designed to be used in non-real-time mode when high-quality information from the constellation of nadir satellite altimeter missions is available. The idea is roughly the same, though, that coefficients for the mean, linear, and quadratic components of the LWC are estimated by optimizing agreement between the corrected SWOT SSHA and other missions where the tracks cross. The various data and smoothness constraints are weighted according to the residual differences between the different sources and the time-offsets between their observations.

One simple assessment of the L2 and L3 LWCs is provided by the spatial variogram of the SSHA.The variogram is computed by averaging over the data from the 1-day-repeat orbit period as,

$$\Gamma^2(\Delta x, \Delta y, x_0) = \langle (\eta(x_0 + \Delta x, y + \Delta y) - \eta(x_0, y))^2 \rangle \tag{1}$$

where $\eta(x, y)$ represents the SSHA at across-track pixel x and along-track pixel y, Δx and Δy represent across- and along-track separation, the angle brackets represent averaging over all orbit cycles in the 1-day-repeat, and x_0 represents a reference pixel in

[3] https://archive.podaac.earthdata.nasa.gov/podaac-ops-cumulus-docs/web-misc/swot_mission_docs/pdd/D-56407_SWOT_Product_Description_L2_LR_SSH_20220902_RevA.pdf Accessed 2024-02-05.

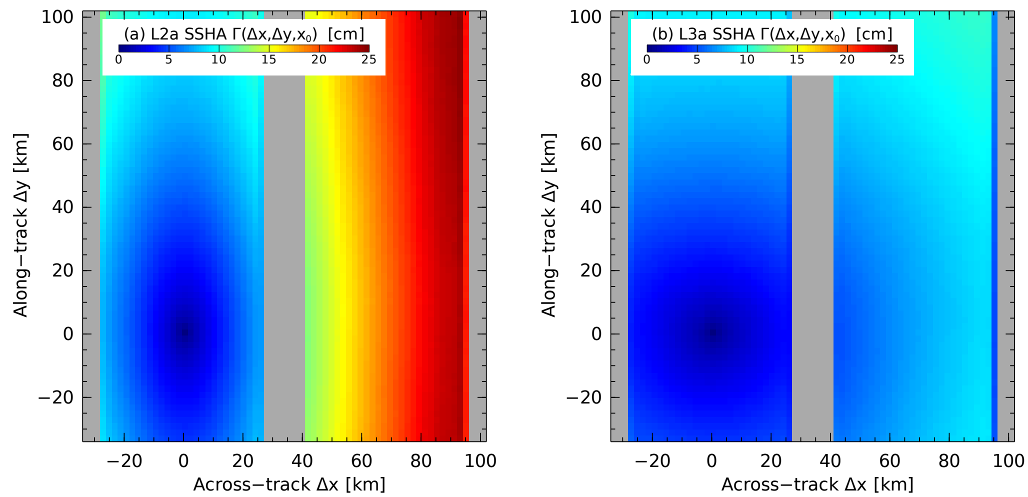

Figure 2. The square root of the spatial variogram ($\Gamma(\Delta x, \Delta y, x_0)$, equation 1) is shown for the (a) L2a and (b) L3a SSHA products. Note that the global minimum of $\Gamma(\Delta x, \Delta y, x_0)$ occurs at $\Delta x = \Delta y = 0$ at the x_0 pixel in the middle of the left swath.

the middle of the left-side swath. The averaging to compute the variogram is restricted to use data from pixels with the best quality flags. Furthermore, averages have been computed over the four latitude ranges, $(-72°, -45°)$, $(-45°, 0°)$, $(0°, 45°)$, $(45°, 72°)$, to look for differences related to the winter/summer hemispheres; however, only the results for the $(0°, 45°)$ range are shown here as no significant differences between the regions were noted.

A comparison of the L2a and L3a SSHA variograms is shown in Figure 2, and it reveals the large-scale structure which is a consequence of the different LWC algorithms (note that Γ is shown, units of cm). The magnitude of the L3a variogram is smaller than that of the L2a variogram, which is consistent with a more accurate LWC for the L3 product and consistent with the larger amount of data involved in computing the L3 LWC. The L2a variogram is notably anisotropic and it also exhibits a jump from one side of the swath to the other, both of which are unrealistic features. Both products also exhibit unrealistic variogram structures in the outermost and innermost pixels on each side; this is apparently related to how SSHA is interpolated onto the geographically fixed grid (Clement Ubelmann, personal communication).

The isotropy of the L3 variogram is quite remarkable, considering that the LWC is modeled with separate parameterizations in the along- and across-track directions. The differences between the L2 and L3 products and their “a” and “b” versions are revealed in cross-sections through the Γ function in Figure 3. The L2b product contains slightly more variance than the L2a product, visible at the largest Δx and Δy as larger values for the red versus the black lines in Figure 3a. In other words, the model-based corrections are slightly noisier than the radiometer-based corrections, as would be expected. In contrast, the L3a and L3b products only differ significantly at scales smaller than about 20 km, and the along- and across-track partitioning of the variance is very nearly equal (Figure 3b). The features noted are consistent with the intended performance of the L3 LWC and small-scale de-noising algorithms.

From the variogram it is not possible to assess whether the LWC is absorbing any ocean signals. The comparisons simply reveal that the L3 product exhibits a plausible rendition of the spatial covariance of ocean surface topography and the L2 product does

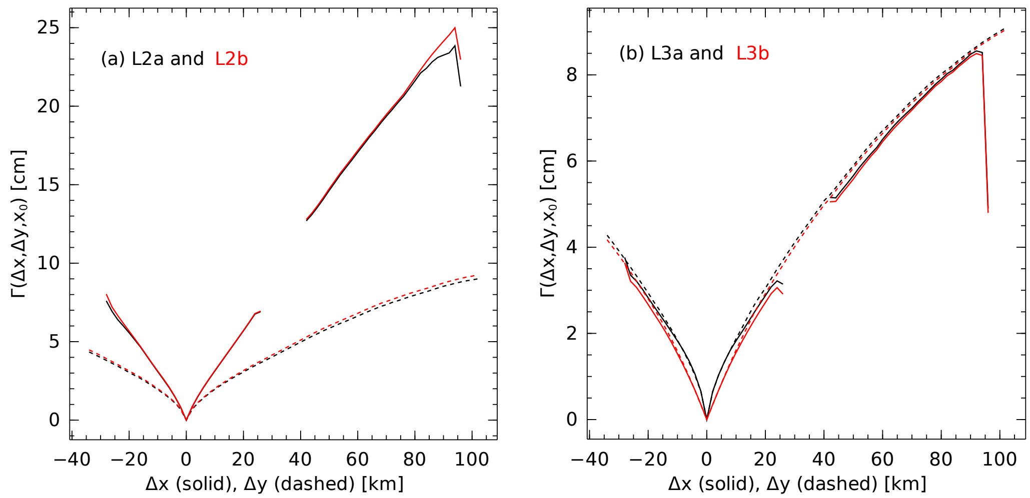

Figure 3. Cross-sections through $\Gamma(\Delta x, \Delta y, x_0)$ in the across-track (Δx, solid lines) and along-track (Δy, dashed lines) reveal the anisotropy of of the L2 products. The panels compare (a) the L2a and L2b products (black and red, respectively) and (b) the L3a and L3b products. Note the very different vertical scales used in each panel.

not. Recall that only data from the 1-day-repeat orbit phase are used here. Due to the sparse crossovers during this orbit phase, the findings shown in Figures 2a and 3a are not unexpected.

Details of the above L2 and L3 algorithms are provided in Dibarboure et al. (2022); however, the algorithms are implemented using weighted least-squares, with the weights derived from covariances estimated using crossover residuals. The data-driven character of the algorithms introduces nonlinearity; therefore, the LWC is not a simple linear filter acting on the measured data. This aspect of the LWC makes it impossible to assess whether it could distort or filter out signals of real geophysical processes, such as the baroclinic tides.

3 A Proxy LWC for Assessing the L3 LWC Using L2 Data

This section describes the development of a linear approximation to the L3 LWC which will be called the "proxy LWC."

The notation used in this section expresses the observed SSHA, $\eta(x, y)$—a function of cross-swath coordinate x and along-swath coordinate y—using the boldface vector notation, $\boldsymbol{\eta}$. The notation connotes that $\boldsymbol{\eta} \in R^N$ is regarded as a 1-dimensional vector containing the N values of $\eta(x, y)$ at each valid pixel for a given orbit cycle. The presence of missing data, which could be caused by environmental effects or instrument anomalies, causes the number of valid pixels, N, to vary in time. Linear operators which act on $\boldsymbol{\eta}$ will be written in matrix notation as bold capital letters, e.g., $\mathbf{U}$, where $\mathbf{U} \in R^{M \times N}$. The rank M will generally be unstated but implied by the context or dimension of the vectors comprising $\mathbf{U}$. When $\mathbf{U}$ consists of M ortho-normalized basis vectors on R^N, the notation, $\mathbf{U}^T\mathbf{U}\boldsymbol{\eta}$, shall be used to denote the projection of $\boldsymbol{\eta}$ onto this M-dimensional subspace.

The proxy LWC is a linear operator acting on the uncorrected SSHA, $\boldsymbol{\eta}_0$. It consists of two stages which are applied to 5000-km segments of SWOT data.

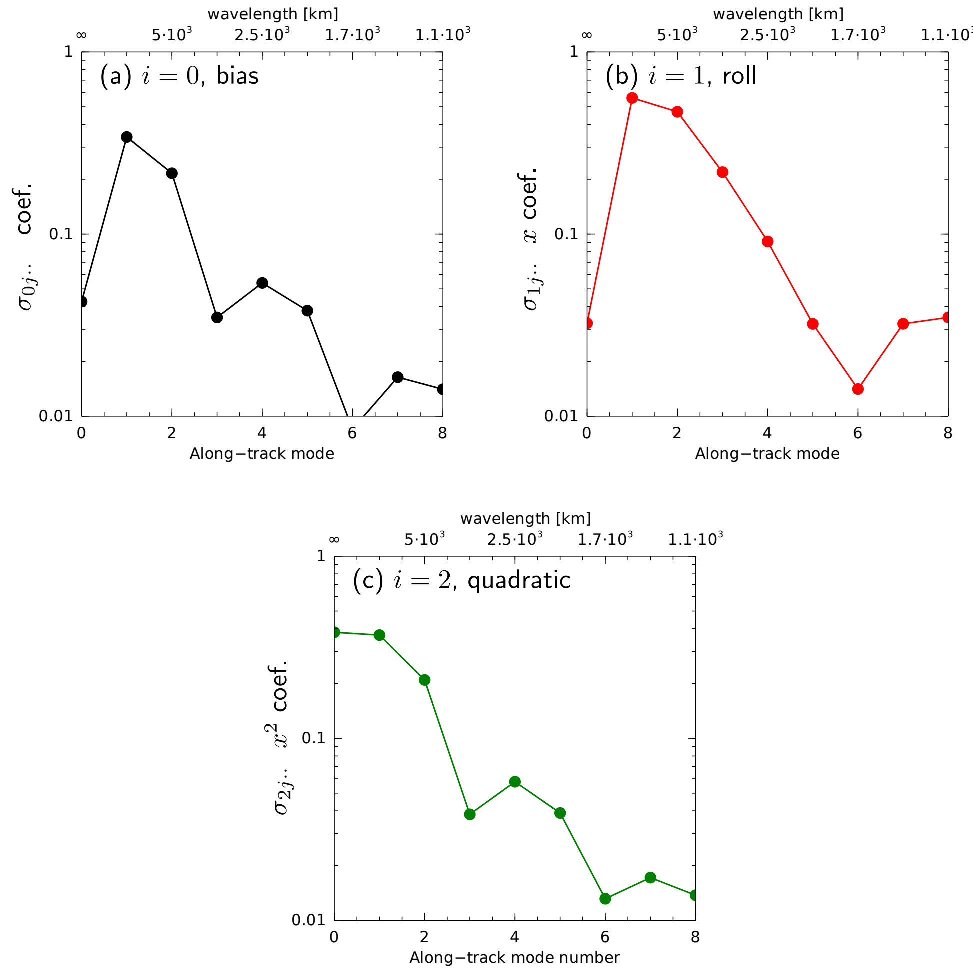

Figure 4. Proxy LWC weight coefficients, σ_m (equation (7)), for a pass #15 across the central Pacific. The coefficients of the (a) bias ($i = 0$), (b) roll ($i = 1$), and (c) quadratic ($i = 2$) terms, (5), are summarized by the mean of the left- and right-swaths (l) and sine and cosine terms (s) as a function of along-track mode (j) on the x-axis. The leading along-track modes ($j = 0$) are small for $i = 0$ and $i = 1$ because these terms are not orthogonal to the bilinear interpolant subtracted in the first stage of the proxy LWC. Overall, the along-track modes are considerably damped compared to the unweighted projection.

The first stage is designed to mimic the alignment of the observed SSHA with the very long wavelength SSHA observed by the nadir altimeter constellation. Because the internal tides are not spatially coherent across such planetary scales, this first step hardly alters the observed internal tides, but it is needed to isolate the SSHA to which a scale-dependent filtering is subsequently applied. This component of the proxy LWC is represented with a least-squares projection onto a bilinear interpolant,

$$\boldsymbol{\eta}_1 = \boldsymbol{\eta}_0 - \mathbf{V}^T\mathbf{V}\boldsymbol{\eta}_0 \tag{2}$$

where $\mathbf{V}$ consists of orthonormalized basis elements spanning the set,

$$\mathbf{v}_0 = 1, \quad \mathbf{v}_1 = x, \quad \mathbf{v}_2 = y, \quad \mathbf{v}_3 = xy, \tag{3}$$

on all valid pixels in the left- and right swaths. $\boldsymbol{\eta}_1$ denotes the modified SSHA which is used in the second stage of the proxy LWC.

The second stage of the proxy LWC is a weighted least-squares projection,

$$\boldsymbol{\eta} = \boldsymbol{\eta}_1 - \mathbf{U}^T\boldsymbol{\Sigma}\mathbf{U}\boldsymbol{\eta}_1, \tag{4}$$

where $\mathbf{U}$ consists of M orthonormalized basis functions which span the set,

$$u_{ijls}(x, y) = \begin{cases} x^i(H(x - x_{mid}) - l)\cos(k_j y), & \text{for} \quad s = 0 \\ x^i(H(x - x_{mid}) - l)\sin(k_j y), & \text{for} \quad s = 1. \end{cases} \tag{5}$$

The basis functions are defined in terms of four indexes: $i \in \{0, 1, 2\}$ indexes the structure of the across-track polynomial; $j \in \{0, \ldots, N\}$ indexes the along-track wavenumber, whether cosine or sine, $s \in \{0, 1\}$; and $l \in \{0, 1\}$ indexes the left and right sides of the swath, where $H(x)$ is the Heaviside function. Taking $L \approx 5000$ km as the length of the 45° pass segment under consideration, the along-track wavenumbers are discretized as, $k_j = j\pi/L$ for $j = 0, \ldots, P$ where $P = 2L/(1000 \text{ km})$. In total there are $M = 6(2P-1)$ spatial basis vectors (rows of $\mathbf{U}$), denoted $\mathbf{u}_m$ where m is a multi-index, $\{ijls\}$. $\boldsymbol{\Sigma}$ is a diagonal $M \times M$ matrix of weight coefficients.

Let $\delta\boldsymbol{\eta}$ denote the observed L3 LWC with large-scale mean removed as in (2), and let $\delta\hat{\boldsymbol{\eta}} = \mathbf{U}^T\boldsymbol{\Sigma}\mathbf{U}\boldsymbol{\eta}_1$ denote the proxy LWC. The weight matrix Σ is chosen to optimize agreement between the proxy LWC and the actual L3 LWC by minimizing,

$$J(\boldsymbol{\Sigma}) = \langle(\delta\boldsymbol{\eta} - \delta\hat{\boldsymbol{\eta}})^T(\delta\boldsymbol{\eta} - \delta\hat{\boldsymbol{\eta}})\rangle, \tag{6}$$

where the angle-brackets denote the time average. In other words, the proxy LWC is a minimum-variance estimator of the actual L3 LWC. Let σ_m denote the entry of $\boldsymbol{\Sigma}$ corresponding to $\mathbf{u}_m$, then the minimizer of (6) is given by,

$$\sigma_m = \frac{\langle(\mathbf{u}_m^T\boldsymbol{\eta}_1)(\mathbf{u}_m^T\delta\boldsymbol{\eta})\rangle}{\langle(\mathbf{u}_m^T\boldsymbol{\eta}_1)^2\rangle}. \tag{7}$$

The weight coefficients σ_m have been determined for all 28 SWOT orbit passes in orbit segments of 45° arclength, $\{(-45°, 0°), (-22.5°, 22.5°), (0°, 45°)\}$. This subset of pass segments is sufficient to evaluate the impact of the correction on baroclinic tidal signals. In general, the optimal weight coefficients exhibit scale-dependence compared to simple unweighted projection onto the basis (the unweighted projection would correspond to $\sigma_m = 1$), as shown in Figure 4. The weight coefficients vary between passes due to the geographically-dependent partition of observed SSHA variance between sea level variability and long-wavelength errors (Figure 5).

4 Results: Performace of the Proxy LWC

Using the proxy LWC it is possible to assess the approximate linear response of the L3 LWC to geophysical signals of interest such as the internal tides.

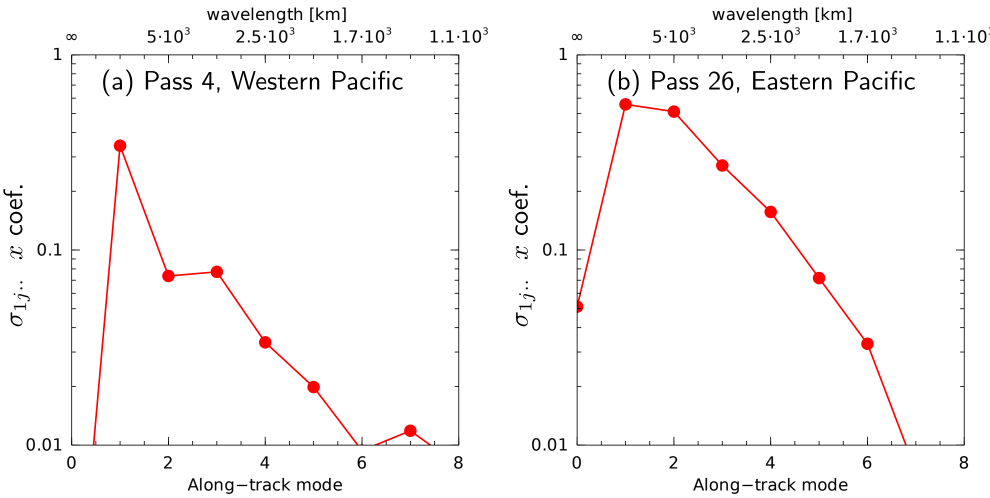

Figure 5. Mean proxy LWC weights for the roll coefficients ($\sigma_{1j\cdot\cdot}$) for two pass segments $(0°, 45°)$ from (a) the Western Pacific and (b) the Eastern Pacific illustrate how the different levels of mesoscale SSHA variability in these regions influence the optimal weight coefficients.

The overall performance of the proxy LWC is shown in terms of the spatial variogram in Figure 6 when the proxy LWC is applied to L2 data. Because the L2 and L3 data are both available for the same time period on the same reference grid, in practice there is little reason to use the proxy LWC on the L2 data; however, the L2 and L3 products differ in regard to how valid data have been flagged, so there may be some specialized applications where the proxy LWC can be usefully applied to the L2 data. In any case, the variogram of the L2 data with the proxy LWC is nearly isotropic and it closely approximates the L3 variogram (Figure 6).

The linear amplitude response of the proxy LWC is evaluated in Figure 7 for pass 15 in the Central Pacific $(0° - 45°)$, a representative example. To compute the response curves, the proxy LWC is applied to an idealized test wave propagating along the azimuthal direction indicated (x-axis) with a wavelength of (a) 150 km or (b) 300-km. The three curves show the impact of the unweighted projection (blue), the proxy LWC (red), and the larger-scale bilinear interpolant (black). Pass 15 is oriented at approximately 78° azimuth, and waves propagating in this direction are filtered the least. Overall, the proxy LWC (and large-scale bilinear interpolant) absorb an insignificant amount of variance from the test waves, except in a narrow range of propagation directions perpendicular to the pass. The simplied unweighted projection absorbs about 10 times more variance than the weighted proxy LWC, except within a band perpendicular to the pass where its performance is much worse.

The linear amplitude response curves for the proxy LWC in Figure 7 indicate that the LWC ought to absorb a negligible fraction of variance of typical baroclinic tidal waves (semidiurnal waves have a wavelength of about 150 km throughout the tropics and mid-latitudes). The amplitude response function does not capture phase errors and may not be the best metric for evaluating the impact of the LWC on the complex wavefield of the tides. A potentially more relevant assessment is provided by appling the proxy LWC to the predicted baroclinic tides, where the latter are provided by the High Resolution Empirical Tide model (HRET, Zaron (2019)), as provided by the SWOT products.

Figure 8a shows a snapshot of the tidal prediction for a representative example, pass 28 $(-45°, 0°)$, from a region with relatively large internal tides. Note the extreme dis-

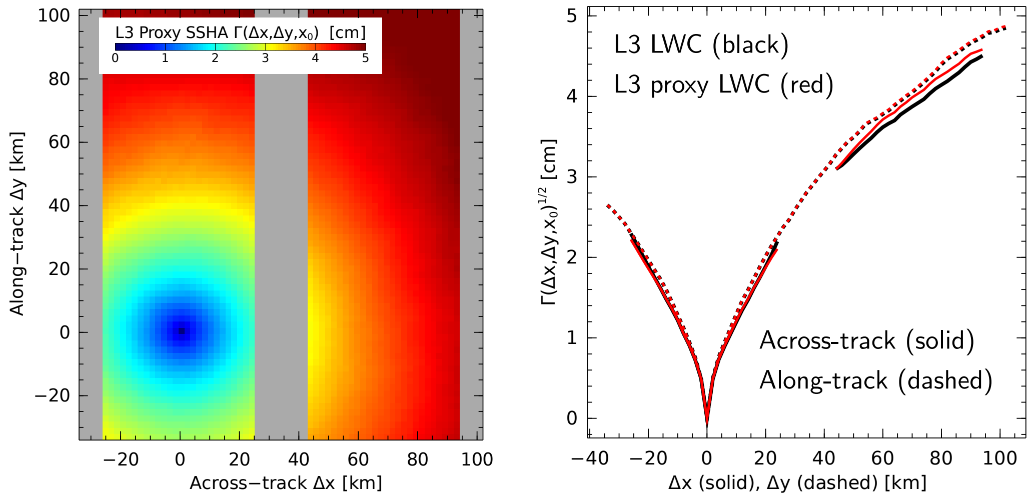

Figure 6. Left panel: The two-dimensional variogram (equation (1)) for the L2 data computed using the proxy LWC is nearly isotropic (cf., Figure 2a). Right panel: Slices through the variogram show that statistical structure of the proxy-LWC-corrected L2 data closely agrees with that of the (nonlinear) LWC-corrected L3 data. The data in this figure are taken from pass 15 in the Central Pacific ($0° - 45°$).

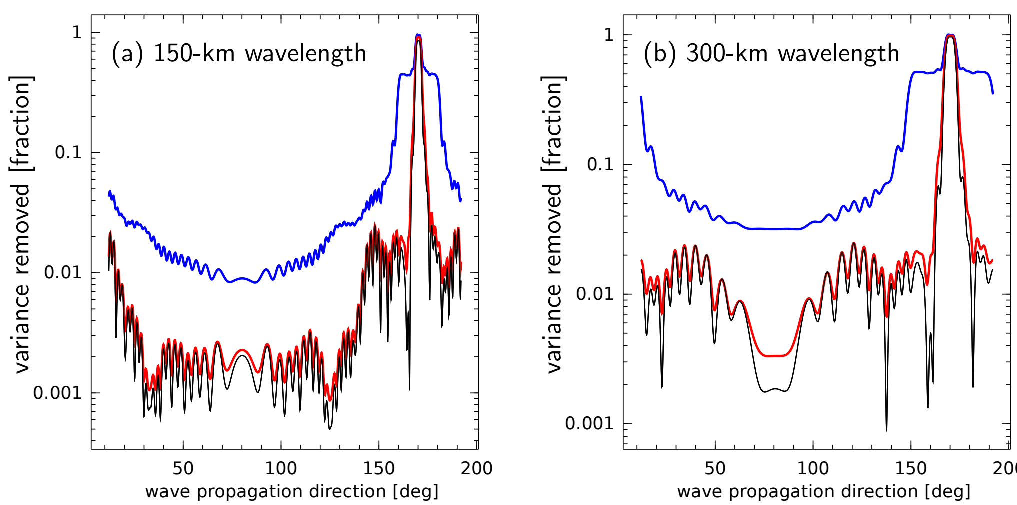

Figure 7. The amplitude response of the proxy LWC (red line) compared with an unweighted projection ($\sigma_m = 1$, blue line), and the large-scale bilinear interpolation (the first stage of the proxy LWC, black line). The response is evaluated with test waves of (a) 150-km wavelength and (b) 300-km wavelength, representative of the wavelengths of mode-1 semidiurnal and diurnal internal tides throughout the tropics and subtropics. The propagation direction of the test waves is indicated on the x-axis; note that the amplitude response is minimal for waves propagating in the same direction as the pass, 78°, and maximal for waves propagating perpendicular to the pass.

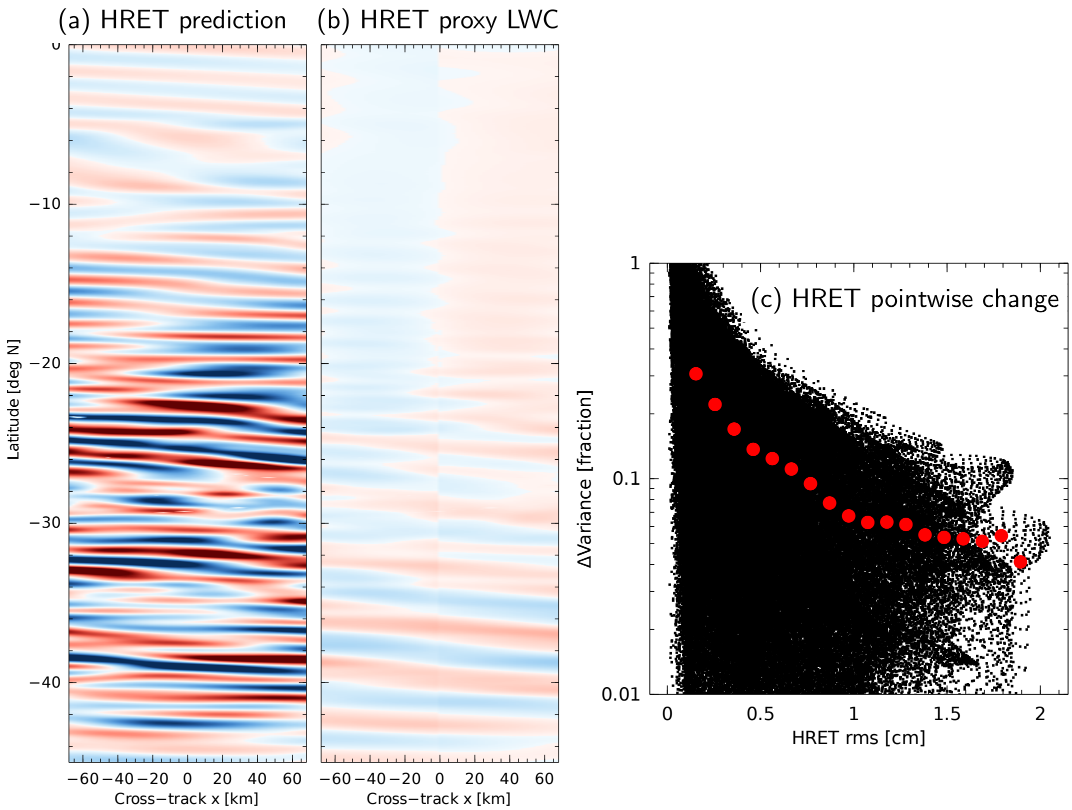

Figure 8. Impact of the proxy LWC on the HRET tide prediction (pass 28, $-45°, 0°$; a representative example). (a) A snapshot of the predicted internal tide SSHA; ±1.5 cm colorscale range. (b) The proxy LWC from the HRET field given in panel (a); same ±1.5 cm colorscale range. (c) The fractional variance of the proxy LWC applied to HRET at every pixel of the swath, averaged over all available orbit cycles. When the variance change is binned over ranges of HRET rms amplitude (red dots), the variance change exceeds 10% for pixels with amplitude less than about 0.8 cm. For larger amplitude tides, the error is approximately 5% to 10%.

tortion of the across- and along-track aspect ratio in the Figure, a factor of 50. The impact of the proxy LWC correction is visually small compared to the predicted tide (Figure 8b). In Figure 8c the variance absorbed by the proxy LWC has been computed pointwise over the pass segment and plotted as a function of the internal tide amplitude, averaging over all available orbit cycles. It is apparent that the phase errors caused by the LWC are not completely negligible and lead to changes of 10% or more for pixels with r.m.s. tidal amplitude less than about 0.8 cm. Pixels with larger tidal signals are influenced less, but errors on the scale of 5% to 10% may be anticipated for centimeter-amplitude baroclinic tidal waves in corrected L3 SWOT data.

5 Conclusions

A linear approximation of the nonlinear, data-driven, L3 LWC has been developed for SWOT data during the one-day repeat orbit phase of the mission. This correction, the so-called "proxy LWC", was implemented as a minimum-variance estimator of the L3 LWC. It is constructed from a set of spatial basis functions describing a plausible subset of possible long-wavelength errors. The proxy LWC was developed soley to characterize, approximately, the linear response of the L3 LWC and to quantify the extent to which the L3 LWC might remove baroclinic ocean tidal signals from the SSHA. The proxy LWC could be applied to other oceanographic fields or SWOT SSHA corrections; or it could be used as a LWC for the L2 data in specialized applications where the L3 product is, for whatever reason, not satisfactory.

The proxy LWC was shown to lead to corrected SSHA which closely mimics the spatial covariance properties of the L3 product. Of course, in the L3 product the large-scale SSHA is aligned with data from the existing nadir altimeter constellation, while the proxy-LWC-corrected L2 data would not be aligned with large-scale sea level. Furthermore, the L3 LWC contains a sophisticated model for the temporal structure of the LWC, a feature which is absent from the proxy LWC. Users of the proxy LWC should keep in mind these limitations if they attempt to use it with L2 data.

The proxy LWC was applied to both idealized and realistic waveforms representative of the low-mode baroclinic tidal SSHA. When measured in terms of a spatial filter response function, the proxy LWC generally absorbs a small fraction of variance from such waveforms. In particular, for most propagation directions, the proxy LWC absorbs about a factor of 10 less signal than would be absorbed by a simpler, unweighted, projection onto the LWC basis functions. For realistic tidal SSHA, the proxy LWC can lead to typical errors of 5% to 10% for waves with amplitude of a centimeter or larger. Overall, it is concluded that the L3 LWC should have a small, mostly insignificant, impact on baroclinic tidal signals,

Appendix A Interpretation of $\boldsymbol{\sigma_m}$ as weight coefficients

Regard the SSHA observations, $\boldsymbol{\eta}_0$, as the sum of a LWC, $\hat{\boldsymbol{\eta}}$, and an oceanographic signal, $\boldsymbol{\eta}_0-\hat{\boldsymbol{\eta}}$. Assuming Gaussian statistics and Bayesian priors for their variances, $\mathbf{E}_{LWC}$ and $\mathbf{R}$, respectively, the maximum likelihood estimator of $\hat{\boldsymbol{\eta}}$ is the minimizer of,

$$J(\hat{\boldsymbol{\eta}}) = \hat{\boldsymbol{\eta}}^T \mathbf{E}_{LWC}^{-1} \hat{\boldsymbol{\eta}} + (\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0)^T \mathbf{R}^{-1} (\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0). \tag{A1}$$

If it is assumed that $\hat{\boldsymbol{\eta}}$ can be represented as a linear combination of orthonormalized basis vectors, $\hat{\boldsymbol{\eta}} = \mathbf{U}^T \boldsymbol{\alpha}$, then J can be expressed in terms of $\boldsymbol{\alpha}$ as

$$J(\boldsymbol{\alpha}) = \boldsymbol{\alpha}^T \mathbf{E}^{-1} \boldsymbol{\alpha} + (\mathbf{U}^T \boldsymbol{\alpha} - \boldsymbol{\eta}_0)^T \mathbf{R}^{-1} (\mathbf{U}^T \boldsymbol{\alpha} - \boldsymbol{\eta}_0), \tag{A2}$$

where $\mathbf{E} = \mathbf{U}\mathbf{E}_{LWC}\mathbf{U}^T$.

The optimal estimate of $\boldsymbol{\alpha}$ satistfies,

$$\frac{1}{2}\frac{\partial J}{\partial \boldsymbol{\alpha}} = \mathbf{E}^{-1}\boldsymbol{\alpha} + \mathbf{U}\mathbf{R}^{-1}(\mathbf{U}^T\boldsymbol{\alpha} - \boldsymbol{\eta}_0) = \mathbf{0}, \tag{A3}$$

or

$$(\mathbf{E}^{-1} + \mathbf{U}\mathbf{R}^{-1}\mathbf{U}^T)\boldsymbol{\alpha} = \mathbf{U}\mathbf{R}^{-1}\boldsymbol{\eta}_0. \tag{A4}$$

Suppose the oceanographic variability is uncorrelated and homogeneous, $\mathbf{R} = r\mathbf{I}$,

$$(\mathbf{E}^{-1} + r^{-1}\mathbf{I})\boldsymbol{\alpha} = r^{-1}\mathbf{U}\boldsymbol{\eta}_0, \tag{A5}$$

and assume $\boldsymbol{E}$ is diagonal with entries, e_m, then the optimizing coeficients of $\boldsymbol{\alpha}$ may be expressed simply as

$$\alpha_m = \frac{r^{-1}}{e_m^{-1} + r^{-1}} u_m^T \boldsymbol{\eta}_0 = \frac{e_m}{r + e_m} u_m^T \boldsymbol{\eta}_0. \tag{A6}$$

Finally, the coefficient vector $\boldsymbol{\alpha}$ can be elimated in favor of $\boldsymbol{\Sigma}$, and the LWC can be written as,

$$\hat{\boldsymbol{\eta}} = \mathbf{U}^T \boldsymbol{\Sigma} \mathbf{U} \boldsymbol{\eta}_0, \tag{A7}$$

where the weights, σ_m, used in the text are given by,

$$\sigma_m = \frac{e_m}{r + e_m}. \tag{A8}$$

The weighting coefficients can be understood in relation to Bayesian priors for the variance of the LWC basis functions (e_m) and the oceanographic component of the SSHA variance (r). The expectation is that σ_m will tend to be smaller in regions of larger oceanic SSHA variance. Indeed, the example of shown in Figure 5 agrees; the σ_m coefficients are smaller in the Western Pacific, a region of enhanced mesoscale eddy activity due to its proximity to the western boundary current.

Open Research Section

The Level-2 SWOT data are publicly available from the NASA Jet Propulsion Laboratory PO.DAAC through the Earthdata website (`https://search.earthdata.nasa.gov`) using the PO.DAAC data-subscriber tool (`https://github.com/podaac/data-subscriber`). The Level-3 SWOT data are available from the AVISO website, with support from CNES, at `https://www.aviso.altimetry.fr/en/data/products/sea-surface-height-products/global/swot-l3-ocean-products.html`. NetCDF-format files containing the weighting coefficients, σ_m, defining the proxy LWC (2)-(7) are available from the Zenodo website `doi:10.5281/zenodo.10914546`.

Acknowledgments
The research reported in this manuscript was supported by NASA award 80NSSC21K0346 to Oregon State University.

References

Dibarboure, G., & Ubelmann, C. (2014). Investigating the performance of four empirical cross-calibration methods for the proposed SWOT mission. *Remote Sens.*, *6*(6), 4831–4869. doi: 10.3390/rs6064831

Dibarboure, G., Ubelmann, C., Flamant, B., Briol, F., Peral, E., Bracher, G., ... Picot, N. (2022). Data-driven calibration algorithm and pre-launch performance simulations for the SWOT mission. *Remote Sens.*, *14*(23), 6070. doi: 10.3390/rs14236070

Esteban-Fernandez, D., Rodriguez, E., Fu, L.-L., Alsdorf, D., & Vaze, P. (2010). The Surface Water and Ocean Topography Mission: centimetric spaceborne radar interferometry. In R. Meynart, S. P. Neeck, & H. Shimoda (Eds.), *Sensors, Systems, and Next-Generation Satellites XIV* (Vol. 7826, p. 782615). SPIE. doi: 10.1117/12.868535

Fjørtoft, R., Gaudin, J.-M., Pourthie, N., Lion, C., Mallet, A., Souyris, J.-C., ... Arnaud, A. (2010). KaRIn - the Ka-band radar interferometer on SWOT: Measurement principle, processing and data specificities. In *2010 IEEE International Geoscience and Remote Sensing Symposium* (p. 4823-4826). doi: 10.1109/IGARSS.2010.5650601

Morrow, R., Fu, L.-L., Ardhuin, F., Benkiran, M., Chapron, B., Cosme, E., ... Zaron, E. D. (2019). Global observations of fine-scale ocean surface topography with the Surface Water & Ocean Topography (SWOT) mission. *Front. Mar. Sci.*, *6*, 232.

Zaron, E. D. (2019). Baroclinic tidal sea level from exact-repeat mission altimetry. *J. Phys. Oceanogr.*, *49*(1), 193–210.