

Article

# Investigating the Performance of Four Empirical Cross-Calibration Methods for the Proposed SWOT Mission

Gerald Dibarboure <sup>1,\*</sup> and Clement Ubelmann <sup>2</sup>

<sup>1</sup> Collecte Localisation Satellite (CLS), 8-10 Rue Hermes, 31520 Ramonville St-Agne, France

<sup>2</sup> Jet Propulsion Laboratory (JPL), 4800 Oak Grove Dr, Pasadena, CA 91109, USA;

E-Mail: Clement.Ubelmann@jpl.nasa.gov

\* Author to whom correspondence should be addressed; E-Mail: gerald.dibarboure@cls.fr;

Tel.: +33-56-139-4783; Fax: +33-56-139-3782.

Received: 21 March 2014; in revised form: 12 May 2014 / Accepted: 19 May 2014 /

Published: 27 May 2014

---

**Abstract:** The proposed surface water and ocean topography (SWOT) mission aims at observing short scale ocean topography with an unprecedented resolution and accuracy. Its main proposed sensor is a radar interferometer, so a major source of topography error is the roll angle: the relative positions of SWOT's antennas must be known within a few micrometers. Because reaching SWOT's stringent requirements with onboard roll values is challenging, we carried out simulations as a contingency strategy (*i.e.*, to be ready if roll is larger than anticipated) that could be used with ground-based data. We revisit the empirical calibration algorithms with additional solving methods (e.g., based on orbit sub-cycle) and more sophisticated performance assessments with spectral decompositions. We also explore the link between the performance of four calibration methods and the attributes of their respective calibration zones: size and geometry (e.g., crossover diamonds), temporal variability (e.g., how many days between overlapping SWOT images). In general, the so-called direct method (using a single SWOT image) yields better coverage and smaller calibrated roll residuals because the full extent of the swath can be used for calibration, but this method makes an extensive use of the external nadir constellation to separate roll from oceanic variability, and it is more prone to leakages from oceanic variability on roll (*i.e.*, true topography signal is more likely to be corrupted if it is misinterpreted as roll) and inaccurate modeling of the true topography spectrum. For SWOT's baseline orbit (21 days repeat and 10.9 days sub-cycle), three other methods are found to be complementary with the direct method: swath crossovers, external nadir crossovers, and sub-cycle overlaps are shown to provide an additional calibration capability, albeit with complex latitude-varying

coverage and performance. The main asset of using three or four methods concurrently is to minimize systematic leakages from oceanic variability or measurement errors, by maximizing overlap zones and by minimizing the temporal variability with one-day to three-day image differences. To that extent, SWOT's proposed "contingency orbit" is an attractive risk reduction asset: the one-day sub-cycle overlaps of adjoining swaths would provide a good, continuous, and self-sufficient (no need for external nadirs) calibration scheme. The benefit is however essentially located at mid to high-latitudes and it is substantial only for wavelengths longer than 100 km.

**Keywords:** SWOT; empirical calibration; roll; error budget; interferometry; topography

---

## 1. Introduction and Context

### 1.1. SWOT

The proposed Surface Water and Ocean Topography (SWOT) mission from NASA (National Aeronautics and Space Administration), CNES (Centre National d'Etudes Spatiales), and CSA (Canadian Space Agency) would provide two-dimensional topography information over the oceans and inland fresh-water bodies. Dibarboure *et al.* [1] and Durand *et al.* [2] give a description of SWOT's objectives, principle, and scientific requirements.

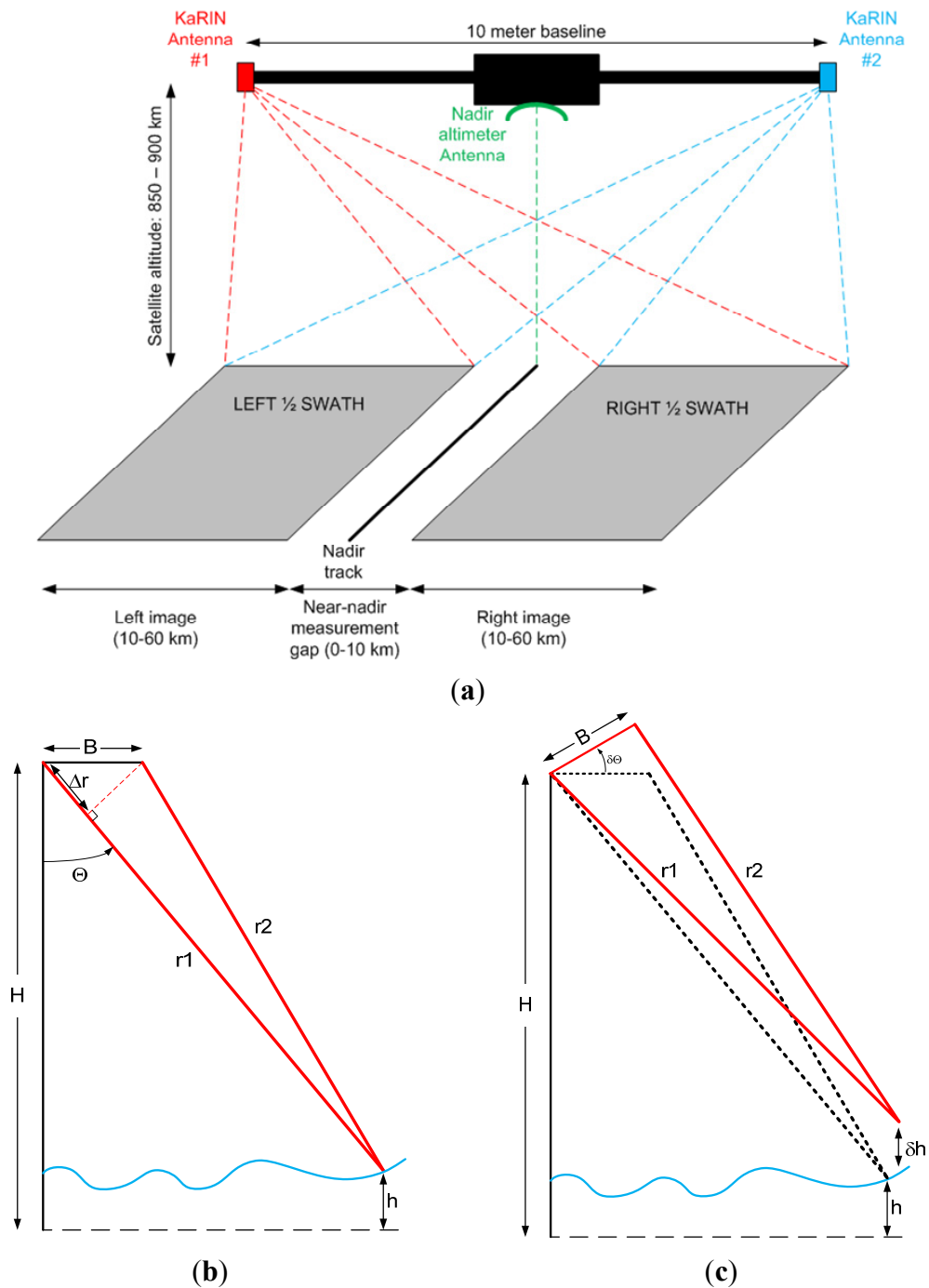
SWOT's main instrument would be the Ka-band radar interferometer (KaRIN), a synthetic aperture radar interferometer with a ground swath that is 120 km wide (Figure 1a). KaRIN would be complemented by a Jason-class nadir-looking altimeter, a microwave radiometer for the wet troposphere correction, and a precise orbit determination payload.

SWOT has two main objectives: to observe mesoscale and submesoscale processes over the oceans and to observe the water cycle over land. Fu and Rogriduez [3] and Durand *et al.* [2] highlight the stringent requirements in terms of error control. SWOT's error budget is required to be one order of magnitude below the signal. In other words, the error budget must be one decade below the signal spectrum, *i.e.*, less than 2 cm for 1 km on ocean, and 10 cm height accuracy and 1 cm/1 km slope accuracy for hydrology. SWOT's error budget is anticipated to be approximately five times smaller than the accuracy observed on Jason-class pulse-limited altimeters.

### 1.2. SWOT's Proposed Orbits

The baseline orbit envisioned for SWOT is circular with  $13 + 19/21$  revolutions per nodal day (*i.e.*, an altitude of about 890 km) and an inclination of approximately  $77.6^\circ$ . It features a 20.9-day revisit time (*i.e.*, repeat cycle), and each cycle is composed of 292 revolutions or 584 tracks. At  $0^\circ\text{N}$ , the distance between adjoining tracks is approximately 130 km, thus providing global coverage of latitudes ranging from  $25^\circ$  to  $77.6^\circ$  and near complete coverage at lower latitudes. The orbit also features a 10.9-day near-repeat cycle (*i.e.*, sub-cycle).

**Figure 1.** Conceptual overview of KaRIN, SWOT’s proposed Ka-band radar interferometer (a). KaRIN’s interferometric measurement concept (b) and consequence of a non-zero roll angle (c).



A different orbit would be used for the CalVal and commissioning phase (circular, approximately 857 km,  $77.6^\circ$ ) featuring exactly 14 revolutions per nodal day with a one-day exact repeat cycle.

Lastly, the SWOT community has defined a so-called “contingency orbit” that is discussed in this paper: the orbit is circular with  $13+20/21$  revolutions per nodal day (873 km) and an inclination of  $77.6^\circ$ . The repeat cycle is 20.9 days but instead of the 10.9-day sub-cycle of the baseline orbit,

the contingency orbit features a one-day sub-cycle only. In other words, for mid to high latitudes, overlapping images are separated by one day (as opposed to 10 days for the baseline orbit).

Dibarboure *et al.* [4] illustrate in details the nature and importance of sub-cycles for space/time sampling in a radar altimetry context. Their impact on roll calibration is discussed in Sections 3 and 4.

### 1.3. The Roll Paradigm

Rosen *et al.* [5] explain how radar interferometry uses a measurement of the relative delay between the signals measured by two antennas separated by a known distance (baseline), together with the system ranging information, to determine surface elevations in a cross-track swath. The interferometric triangle (Figure 1b) formed by the baseline  $B$  and the range distance to the two antennas  $r_1$  and  $r_2$  can be used to geolocate off-nadir points in the plane of the observation. The range difference between  $r_1$  and  $r_2$  is determined by the relative phase difference between the two signals as given by  $\Phi = kr_1 - kr_2 \approx kB \cdot \sin(\theta)$  where  $k$  is the electromagnetic wavenumber. From these measurements, the height  $h$  above a reference plane can be obtained using the equation  $h = H - r_1 \cos(\theta)$ .

If the instrument is not oriented precisely in the Nadir direction (Figure 1c), the topography measurement  $h$  is corrupted by an error  $\delta h$  that is directly proportional to the sine of the roll angle. For KaRIN's very small angles (0 to 4°), the roll error can be approximated by  $\delta h = x \times R$ , where  $R$  is the roll angle value and  $x$  the measurement position in the cross-track direction. Assuming an outer edge position of 60 km and a very small pointing error of 1 arcsec (*i.e.*, about 0.00028° or 5  $\mu$ rad) the roll topography signal is as large as 30 cm on the outer edges of the swath.

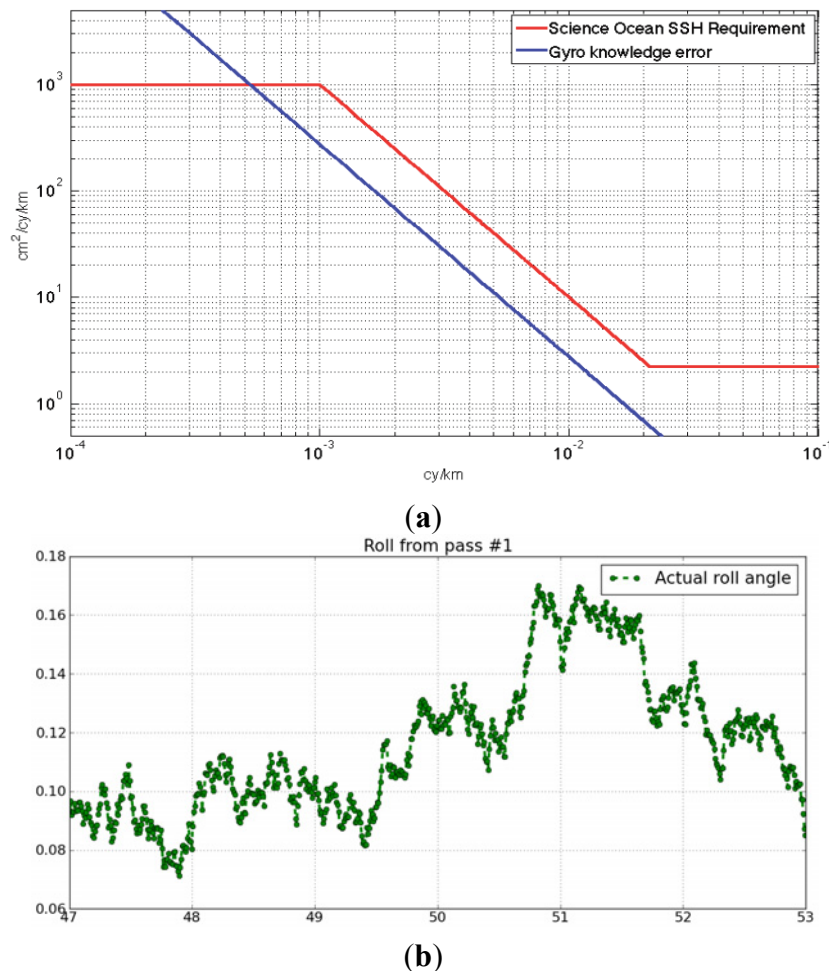
Fortunately, since the error has a predetermined linear signature in the cross-track direction, it can be perfectly corrected if the true value of the roll angle is known. Consequently the so-called “roll error” is not an error due to the mispointing of the antennas, but an error stemming from an imperfect knowledge of their true roll angle. To be within SWOT's proposed centimetric requirements on topography, it is critical to know the relative position of each KaRIN antenna within a few micrometers at all times.

SWOT's error budget is still being consolidated, and early definitions highlighted that imperfect roll knowledge might be a substantial source of error (Esteban-Fernandez [6]). To minimize this effect, the design of KaRIN might feature a stringent control of the mast distortion and angle (very rigid baseline), and high performance on-board roll determination (*e.g.*, gyroscope on the baseline).

The error made on the true roll angle of KaRIN's antennas could either originate in a non-perfect information about the attitude of the satellite (*e.g.*, residual errors from gyroscope measurements), or from an angular deformation of the instrument baseline and antennas (*e.g.*, roll angle measured at the center of the baseline, instead of the outer edges where the antennas are located).

The residual roll error (Figure 2a, blue) would represent 20% to 30% of the total error of the ocean products, with more impact on SSH wavelengths ranging from 100 to 1000 km (Esteban-Fernandez [6]). Figure 3 shows that residual roll remains a significant contributor to the topography error. Furthermore the higher-frequency components of roll would have stronger impact on topography derivatives (*e.g.*, geostrophic velocities). In contrast, hydrology products are less affected by roll as long as the long wavelength component is controlled by a relatively simple calibration method (*e.g.*, Enjolas *et al.* [7]).

**Figure 2.** (a) shows an example of roll error budget allocation (blue) with respect to the total SSH requirement (red). (b) shows a sample of along-track roll segment (in arcsec) as function of latitude. The roll sample from (b) follows the blue spectral law from (a).



#### 1.4. Empirical Cross-Calibration

Furthermore at the early stage of the mission development, one cannot entirely rule out a scenario where the roll error is higher than anticipated. To that extent, mitigation algorithms were developed as a contingency plan. Enjolras *et al.* [7] or Dibarboure *et al.* [1] developed a cross-calibration mechanism that could be used on ground to further reduce SWOT's roll error using empirical methods. Their approach is essentially to use topography measurements from multiple swaths anticipated from SWOT or from the nadir constellation to solve the problem and to provide roll estimates. Their methodology is summarized in Section 2. They demonstrated that their proof-of-concept could substantially reduce roll signals with correlations ranging from 800 km to thousands of kilometers: the residual roll signature on topography after their cross-calibration was 1.5 to 2 cm RMS (down from 70 cm before calibration).

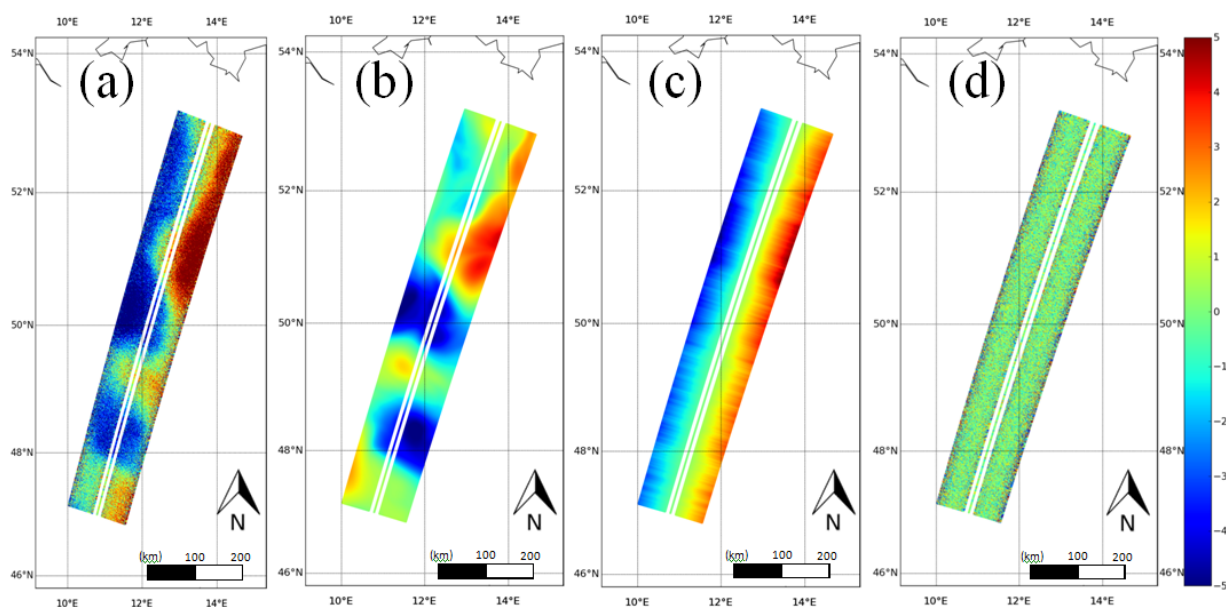
In their study, they discussed but did not answer the following questions:

- Their residual error was quantified as a total RMS, whereas the proposed SWOT mission's error budget is now addressed as the sum of the power spectral density (PSD) laws for each

error source: for which range of wavelengths can we expect ground calibration to be efficient and what is the limit?

- Would SWOT's roll calibration require inputs from the constellation of nadir altimeters?
- What is the performance of the two new methods they describe but did not implement: 1/the collinear method and 2/the neighbor or sub-cycle method?
- Lastly, they discussed the importance of the orbit, but they did not address the rationale and possible benefits of using a “roll contingency” orbit that would more suitable for roll calibration.

**Figure 3.** Decomposition of a simulated KaRIN topography image (a) as the sum of the “true” topography derived from a high-resolution ocean model from Klein *et al.* [8] (b), the roll error (c) derived from the roll sample from Figure 2b, and instrumental white noise (d). Unit: cm.



### 1.5. Objective of This Paper

This paper extends the analysis from Dibarboure *et al.* [1] and investigates the above questions. The objectives are:

- To carry out performance assessments in the wavenumber domain, and to try and understand why some methods are more efficient for long/short wavelengths,
- To implement all four calibration methods and compare their relative performance as a function of the density and geometry of different calibration zones,
- To quantify the benefits of using the so-called “contingency orbit” and the nadir constellation to improve the empirical roll correction,
- To use more recent SWOT design items: orbits, white noise level, input roll before calibration.

After a methodology introductions in Section 2, Section 3 explains the main pitfall of empirical methods: to correctly separate the signal of interest (roll) from other sources that should be left untouched in the input data (e.g., oceanic variability). Section 4 then gives some insights on the strengths and limits of each cross-calibration method.

To understand the influence of simple effects such as the geometry of calibration zones and temporal variability, we chose to use a relatively simple simulation context so the simulations presented in this paper are not realistic as we do not use the complex error decomposition from Esteban-Fernandez [6]. This paper, thus, intends to provide a methodology which would allow the definition of more sophisticated algorithms once risk mitigation and contingency scenarios are defined by the agencies and the SWOT community (discussed in Section 5.3).

## 2. Methodology

All simulations are carried out on ocean using a sea surface height (SSH) “reality” from a high-resolution model simulation described by Klein *et al.* [8]. This “reality” features very small mesoscale and sub-mesoscale features that are at the core of SWOT’s ocean requirements, but the surface covered by the model is small, and the model is not using a realistic bathymetry. To that extent, we perform only limited simulations (approximately 600 km) in the along-track direction, and we re-locate the model outputs to latitudes of interest with random space and time offsets. Note that we do not use any land/sea mask because we analyze only the impact of geometry, latitude, and temporal variability. To that extent we sometimes allow an ocean model output to be re-located on land.

The SSH structures of this model are very intense, representing typical western boundary currents. To get a reality more consistent with the global ocean, we reduce the SSH by a factor of 7 so that the large scale SSH variability is of the order of the median global oceanic variability observed by nadir altimetry (e.g., Dibarboure *et al.* [9]). However, this “ocean reality” is not representative of the geographical variations of the true oceanic variability, nor does it contain the signature on topography of various ocean processes (e.g., high-frequency waves, tides) that Klein *et al.* [8] did not consider.

Because SWOT’s errors are decomposed into cross-track and along-track components, our simulations are carried as a function of  $t$  (time, *i.e.*, position in the along-track direction) and  $x$  (position of the measurement in the cross-track direction). The orbit and attitude of the satellite are known so  $[x, t]$  and  $[\text{latitude}, \text{longitude}, \text{time}]$  are effectively interchangeable.

We interpolate this series of “reality” topography fields  $H_{\text{real}}(x, t)$ :

- along 1D profiles of nadir altimetry missions (Jason-CS and Sentinel-3 would be concurrent with SWOT) with a resolution of 7 km
- along 1D profile of SWOT’s nadir with a resolution of 7 km
- in the 2D swath of KaRIN (geometry from Figure 1a) with resolution of 5 km.

In practice, the resolution used for our SWOT simulation is 5 km because: (1) our sensitivity tests with 1 km and 2 km simulations did not exhibit a significant benefit to use the higher resolution (Esteban-Fernandez [6] describes that roll is a major error source for 100–1000 km and that the error budget for 1–5 km is dominated by random noise), (2) some processing steps are 20 to 100 times faster at this resolution (our software implementation is not numerically optimized yet).

Regarding SWOT’s nadir, this dataset is optional as per the Phase A description of the SWOT mission: we use a worst-case scenario without this dataset and we performed a few sensitivities studies to verify that our conclusions remain valid with the nadir instrument. CNES recently confirmed that SWOT would include a Jason-class altimeter and this change is discussed in Section 5.3.

Then we add an instrumental white noise  $\varepsilon(x,t)$ . In the case of nadir measurements,  $\varepsilon$  has a constant standard deviation (STD) of 1 cm for Jason-CS and Sentinel-3 (*i.e.*, we assume they are in Cryosat-2's synthetic aperture radar mode), whereas the STD for SWOT's nadir is 3 cm (*i.e.*, assumed to be in Jason-2's low resolution mode). The noise STD for KaRIN's image is 1.7 cm RMS for 1 km postings and spatially varying with the cross-track distance: the STD-law is U-shaped and ranging from 1.2 cm RMS in the center of each 1/2 swath to 3 cm on the outer edges [6]. Because our synthetic observations have a  $5 \text{ km} \times 5 \text{ km}$  resolution, the effective noise STD is reduced by a factor of 5 (2D averaging of random process).

For KaRIN's roll, we use two input scenarios: baseline and worst-case. The baseline scenario is the spectral allocation for roll in SWOT's error budget (blue PSD from Figure 2a), and our worst-case scenario pessimistically assumes that KaRIN's roll PSD is ten times larger than the baseline scenario for all frequencies. From the roll PSD, we generate along-track samples of the roll angle  $R(t)$  using inverse FFT with a random phase (e.g., Figure 2b). To measure an average performance, we generate 25 to 100 samples with random realizations of noise  $\varepsilon$ ,  $H_{\text{real}}$ , and roll  $R$ .

The main equations used to estimate roll are summarized in Sections 2.1 and 2.2. Further details about implementation are given by Dibarboure *et al.* [1].

### 2.1. Observations

Our topography measurements  $H_{\text{obs}}(x,t)$  can be decomposed as follows:

$$H_{\text{obs}}(x,t) = H_{\text{real}}(x,t) + x \cdot R(t) + \varepsilon(x,t) \quad (1)$$

The most basic roll estimation method (“direct method” hereafter) consists in a linear model adjustment on  $H_{\text{obs}}$  in the cross-track direction  $x$  to obtain  $R(t)$ . This approach is valid only when  $H_{\text{real}}$ ,  $R$  and  $\varepsilon$  are spectrally separable (either along-track or cross-track or in time).

In Section 3, we improve the spectral separation of the direct method by using the nadir constellation to construct a proxy  $H_{\text{ref}}$  of  $H_{\text{real}}$ . The proxy is similar to AVISO maps described by Dibarboure *et al.* [9]. While this proxy is limited by the space/time resolution of nadir altimetry (e.g., discussed by Chelton and Schlax [10]), it still removes a fraction of the oceanic variability in SWOT images: instead of  $H_{\text{obs}}(x,t)$  from Equation (1), the measured observation is now  $Y(x,t)$  from Equation (2), where roll  $R$  must be separated from  $\delta H(x,t)$ , *i.e.*, the rapid and small scale components of  $H_{\text{real}}$  that are not resolved by nadir altimetry  $H_{\text{ref}}$ .

$$Y(x,t) = H_{\text{obs}}(x,t) - H_{\text{ref}}(x,t) = x \cdot R(t) + \delta H(x,t) + \varepsilon(x,t) \quad (2)$$

Similarly, when a KaRIN pixel  $H_{\text{obs}}(x,t)$  and a nadir altimeter  $H_{\text{nadir}}(x,t')$  measurements are overlapping on position  $x$  (e.g., crossover segment), we can use  $Y$  the difference between  $H_{\text{obs}}$  and  $H_{\text{nadir}}$ . The benefit of this method (“Nadir  $\times$  KaRIN crossover method” hereafter) is that we replace  $H_{\text{real}}$  from Equation (1) (total signal), by  $\delta H(x,t' - t)$ , the rapid components of  $H_{\text{real}}$  that are not constant between  $t$  and  $t'$ . We also replace  $\varepsilon(x,t)$  by  $\delta \varepsilon_{\text{hf}}$ , the rapid components that are not constant between  $t$  and  $t'$  (e.g., tide residuals or internal waves), or different in KaRIN and nadir altimetry (e.g., sea state bias residuals).

$$Y(x,t,t') = H_{\text{obs}}(x,t) - H_{\text{nadir}}(x,t') = x \cdot R(t) + \delta H_{\text{hf}}(x,t' - t) + \delta \varepsilon_{\text{hf}}(x,t' - t) \quad (3)$$



Lastly when two KaRIN pixels are overlapping (e.g., cycle  $N$  minus cycle  $N - 1$ ), we can use  $Y'$  the difference between  $H_{\text{obs}}(x, t)$  and  $H_{\text{obs}}(x', t')$ , to replace  $H_{\text{real}}$  from Equation (1) (total signal), by  $\delta H(x, t - t')$ , *i.e.*, the rapid components of  $H_{\text{real}}$  that are not constant between  $t$  and  $t'$ . Furthermore, in place of a direct measurement of the roll in  $H_{\text{obs}}$ , KaRIN overlaps give an observation of the difference between the signature of  $R(t)$  on the first swath, and  $R(t')$  on the second swath. Like in Nadir  $\times$  KaRIN crossovers, the cross-calibration input  $Y'$  is altered only by rapid topography changes  $\delta H_{\text{hf}}$  or rapid error changes  $\delta \varepsilon_{\text{hf}}$ . This is the equation used for the so-called Sub-cycle, Collinear, and KaRIN  $\times$  KaRIN crossover methods.

$$Y'(x, t, t') = H_{\text{obs}}(x, t) - H_{\text{obs}}(x', t') = x \cdot R(t) - x' \cdot R(t') + \delta H_{\text{hf}}(x, t' - t) + \delta \varepsilon_{\text{hf}}(x, t' - t) \quad (4)$$

## 2.2. Problem Solving

To get an estimate  $R_{\text{est}}$  of the true roll  $R$  from our observation ( $H_{\text{obs}}$ ,  $Y$  or  $Y'$ ), we either use least squares (Equation (5)) or an inverse method described by Bretherton *et al.* [11] (Equation (6)). In these equations  $M$  is the observation model mapping the state space to the observed space (e.g., linear signature models).

$$R_{\text{est}} = (M^T \cdot M)^{-1} \cdot M^T \cdot Y \quad (5)$$

$$R_{\text{est}} = C_{\text{xx}} \cdot M^T \cdot (M \cdot C_{\text{xx}} \cdot M^T + C_{\text{vv}})^{-1} \cdot Y \quad (6)$$

The matrix  $C_{\text{xx}}$  is the covariance of the unknown variable ( $R(t)$ ) which is linked to the roll spectrum from Figure 2a (blue). Matrix  $C_{\text{vv}}$  is the observation error covariance: any a priori knowledge on the covariance or spectrum of  $H_{\text{real}}$  and  $\varepsilon$  (Equation (1)), or  $\delta H$  and  $\delta \varepsilon$  (Equations (2)–(4)) is modeled in this matrix.

## 3. Spectral Separation of Roll from Other Cross-Track Gradients

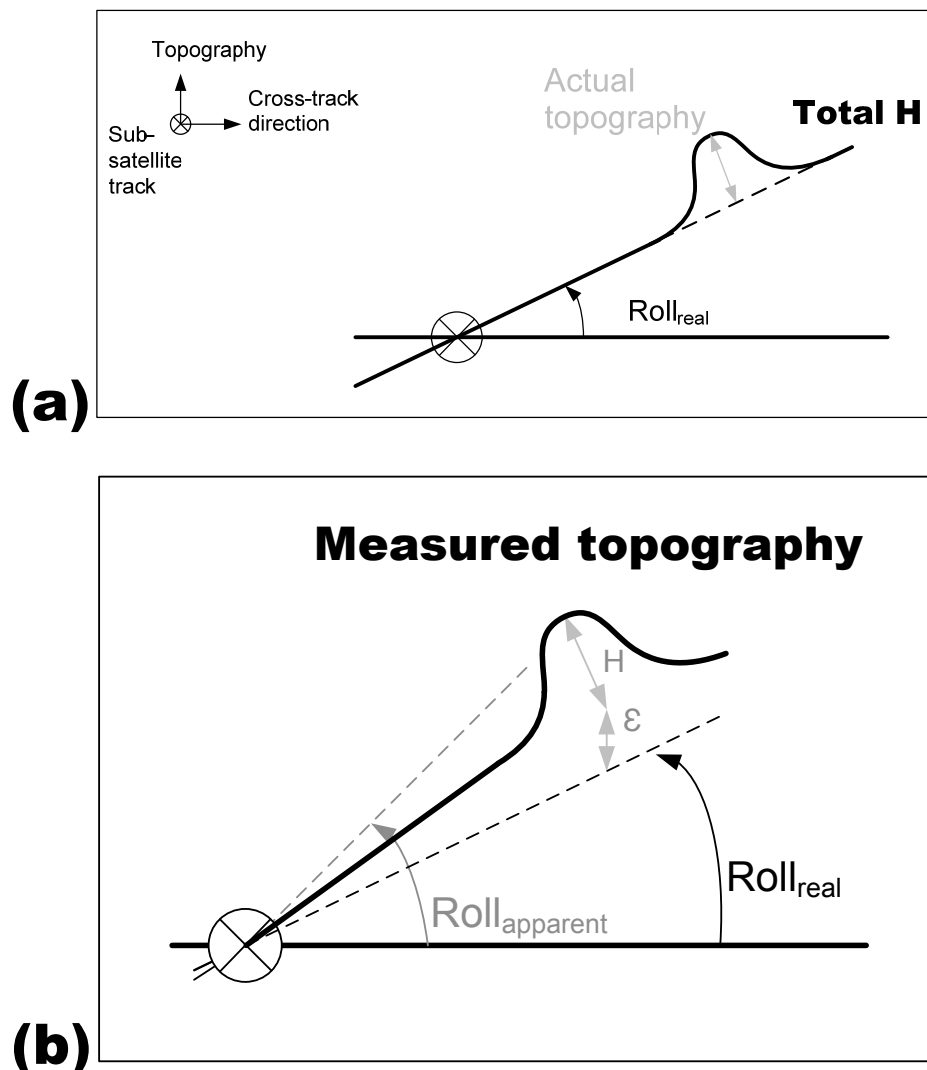
### 3.1. Pitfalls of Empirical Roll Calibration

Roll cross-calibration of SWOT's level 2 data is addressed here with an empirical approach: we do not have access to external measurements or mechanical models since we assume that both types of corrections have been applied in upstream processing steps. We are trying to correct for *residual* roll, *i.e.*, imperfections of other roll reduction methods.

The premise when using such an empirical method is that the signal of interest (here roll) is spectrally separable from other sources (e.g., noise, oceanic variability) in the observation data (topography). However, this premise is not met if we use Equation (1).

To illustrate this point, Figure 3 shows the decomposition from Equation (1): Figure 3a is  $H_{\text{obs}}$ , *i.e.*, the sum of the true topography  $H_{\text{real}}$  (Figure 3b), the topography signature of roll  $R$  (Figure 3c) and white noise  $\varepsilon$  (Figure 3d). In this example, it is clear that there are cross-track gradients from true topography or from white noise that could be misinterpreted as roll like in the schematics Figure 4.

**Figure 4.** Empirical estimation of the roll angle from topography. (a) shows an ideal case where the actual topography  $H$  can be perfectly separated from the linear signature of roll, and (b) shows a more realistic situation where the measured topography contains cross-track gradients from  $H$  or measurement errors  $\varepsilon$ . If these gradients are misinterpreted as roll, then roll estimates would yield  $\text{Roll}_{\text{apparent}}$  instead of  $\text{Roll}_{\text{real}}$ .

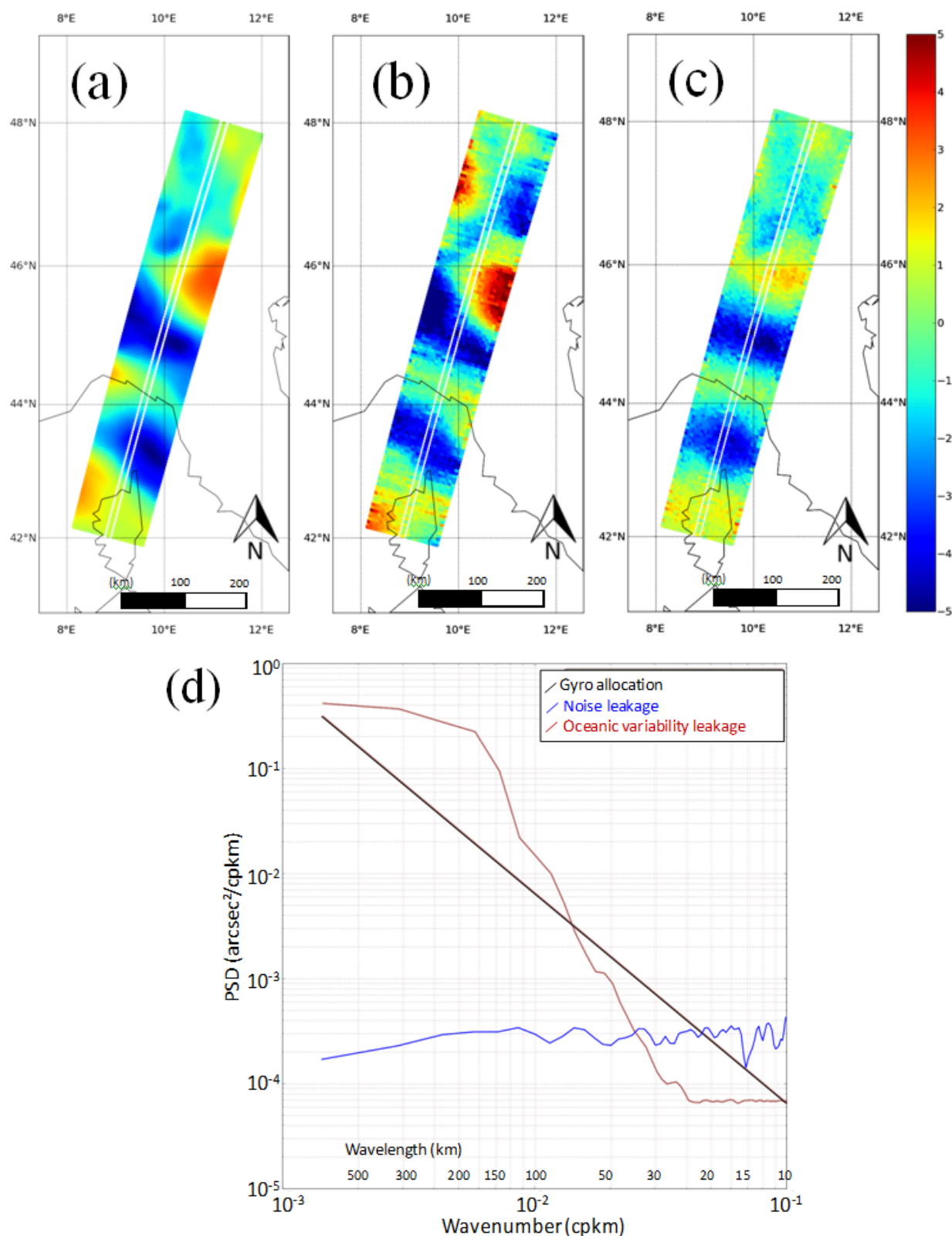


Consequently, if we try to inverse Equation (1) directly with Equation (5), then noise and oceanic variability would leak on our estimated roll. And if we correct the topography with the linear gradient from this estimated roll, the correction would be erroneous.

Figure 5 illustrates this phenomenon with an extreme example where the inversion is done with simple unweighted least squares. Figure 5a shows the true error-free topography, and Figure 5b is when noise and roll are added. Figure 5c shows the residual after Equation (1) is solved for roll with Equation (5). While the input roll from Figure 5b has been effectively removed, the inversion has also destroyed all cross-track gradients from the true topography, thus doing more harm than good to this SWOT simulated sample.

To better quantify the influence of each source of error, Figure 5d shows the output from two sets of simulations where our input roll was exactly zero (*i.e.*, estimated roll is due to leakage only):

**Figure 5.** Spectral leakage generated by a linear fit with unweighted least squares on a simulated topography image. (a) is the true topography, (b) the measured topography (with roll and white noise), and (c) the topography after roll calibration (unit: cm). (d) shows the roll leakage PSD where KaRIN's white noise (blue) or oceanic variability (red) are misinterpreted as roll noise because they have a non-zero cross-track component.



- In the first experiment (blue spectrum), we set the oceanic variability to zero in order to measure how instrumental noise (e.g., Figure 3d) would leak on our roll estimate if it was solved with Equation (5). As expected, the leakage artifact looks like Gaussian noise on our roll estimate, and it is primarily a problem for higher frequencies. This effect is amplified because the standard deviation of the instrumental noise is higher in the outer edges of the swaths, thus also increasing the probability to get a non-zero cross-track component that cannot be separated from actual roll.
- In the second experiment (red spectrum), we set the instrumental noise to zero to measure how oceanic variability would leak on our roll estimate with Equation (5). The red spectrum indicates that the leakage primarily happens for wavelengths larger than 100 km. This is caused by the width of SWOT's swath (120 km): large mesoscale structures (150–300 km) have a linear cross-track component hardly separable from actual roll because the instrumental field of view is not large enough.
- In the following sections, we try and understand how such leakage can be mitigated by using a nadir constellation proxy or image-to-image differences (problem described with Equations (2) to (4) instead of Equation (1)), and by using an inverse method that exploits our statistical knowledge of roll and oceanic signal spectra (problem solved with Equation (6) instead of Equation (5)).

### 3.2. Benefit and Limit of Using Image Differences

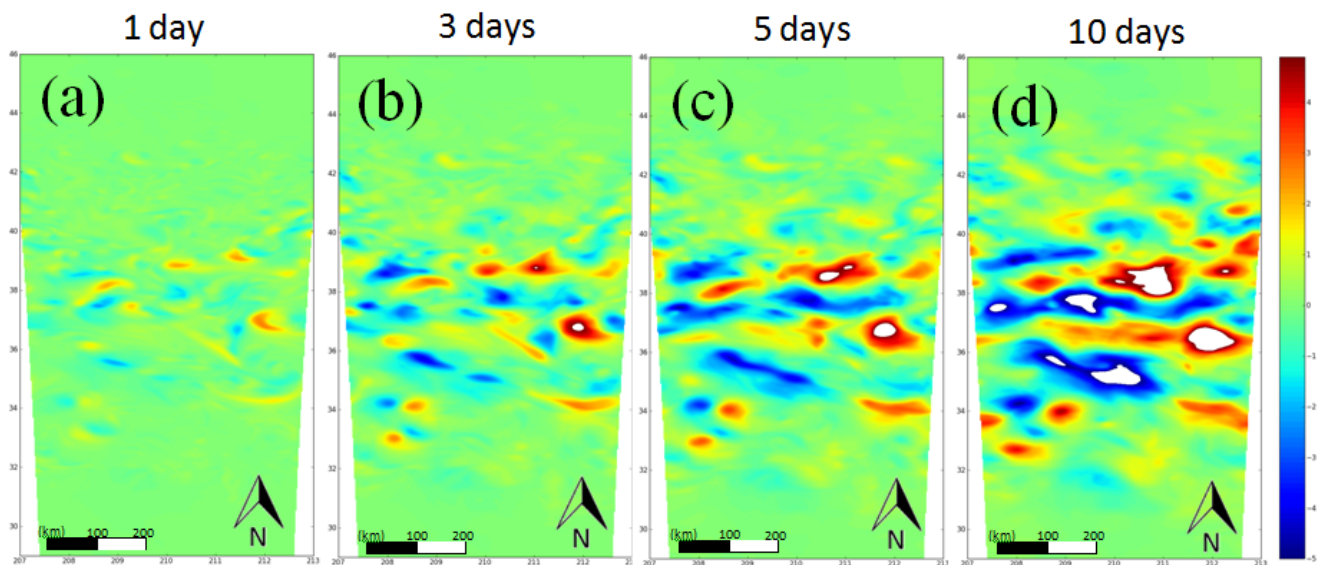
In the red spectrum from Figure 5d, the full extent of the ocean variability of our “reality” is in the observation  $H_{\text{real}}$ . Minimizing the leakage from  $H_{\text{real}}$  is the rationale behind the crossover method presented by Dibarboure *et al.* [1]: a difference between two SWOT images (Equation (4)) contains only the high-frequency fraction of the oceanic variability.

Figure 6 shows that if roll is estimated on image differences (e.g., separated by one to three days), spectral leakage would be smaller for two reasons: (1) because the amplitude of each structure is 3 to 10 times smaller, and (2) because rapid structures/changes are often smaller in radius as well (e.g., smaller structures in Figure 6b for three days than Figure 6d for 10 days). The second reason is probably the most important because if mesoscale features are smaller than KaRIN's field of view (120 km) it becomes easier to separate them from linear gradients of roll in an image.

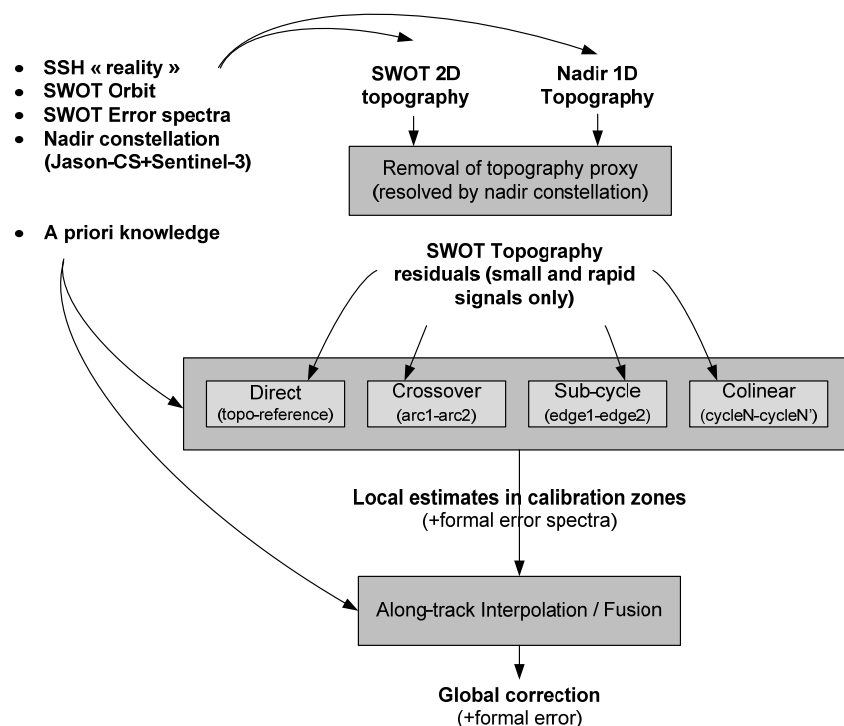
Figure 7 illustrates the three methods described by Dibarboure *et al.* [1] to get co-located differences between  $H_{\text{obs}}(t)$  and  $H_{\text{obs}}(t')$  like in Equation (4):

- Crossovers: when the swath from an ascending (respectively descending) arc crosses the swath of a descending (resp. ascending) arc.
- Sub-cycle: when the proposed SWOT mission would be on its final orbit with a 21-day repeat cycle, adjoining swaths would start to overlap at low to mid latitudes.
- Collinear: SWOT's revisit time and reference track would be controlled, so SWOT images would be almost perfectly co-located every repeat cycle (e.g., one-day cycle with the commissioning/CalVal orbit).

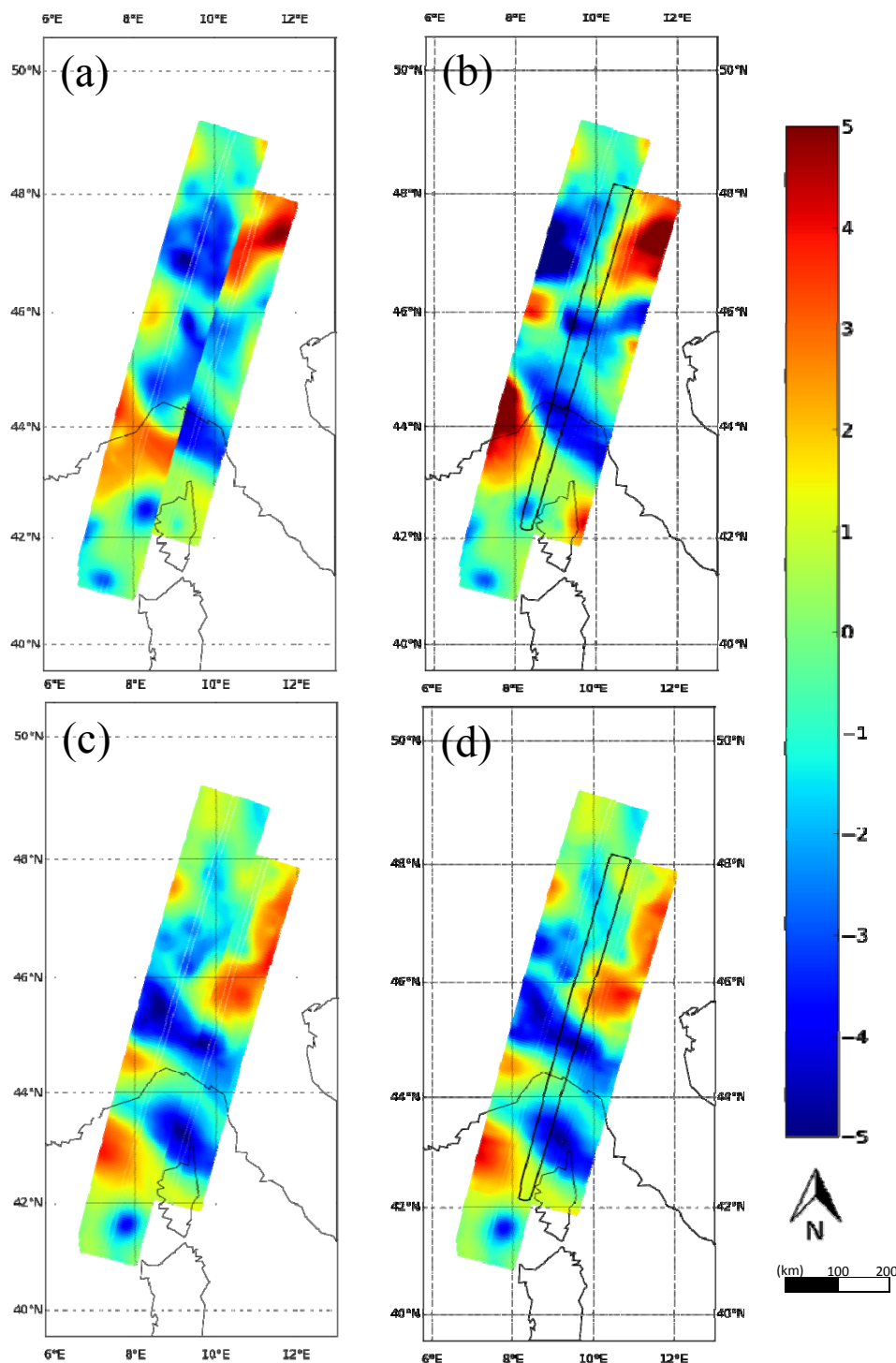
**Figure 6.** Example of ocean surface topography variability derived from the model from Klein *et al.* [8]: amplitude of the difference between two arbitrary model snapshots separated by 1 day to 10 days ((a) 1 day; (b) 3 days; (c) 5 days; (d) 10 days). Unit: cm.



**Figure 7.** Overview of SWOT's empirical cross-calibration scheme. Step 1 uses 1D topography profiles from the nadir altimetry constellation (e.g., Jason-CS + Sentinel-3) to generate a large/slow proxy of the SSH field (*i.e.*, scales resolved by nadir altimetry) that is removed from SWOT's 2D image. In step 2, the residual SSH images are processed by multiple roll estimation methods on independent crosscalibration zones. In step 3, all local roll estimates are merged in an interpolation / fusion process to handle gaps and overlaps.



**Figure 8.** Spectral leakage of oceanic variability on roll with the sub-cycle method. (a) shows the “true” and roll-free topography for two overlapping images separated by a 10-day sub-cycle. If the oceanic variability is not accounted for in the roll inversion, both images are skewed by unweighted least squares to minimize the differences in the overlap zones, thus creating an artificially seamless composite (b). The same phenomenon is observed, albeit with a much smaller amplitude, if the proposed SWOT mission would use an orbit with a one-day sub-cycle (the small discontinuities from the two swaths of (c) disappear entirely in (d)). Unit: cm.



Using differential observations helps mitigate the leakage of oceanic variability, but residual leakages remain. Figure 8 is an extreme example of the leakage obtained when Equation (4) is solved with unweighted least squares with Equation (5). Figure 8a shows the roll and noise-free oceanic variability for two adjoining swaths of the baseline orbit. As expected from Figure 6, many features have changed in 10.9 days, resulting in major differences between both images. Figure 8b shows that solving the problem with Equation (5) still creates a substantial leakage of mesoscale on roll: both images are skewed into an artificially seamless composite because differences in  $Y'(x,t)$  are misinterpreted as roll.

The artifacts from Figure 8b are very large because we assumed the baseline orbit with a 10.9-day sub-cycle. Conversely, if we use SWOT's proposed "contingency" orbit, the one-day sub-cycle provides a much more roll-friendly input dataset as shown by Figure 8c. The one-day discontinuity is exceedingly smaller as per Figure 6 and the leakage on roll is also smaller. However, Figure 8d shows that the images that were not *perfectly* seamless in Figure 8c are now skewed into a *perfect* seamless composite. The leakage of oceanic variability is less pronounced in Figure 8d than in Figure 8b because we use one-day differences instead of 10-day differences but the basic phenomenon is the same.

### 3.3. Further Mitigating Leakage on Roll

#### 3.3.1. Nadir Constellation Proxy

In the red spectrum from Figure 5d, the full extent of the ocean variability of our "reality" is in the observation  $H_{\text{real}}$ . Yet a fraction of the mesoscale features is observed in gridded topography maps from the nadir constellation (e.g., AVISO maps described by Le Traon *et al.* [12] or Dibarboure *et al.* [9]).

If we remove this proxy  $H_{\text{ref}}$  of large mesoscale, we can replace Equation (1) by Equation (2). The proxy can (and should) be used also in Equation (3) or Equation (4) in order to mitigate the leakage of mesoscale. From the analysis from Chelton *et al.* [13], we can anticipate that the gain would be for large and slow mesoscale features because smaller and more rapid structures are not resolved by nadir altimetry and smoothed in the gridding of 1D nadir profiles, even if the mapping process is tuned to resolve slightly smaller scales (e.g., Escudier *et al.* [14]).

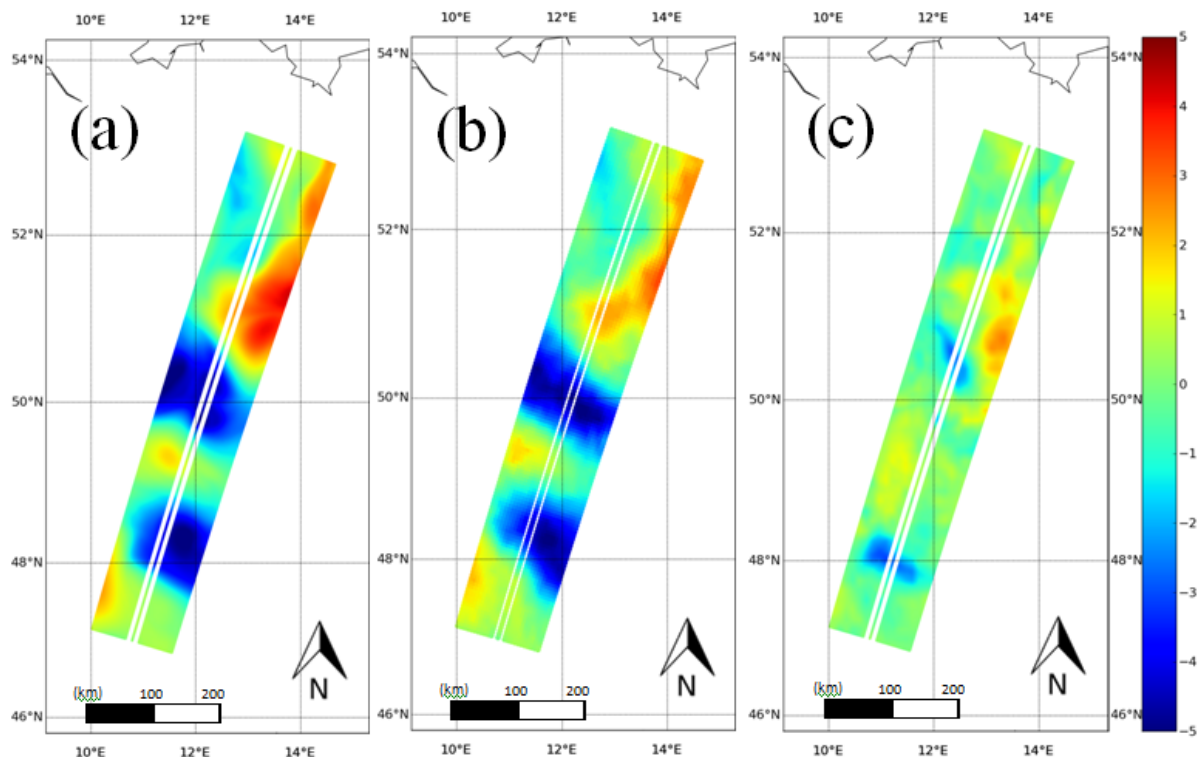
Figure 9 illustrates the proxy approach. Figure 9a is the true topography from previous figures, and Figure 9b is the proxy reconstructed with the AVISO processor described by Dibarboure *et al.* [9] in its short-scale mode: we reconstruct smaller mesoscale features (observed sporadically along-track) even though these structures cannot be resolved continuously in time (not enough altimeter tracks). This strategy is interesting in a roll calibration context to remove as much mesoscale as possible, but this is arguably not appropriate if one wants to create AVISO-like maps where mesoscale structures are consistently observed throughout their life cycle (or simply in subsequent maps).

The proxy is constructed with Jason-CS and Sentinel-3 1D profiles. Figure 9c is the residual not resolved by nadir altimetry, *i.e.*,  $\delta H(x,t)$  from Equation (2). Because KaRIN's projected swath is 120 km, the proxy allows a better separation of the roll from small mesoscale features in Figure 9c: even when the cross-track gradient is not zero, residual mesoscale anomalies are essentially round and



fluid shapes whereas roll is perfectly linear. We substantially improved the 2D spectral separation capability between roll and mesoscale.

**Figure 9.** Usage of nadir altimetry and inverse method to mitigate absorption of oceanic variability in roll. (a) shows the “true” topography derived from the high-resolution ocean model. (b) shows the low resolution proxy estimated from nadir altimetry (Jason-CS + Sentinel-3) using algorithms described by Dibarboure *et al.* [9]. (c) shows the residuals after the proxy from (b) has been subtracted from the topography of (a). Unit: cm.



### 3.3.2. Inverse Method

Furthermore, to better solve the problem and to minimize the leakage on roll, one can replace simple least squares with an inverse method where the spectral information is used (Equation (6)). The mean spectrum of the roll signal can be used in matrix  $C_{xx}$ , and statistical knowledge on oceanic variability or measurement errors can be used in  $C_{vv}$ . One of the simplest forms of  $C_{vv}$  is weighted least squares: because we know that instrumental noise is higher on the outer edges of the swath (e.g., Figure 3d) we can describe the larger variance for these measurements points. For roll or oceanic variability we also account for their correlation (*i.e.*, spectrum) in non-diagonal terms of  $C_{xx}$  and  $C_{vv}$ .

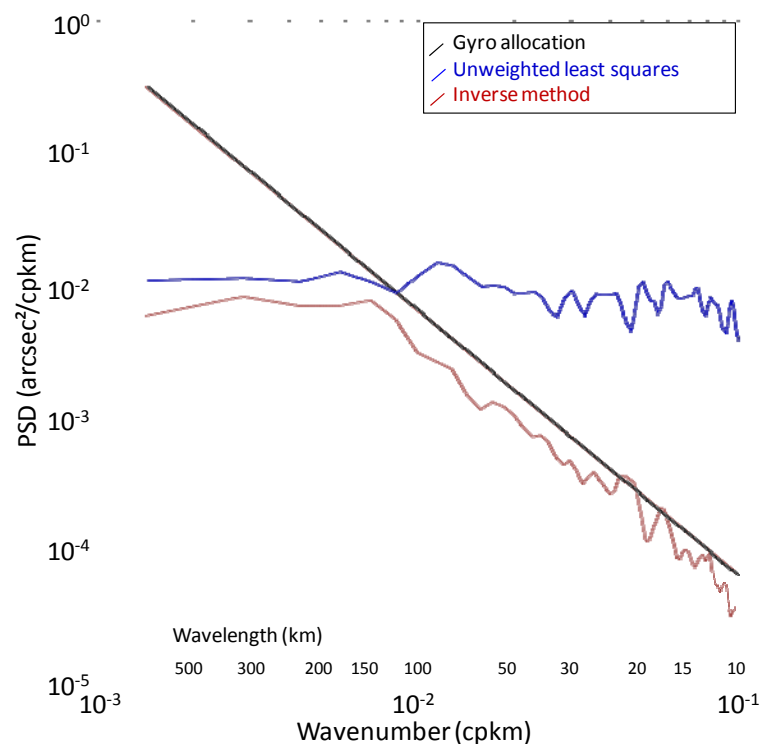
Figure 10 shows the influence of this change on noise leakage in the so-called sub-cycle method (applied on one-day image differences, see Section 4.4): here we set the input roll and oceanic variability to zero so that any non-zero value of the estimated roll is the result from noise leakage. With least squares (blue spectrum), the leakage creates apparent noise on the estimated roll (flat spectrum). The noise magnitude is substantially larger than for the direct method (Figure 5, blue) because the sub-cycle overlap is only 15 km for latitudes of 30° (discussed in Section 4.4). If we use



the inverse method to solve the same problem (Figure 10, red), we force the estimated roll (red spectrum) to be bound by a specific power law (black spectrum). While the higher wavenumbers of roll cannot be estimated because roll is below the noise floor of the outer edges of KaRIN images, the inverse method at least ensures that the cross-calibration method does not actually add undesirable energy by leaking this noise into the estimated roll like in the blue spectrum.

The same phenomenon happens for oceanic variability: if we model the 2D correlation of oceanic variability (coherent in space and time) in  $C_{vv}$ , we mitigate their absorption in the estimated roll because  $C_{vv}$  allows the inversion to distinguish true roll gradients from fluid 2D structures. Figure 11 illustrates the benefit of using both the nadir altimetry proxy and the inverse method to reduce the leakage (true roll is zero so any residual is leakage). The gain is a factor of 5 to 10 for most wavelengths.

**Figure 10.** Spectral leakage of simulated KaRIN’s noise on roll with the sub-cycle calibration method (latitude  $30^\circ$ ) when the problem is solved with unweighted least squares (blue) or with an inverse method (red).



### 3.4. Summary: Observability Rules

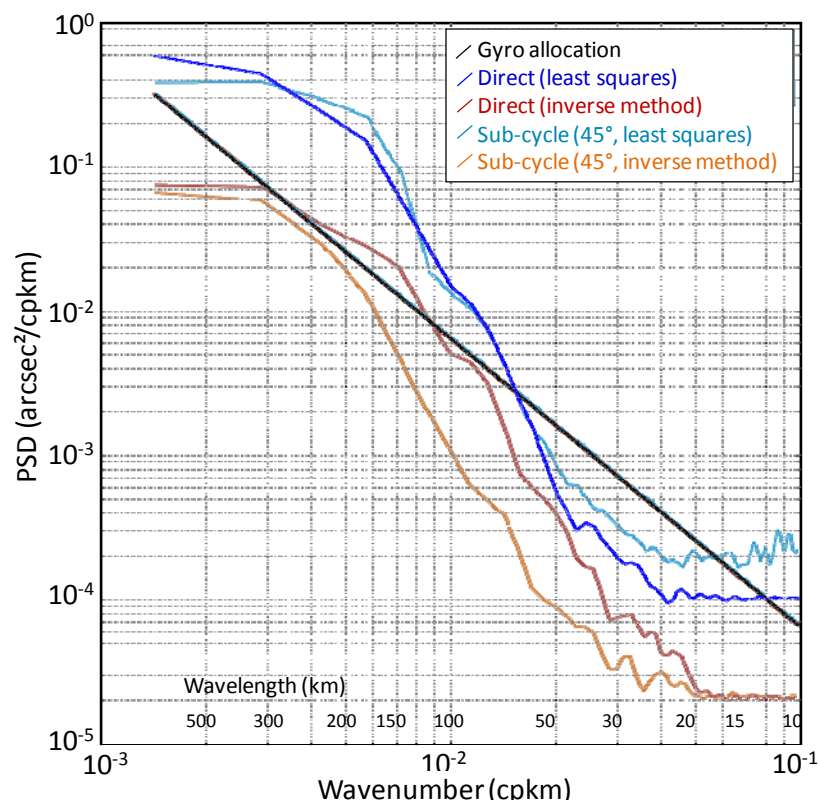
From these simple simulations, it is clear that the difficulty is to separate the linear gradients of roll from spatially coherent mesoscale features or measurement errors. The observability challenge of roll calibration is linked with spectral separation and this can be articulated in a small set of rules:

- **O1:** The larger the calibration scene, the easier it is to achieve the spectral separation. Ideally, one would need a scene with a radius  $> 500$  km (*i.e.*, more than three times the radius from Chelton *et al.* [10] or Le Traon *et al.* [12]), but in practice the limit would be KaRIN’s field of view or less (e.g., sub-cycle and crossover overlaps could be smaller than 60 km).

- **O2:** One can enhance spectral separation with an image-to-image difference. The shorter the temporal distance between co-located pixels, the better the spectral separation is. Because a larger fraction of the variability is removed in the difference (smaller amplitude), and because rapid features are also smaller in size (the relative size of the calibration scene increases).
- **O3:** Spectral separation is enhanced by removing external content derived from the nadir constellation. The benefit to use nadir-derived maps is inversely proportional to the scene/method's compliance to O1 and O2. In other words, if the scene is small or if a substantial fraction of mesoscale cannot be cancelled with an image-to-image difference, then external content from nadir maps would become a necessary input for roll calibration.

The relative performance of each method, largely driven by these rules, is analyzed in the next sections.

**Figure 11.** Usage of nadir altimetry and inverse methods to mitigate leakage of oceanic variability in roll. The dark blue PSD shows how much ocean variability is absorbed as apparent roll when unweighted least squares and the total topography are used in the direct method, and the red PSD shows the same results with an inverse method and a nadir constellation proxy. The cyan PSD shows how much ocean variability is absorbed as apparent roll when unweighted least squares and the total topography are used in the sub-cycle method (latitude of  $45^\circ$ ), and the orange PSD shows the same result with an inverse method and a nadir constellation proxy.



#### 4. Understanding the Performance of the 4 Cross-Calibration Methods

Figure 7 gives an overview of the advanced roll simulation/correction scheme developed from the findings of Section 3. The first processing step is to remove from  $H_{\text{real}}$  the nadir constellation proxy  $H_{\text{ref}}$  using AVISO-like maps computed from external nadir altimeters (Jason-CS and Sentinel-3).

The second step is to estimate roll in local calibration zones; this is done independently for each calibration zone and each method. This strategy results in a large amount of independent roll estimates in segments ranging from tens to thousands of kilometers. The final interpolation/fusion processing step is then used to account for gaps and overlaps between these independent roll calibrations.

Each local calibration provides a reliable error estimate as shown by Dibarboure *et al.* [1] and this formal error is used in the final fusion. In other words: if there are overlapping local calibration zones, the final merging would put a higher weight on the more trustworthy local roll estimates.

The main question raised by this global calibration scheme is the following: is one method substantially better than the others or should one use all four methods concurrently?

To investigate this question we use the “worst-case” scenario where we assume that the input roll before cross-calibration is ten times above the mission requirements for all wavelengths. This scenario is by no means realistic but it is a good way to gauge the calibration capability as a contingency measure if the roll observed during the SWOT’s commissioning phase is higher than the requirements for specific wavelength bands.

Using this input scenario, we repeatedly applied the four calibration methods of Figure 7 in various conditions to compare the spectrum of the calibrated roll with the original one (and the requirements).

##### 4.1. Direct Method

The direct method uses Equations (2) and (6): roll is directly estimated in a single segment of SWOT image (e.g., hundreds to thousands of kilometers along-track) after the nadir SSH proxy from nadir altimetry is removed. By far the main advantage of this method is that it can be used with all orbits, and at all latitudes.

The average residual of direct roll calibration is shown in Figure 12a. The green and orange spectra show the input (uncalibrated) roll for two simulations (with and without the presence of a mesoscale proxy from the nadir constellation): both inputs follow our worst-case spectral power law and they are ten times above the requirements. The blue spectrum is the average PSD of the calibrated roll after we apply the direct method. Note that the blue spectrum includes: 1/residual roll that has not been estimated properly and 2/leakage from noise or oceanic variability.

The roll reduction observed on Figure 12a is substantial with a factor of 5 to 20. The reduction is higher for wavelengths ranging from 20 to 60 km and for wavelengths larger than 200 km. This is explained by the observations from Section 3:

- The direct inversion is done using the full extent of the 120-km wide swath. This large cross-track field of view makes it easier to separate true roll from small mesoscale signatures (rule O1). This is emphasized between 20 and 60 km because the roundish shapes of mesoscale and cross-track roll gradients are visually and numerically easy to separate. For longer

wavelengths, mesoscale structures are observed only partially in the geometry from Figure 1a (two 1/2 swaths of 50 km each) and roll and mesoscale are more difficult to separate.

- For wavelengths longer than 150–200 km, the nadir constellation proxy accounts for the bulk of mesoscale gradients, which could be misinterpreted as roll (rule O3). This ability of the nadir constellation proxy to mitigate leakage is visible in the difference between the red (without proxy) and blue (with proxy) spectra from Figure 12a.
- Incidentally, because our covariance error model  $C_{vv}$  is focused on the smaller scales of mesoscale residuals after the proxy removal (Figure 9c), we also observe an improvement for high-frequency roll as well. When the proxy is not used, the model in  $C_{vv}$  is dominated by large mesoscale, which slightly increases the leakage of smaller structures and noise.

**Figure 12.** Average performance of the direct method (a) and the collinear method (b).

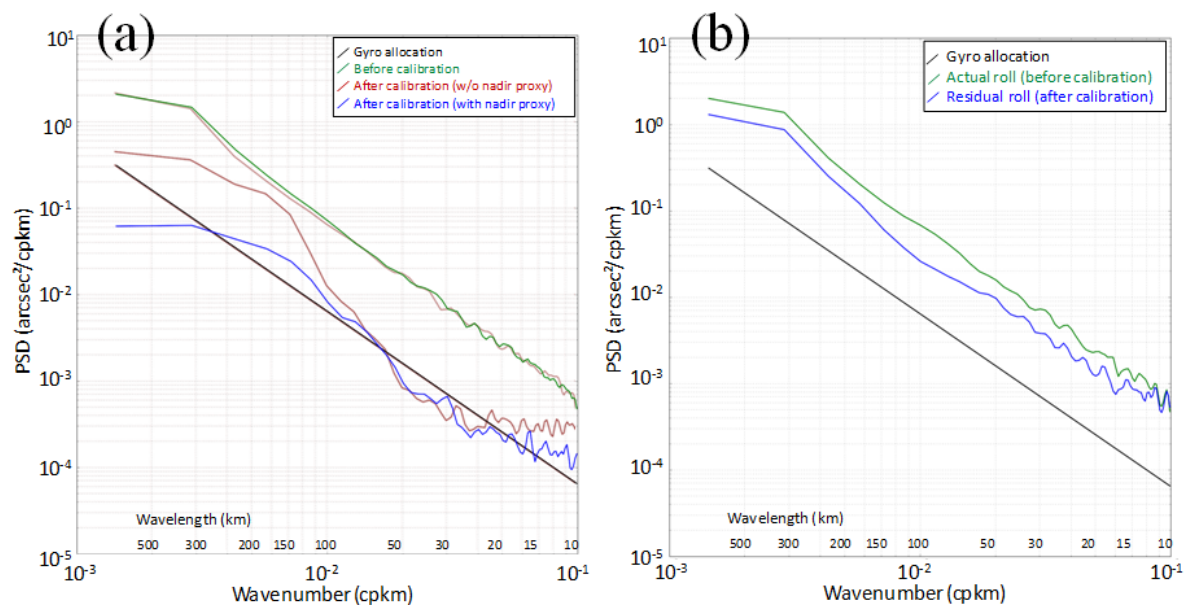


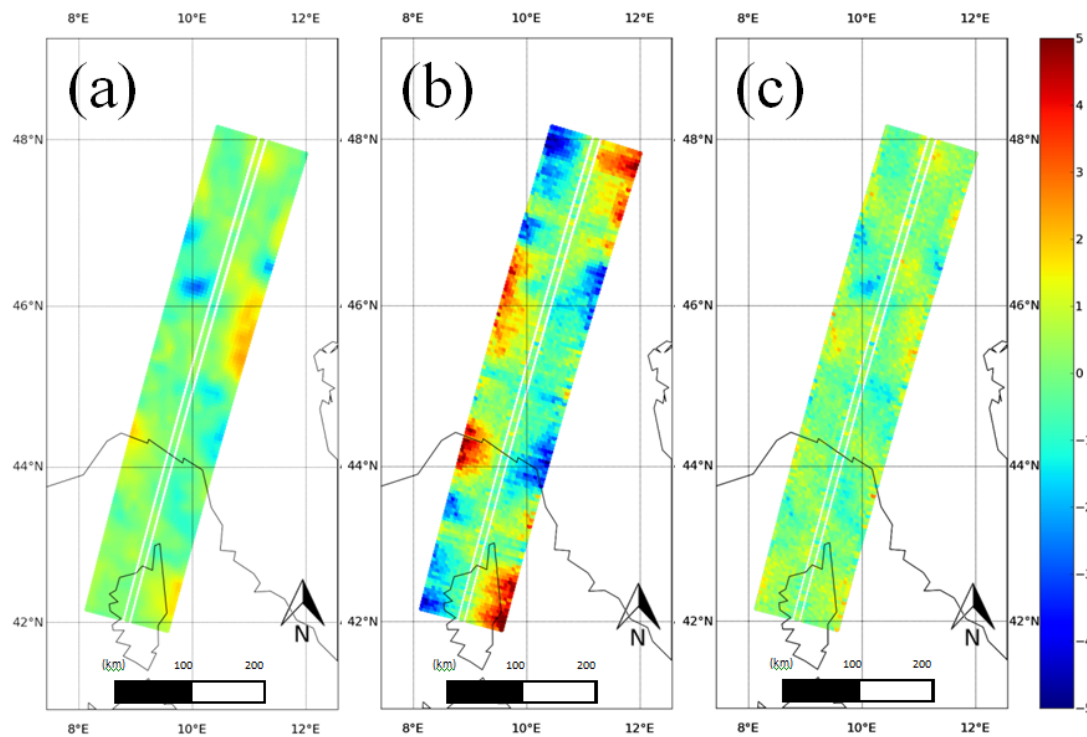
Figure 13 shows an example of KaRIN error-free measurement (Figure 13a), the same scene with the noise and roll errors from our “worst-case” scenario (Figure 13b), and after we apply the direct method calibration (Figure 13c). The residual roll after calibration error is very small, and many features from Figure 13a can be observed in Figure 13c (in contrast with the uncalibrated roll in Figure 13b). Nevertheless, a closer comparison of Figure 13a,c shows that some mesoscale features ranging from 50 to 200 km are affected (altered by the calibration as they leak into the roll estimate).

While this model yields good results in our simulations, it relies on a high-quality proxy from the nadir constellation. In other words, this calibration method would create a dependency between SWOT and other satellites, whereas their availability is not yet guaranteed.

Furthermore, because we use only a single image (Equation (2)), not a difference between images (Equation (4)), there is a high risk of absorbing other error-induced cross-track gradients that are currently not simulated in our experiments (rule O2). For instance, these gradients could originate in imperfect ancillary geophysical corrections and references (e.g., ionosphere, wet troposphere, mean sea surface). Our method would empirically mitigate some of these gradients (e.g., residual of troposphere errors), although without a good control of the covariance model and error residuals.

Lastly, the performance of this method likely depends on the geographical variations of the oceanic variability (discussed in Section 5) and our ability to properly model it in the covariance matrix  $C_{vv}$ .

**Figure 13.** Example of cross-calibration with the direct method. (a) shows the “true” topography signal (after the large scale proxy from nadir altimetry is removed). (b) shows the measured topography once roll and white noise are added. (c) shows the topography after the direct cross-calibration scheme is used. Unit: cm.



#### 4.2. Collinear Method

The collinear method (schematics of Figure 7) uses the complete overlap between the images from two subsequent cycles. It is designed for the CalVal/commissioning orbit (1-day repeat cycle): we use the daily revisit time to cancel out a large fraction of the oceanic variability  $H_{\text{real}}$  and systematic errors  $\epsilon$  (as per rule O2). In contrast, the baseline orbit has a very long cycle of 21 days so there is no benefit to compute image differences between subsequent cycles since the bulk the ocean variability and measurement errors would decorrelate (*i.e.*, not cancel out in the difference). A major advantage of this method is that it can be used at all latitudes.

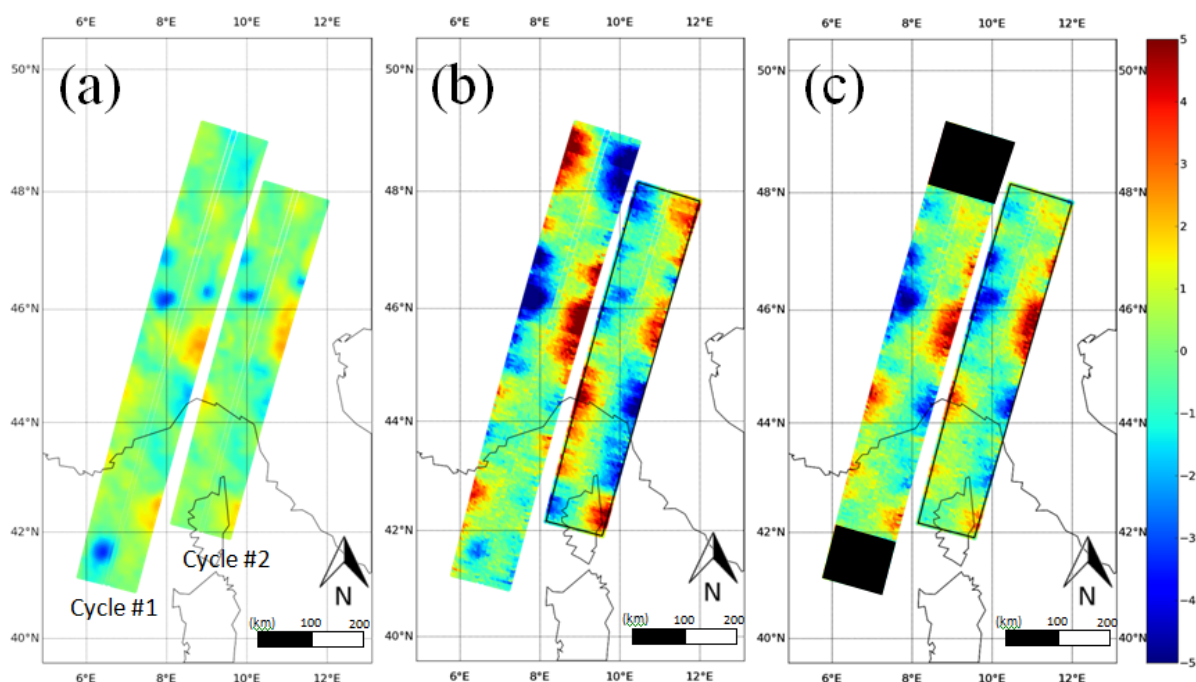
Figure 12b shows the average performance of the collinear method for the Cal/Val orbit: the cross-calibrated roll (blue) is smaller than the uncalibrated roll (green) by a factor of 2 except for small wavelengths. This is significantly less than what we obtained with the direct method whereas the geometry of the scene is the same (entire swath).

The reason for this lower performance is given by the example of Figure 14. Figure 14a shows the true error-free topography, Figure 14b is the topography with noise and roll errors, and Figure 14c is the output of the collinear calibration. While Figure 14c clearly shows a slight reduction of the overall roll error, it is also clear that the residual rolls after calibration are exactly the same for both cycles. This is explained by the geometry of the calibration overlap and Equation (4): for the collinear method,

$x$  and  $x'$  are strictly equal because the one-day revisit is an exact repeat cycle. Consequently, it is impossible to distinguish  $R$  from  $R'$  since their cross-track linear models are identical. As a result, instead of computing  $R$  and  $R'$ , the collinear method can only compute  $\delta R = R(t) - R'(t')$  from  $\delta H_{\text{obs}}$ . In other words, this method projects an average of  $R$  and  $R'$  on each swath, leading to a factor 2 in the mean variance reduction.

The collinear method would be efficient to estimate the spectrum of the roll error during SWOT's CalVal phase, but it cannot be used alone as a major cross-calibration mechanism because of the ambiguity between  $R(t)$  and  $R(t + \Delta t_{\text{cycle}})$ . Moreover, the method is less effective for small wavelengths because our covariance matrixes account for the one-day mesoscale variability and instrumental noise (topography observations are not fully trusted at these wavelengths).

**Figure 14.** Example of calibration with the collinear method. (a) shows the “true” topography signal (after the large-scale proxy from nadir altimetry is removed) for two consecutive one-day cycles on SWOT's fast sampling orbit (cycle #2 is shifted by  $2^\circ$  or both images would be overlapping). (b) shows the measured topography once roll and white noise are added. (c) shows the topography after the collinear cross-calibration scheme is used. Although there is an overall reduction of the roll error, the residual rolls from cycle #1 and #2 are identical because the collinear method is numerically unable to separate roll signals from consecutive cycles. Unit: cm.



#### 4.3. Crossover Method

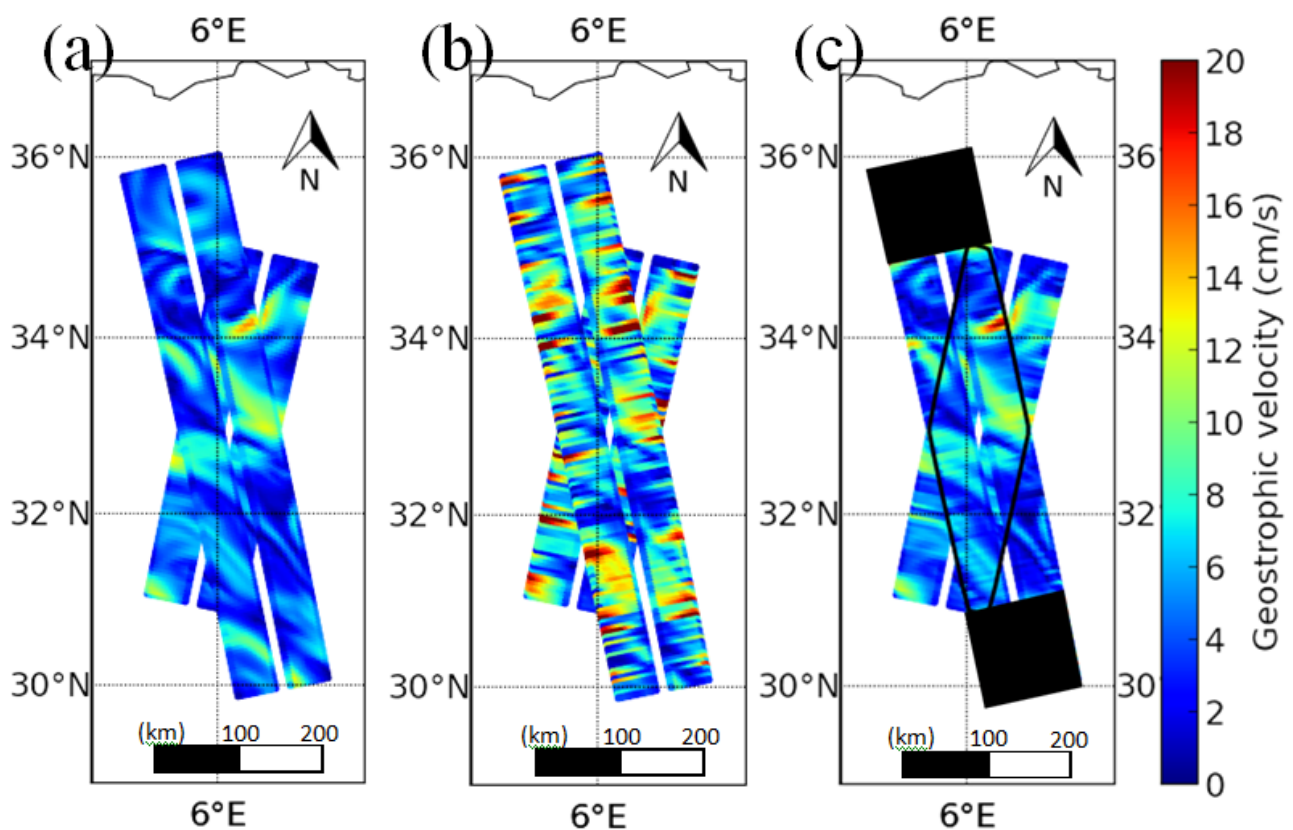
The crossover method was extensively analyzed by Dibarboure *et al.* [1]: it uses overlaps between two KaRIN swaths or between a KaRIN swath and the 1D profile of a nadir altimeter (e.g., Jason-CS). The former produces diamond-shaped calibration zones whereas the latter produces 1D calibration segments (when the Jason-CS track is in the SWOT swath). Figure 15 shows an example of a one-day crossover calibration zone: Figure 15a shows the error-free geostrophic velocities derived from



KaRIN's topography, Figure 15b is affected by the uncalibrated roll error (“worst-case” scenario), and Figure 15c is after cross-calibration. Here the cross-calibration mechanism is able to recover most of the signal of interest in a calibration diamond of approximately 400 km long in the along-track direction.

This example highlights two important parameters affecting the crossover method: the length of calibration (depending on the crossover angle and sub-satellite track geometry) and the time difference  $\delta t = t - t'$  between image acquisitions. Both parameters vary with latitude.

**Figure 15.** Example of calibration with the crossover method. Geostrophic velocity field from the high-resolution model of Klein *et al.* [8]: error-free measurements (a), with uncalibrated roll (b), and after crossover calibration (c).



#### 4.3.1. Calibration Zones

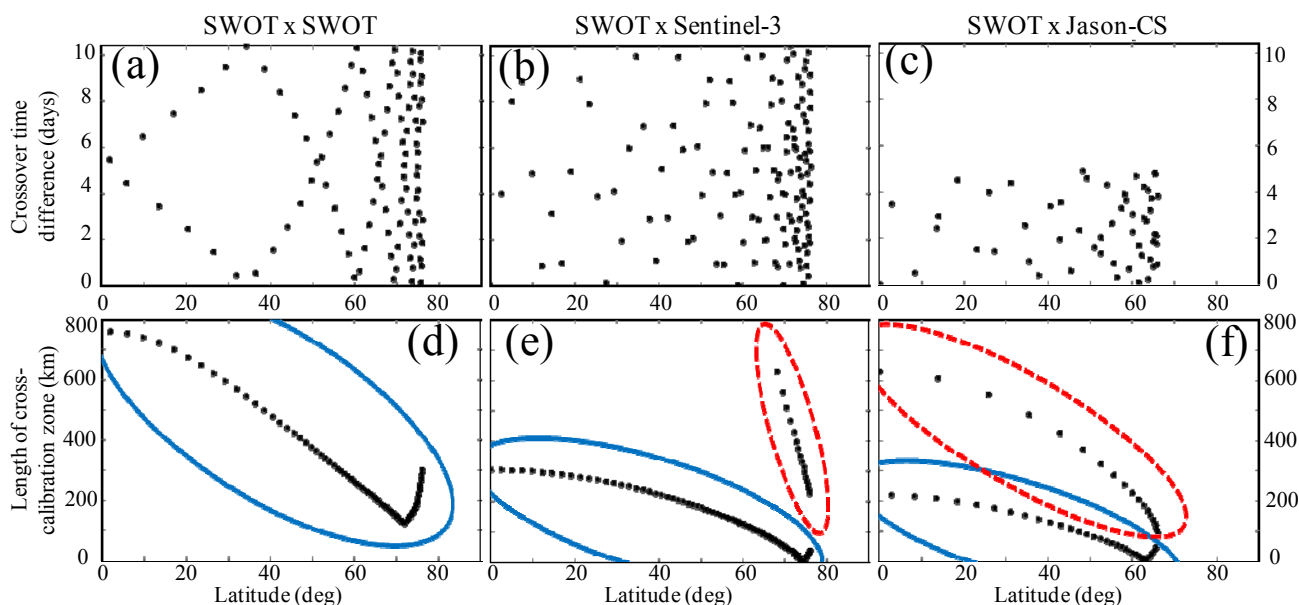
Figure 16 shows the distribution of these parameters for mono and multi-satellite crossovers. Figure 16a–c are for the crossover time difference in days, and Figure 16d–f are for the (along-track) length of the crossover zones in kilometers.

The length of a KaRIN × KaRIN calibration diamonds range from 120 km (*i.e.*, orthogonal crossover), to 750 km (*i.e.*, narrow crossover angle). KaRIN × nadir crossovers are generally shorter as they range from 1 km (*i.e.*, orthogonal crossover) to 600 km. As far as multi-mission crossover segments are concerned, there are ascending/descending crossovers (blue ellipses in Figure 16e,f), which are shorter than 300 km, and ascending/ascending or descending/descending crossovers, which are longer than 100 km (red dashed ellipses).

The crossover time difference of SWOT's baseline orbit is structured with a “butterfly shape” as shown by Figure 16a: each crossover with more than five days or less is located between two crossovers with five days or less. In contrast, multi-mission crossovers are much more random and they range from a few hours to half the shortest cycle (10.5 days for Sentinel-3/SWOT and five days for Jason-CS/SWOT).

There are many combinations of the three parameters controlling crossovers: diamond or segment, geometry and time difference. The complexity and diversity of situations is illustrated by Figure 17, which shows an overview of the mono-satellite (left panels) and multi-mission (right panels) calibration zones as a function of latitude. Figure 17a,b display each crossover zone as a segment (stacked arbitrarily if there are two or more in a given latitude), and the segment color shows the crossover type (purple is for SWOT diamonds, green and red for nadir segments), Figure 17c,d show the time difference in days. Lastly Figure 17e,f show the number of concurrent crossover zones (*i.e.*, size of the stacks from Figure 17a–d).

**Figure 16.** Overview of the crossover distribution for SWOT's 10-day orbit. (a–c) show the crossover delta-time (time difference between co-located pixels in days) and (d–f) show the along-track length of the SWOT segment that is in the calibration zone (unit: km). The blue ellipses (plain) highlight crossovers between ascending/descending arcs and the red ellipses (dashed) highlight crossovers between ascending/ascending or descending/descending arcs.



If crossover calibration is tackled with SWOT data only, each roll angle  $R(t)$  is seen (on average) by two independent cross-calibration zones. This number increases slowly to 3 at mid latitudes and 5 or more above  $70^\circ$ . For almost all latitudes, there is at least one segment with a time difference smaller than five days (as per the “butterfly shape” of Figure 16a). In contrast, if Jason-CS and Sentinel-3 crossovers are used as well, each roll angle  $R(t)$  is seen by 4 zones on average and 5 to 20 for mid-latitudes to high-latitudes respectively. The additional crossover segments from nadir altimetry

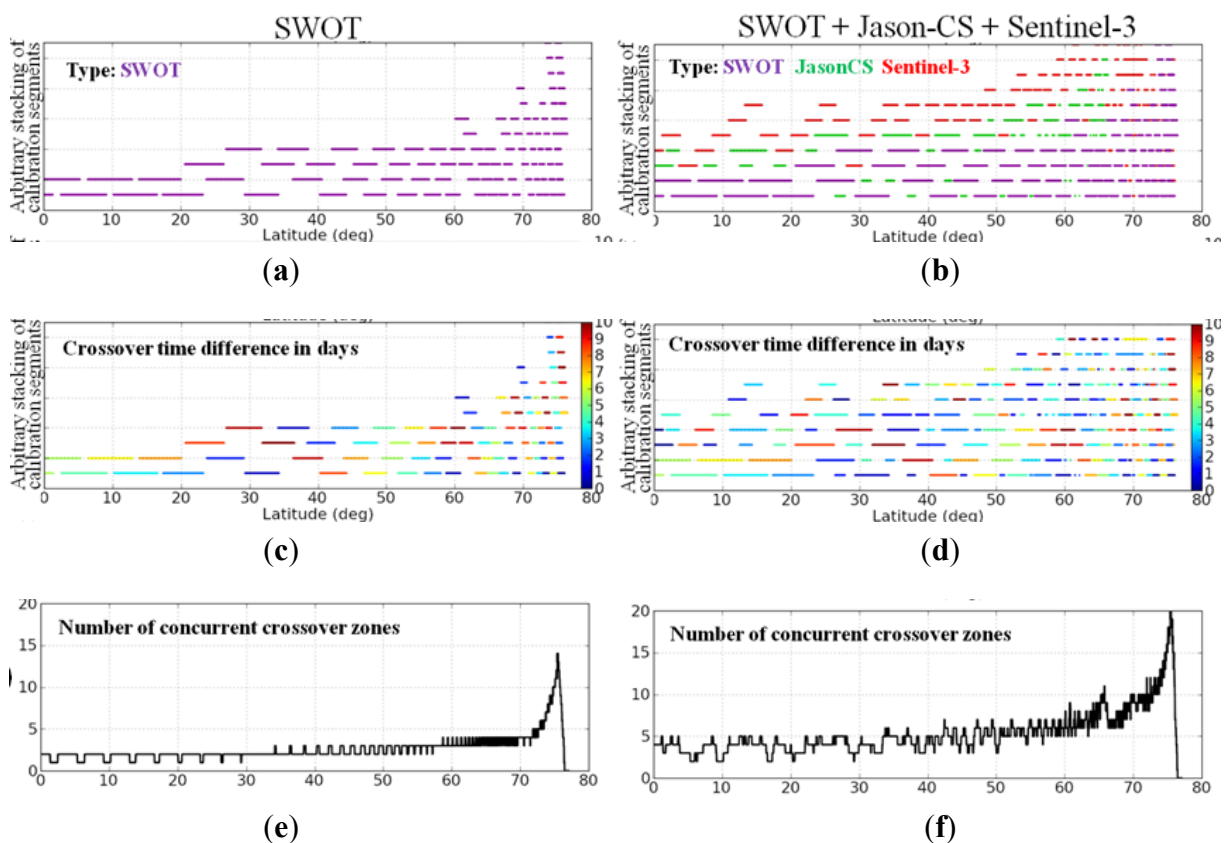


also increase the probability to find multiple crossover zones with a time difference less than two days at all latitudes.

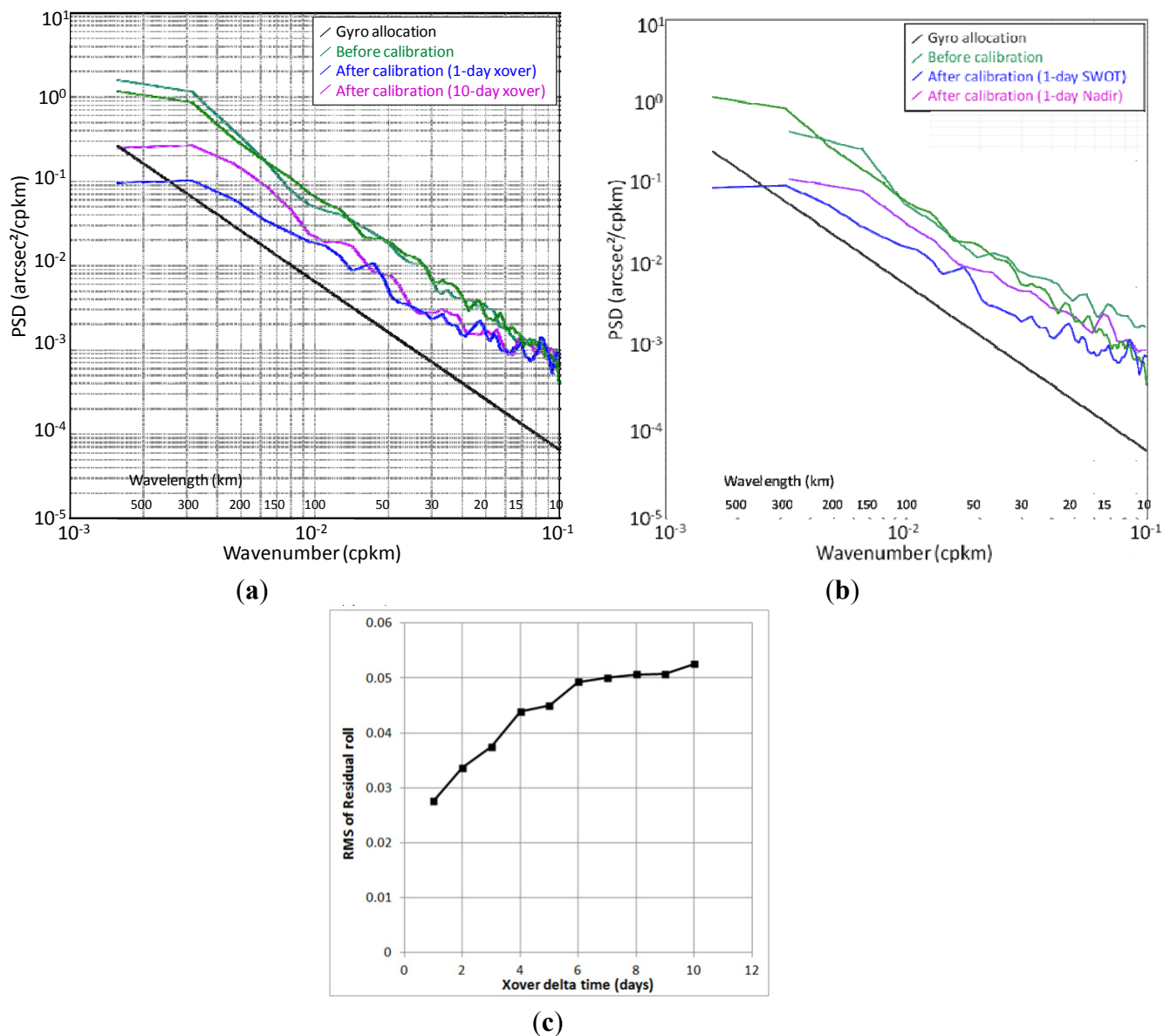
#### 4.3.2. Performance

From the observations of Section 4, we anticipate that the performance of this method highly depends on the crossover time difference (rule O2), and the example of Figure 15 was easy to solve because it was a one-day crossover (oceanic variability leakage was limited). Figure 18a shows the mean spectrum before (green) and after calibration for a one-day crossover (blue) and a 10-day crossover (purple). Both simulations yield a reduction of roll signatures by a factor of 2 to 5 for all wavelengths longer than 20 km. Shorter wavelengths are not well corrected for two reasons: the calibration zones are thinner at the tip of each crossover diamond (extensively discussed in Figure 17 of Dibarboure *et al.* [1]), and because smaller submesoscale features are also faster (*i.e.*, not cancelled in the crossover difference).

**Figure 17.** Overview of crossover calibration zones if SWOT is used alone (**left panel**), or in combination with Jason-CS and Sentinel-3 (**right panel**). In the upper panels, each colored segment is one calibration zone (stacked arbitrarily if there are two or more at a given latitude). New types of crossover zones (**a,b**) can be used with external nadirs, thus increasing the probability to find crossover with a short time difference between co-located measurements (**c,d**), increasing the number of overlapping calibration zones seen by KaRIN's image (**e,f**).



**Figure 18.** Performance of the crossover calibration method. (a) shows the roll error PSD after calibration for a 1-day KaRIN  $\times$  KaRIN crossover (blue) and a 10-day crossover (purple), and the roll error before calibration (green for both simulations). (b) shows the roll error PSD after calibration for a 1-day KaRIN  $\times$  KaRIN crossover diamond (blue) and a one-day KaRIN  $\times$  Jason-CS crossover segment (purple), and the roll error before calibration (green for both simulations). (c) shows the RMS of the uncorrected roll as a function of crossover diamond delta time (unit: arcsec).



Furthermore, there is a factor of two between the 1-day and 10-day configurations, and the improvement is observed primarily in the longer wavelengths. Figure 18c also gives some insights on the increase of residual roll after calibration as a function of time difference. The RMS of roll residuals after calibration increases almost linearly from one to six days then more slowly from 6 to 10 days. This might be explained by the  $H_{\text{ref}}$  proxy here build with shorter temporal correlation scales (less than 10-days) than a global AVISO map (of the order of 15 days) which might allow the proxy to resolve a fraction of the variability from 6 to 10 days (rule O3).

Figure 18b compares the average residual roll for KaRIN  $\times$  KaRIN diamonds (blue) and KaRIN  $\times$  Jason-CS segments (purple) for one-day best-case configurations. The nadir  $\times$  KaRIN residual roll error is larger by a factor of 2 for almost all wavelengths. 1D crossover profiles are less attractive for roll calibration purposes: the spectral separation of roll and mesoscale signals leverages the 2D information in images (2D model  $M$  and 2D covariance  $C_{vv}$  in Equation (6)) whereas this separation is limited in 1D nadir profiles (rule O1).

These results emphasize the limits of the cross calibration presented on Figure 17: using multi-mission crossover does increase the number of calibration zones and/or reduce the average crossover delta time, but nadir/crossover segments are not as efficient as a KaRIN diamonds even when the time difference is the same.

#### 4.4. Subcycle Method

SWOT's baseline orbit features 292 revolutions per repeat cycle, leading to approximately 135 km between adjoining sub-satellite tracks at the equator while KaRIN has a 120 km field of view. However, sub-satellite tracks get closer at higher latitudes so adjoining swaths overlap as illustrated by Figure 19a. The overlap zone can be used to form Equation (4) and to solve the problem with Equation (6) like with crossovers.

Figure 19b–d gives an example of the sub-cycle calibration in a favorable configuration (55°N, contingency orbit). Figure 19b shows the “true” velocity field with submesoscale features. Once the uncalibrated roll error (“worst-case” scenario) is added, velocities cannot be recovered (Figure 19c). After we solve the problem for roll using the overlap between adjoining swaths, the bulk of the roll error is accounted for (Figure 19d), and most of the oceanic features can be observed.

**Figure 19.** SWOT's sub-cycle cross-calibration. (a) shows the principle of sub-cycle overlaps. (b–d) show an example of cross-calibration with the sub-cycle method (simple problem: 55° latitude, one-day sub-cycle orbit). Geostrophic velocity field from the high-resolution model of Klein *et al.* [8]: error-free measurement (b), with uncalibrated roll (c), and after sub-cycle calibration (d).

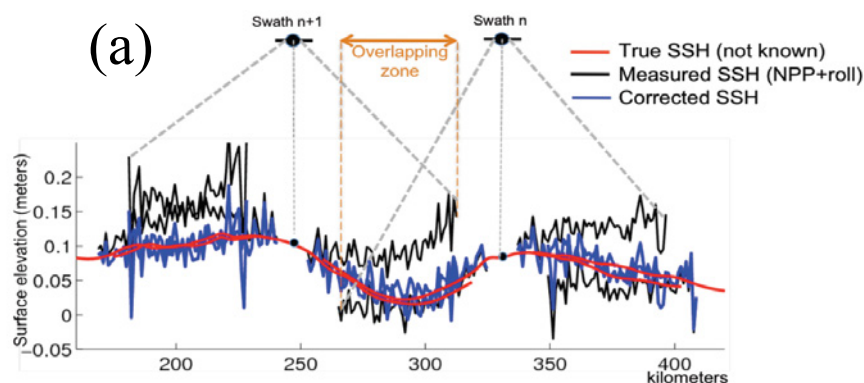
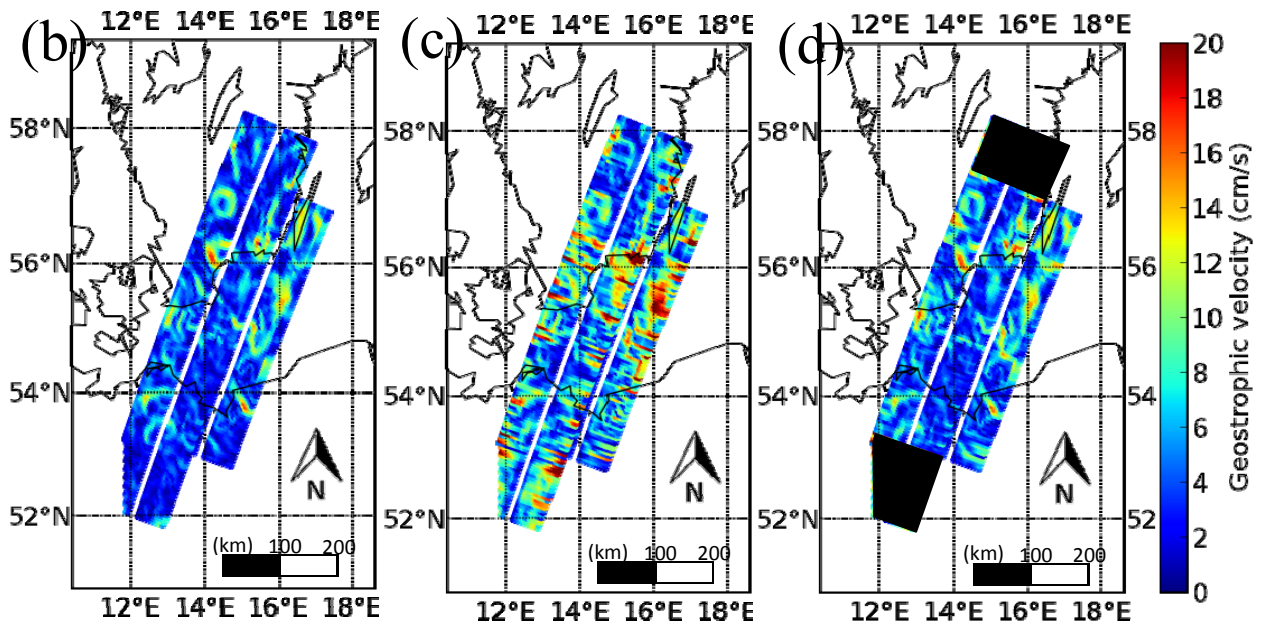


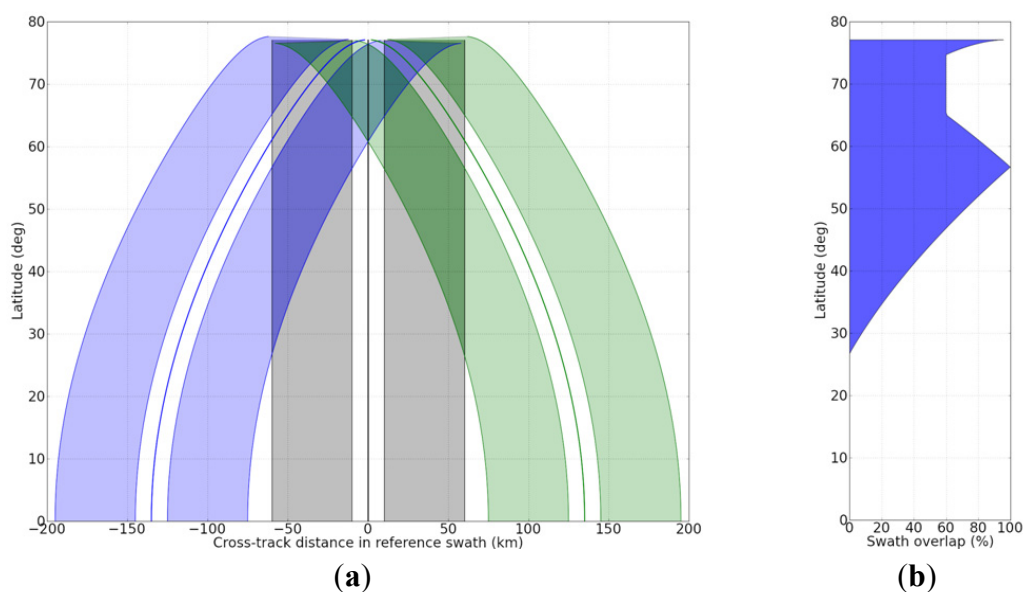
Figure 19. Cont.



#### 4.4.1. Calibration Zones

Figure 20a gives some insight on the geometry of the overlap between adjoining swaths, and Figure 20b on their magnitude. With 21-day repeat orbits and a 120 km field of view, the cross-track overlap starts at a latitude of approximately  $26^\circ$ , then the extent of overlap slowly increases to 50% of the swath (30 km) at  $45^\circ$  and up to 100% (60 km) at  $55^\circ$  where the overlap is complete (each pixel is common to adjoining swaths). Then the extent of the overlap decreases from 100% (60 km) to only 60% (36 km) between  $65^\circ$  and  $75^\circ$  because of the near-nadir gap.

**Figure 20.** Geometry of sub-cycle overlaps. (a) describes their geometry as a function of latitude and (b) shows the percentage of the reference swath (grey in (a)) that is also being measured by adjoining swaths (green and blue in (a)).



In the along-track direction, calibration zones are seamless and span over thousands of kilometers. In the example from Figure 19, we chose to limit the scene to approximately 200 km for practical reasons but the same inversion could be performed an entire pass/revolution if necessary.

#### 4.4.2. Performance

The time difference between overlapping images is defined by the orbit's largest sub-cycle, *i.e.*, 10.9 days for SWOT's baseline orbit and one day of the contingency orbit. From rule O2, we can anticipate that the latter is more favorable because systematic errors and oceanic signals are better removed by the difference between the two images.

This is quantified in Figure 21a: the blue spectrum (contingency orbit) exhibits roll residual up to five times smaller than the purple spectrum (baseline orbit with 10-day sub-cycle). The difference is higher for wavelengths longer than 100 km, but it is significant down to 20 to 30 km.

While using this calibration always reduces roll on average, the benefit is limited to wavelengths longer than 30 km in the case of the baseline orbit. This is an observability limit from rule O1. The calibration zone is thinner (50 to 60 km wide) than for the direct or crossover methods (120 km), and the 10.9-day sub-cycle does not cancel much mesoscale in the overlap difference. Consequently, either input observations are not trusted because they contain mesoscale (*i.e.*, roll is underestimated) or mesoscale leaks on roll (*i.e.*, roll is removed, but mesoscale is distorted).

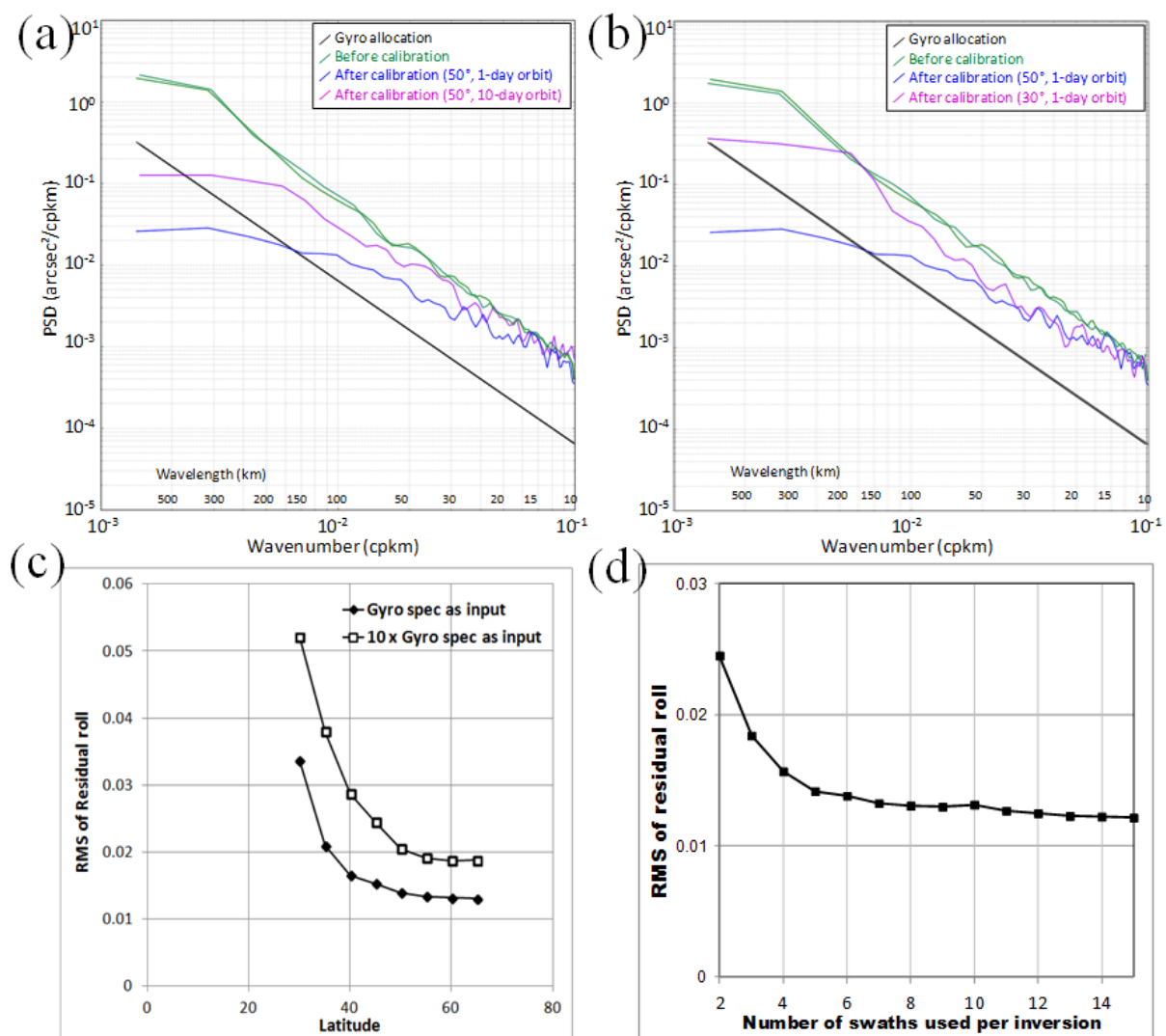
The same rule applies to the contingency orbit, but the short temporal difference eliminates a larger fraction of the signal, essentially achieving the spectral separation between mesoscale and roll down to 15–20 km, and enhancing the calibration's ability to correct roll with little leakage on small/sub mesoscale.

Nevertheless, Figure 21a is computed at 50°, of latitude where the overlap is maximum (Figure 20b). This case is representative of the performance at higher latitudes. However, for lower latitudes, the extent of the overlap is limited (30 km or less) and Figure 21b shows that the performance of the sub-cycle methods decreases as per rule O1.

Even though the one-day contingency orbit is used for both the blue (latitude 50°) and purple (latitude 30°) spectra, the latter yields roll residuals that are an order of magnitude larger at low latitudes. The difference primarily affects wavelengths larger than 30 km because of the geometry: at 30° the overlap is limited to 10 to 20 km which is too short to separate roll cross-track trends from biases or very small changes in mesoscale. At low latitudes the baseline 10-day sub-cycle orbit is barely efficient (not shown) because the degradations from spatially limited overlap (rule O1) and the 10-day distance (rule O2) add up.

Lastly, Figure 21c gives some insights on the changes in roll residuals as a function of latitude for the baseline roll scenario (*i.e.*, current allocation in SWOT's requirements) in black and the "worst-case" scenario (roll ten times larger for all wavelengths) in white. The pessimistic worst-case scenario yields larger roll residuals: the residual RMS is 50% to 80% larger (down from 300% before calibration), but the latitude dependency is somewhat similar: roll residuals are almost three times larger at low latitudes than at mid to high-latitudes, and the performance becomes stable between 45% and 55° when the extent of the overlap reaches 50% to 100% of the swath (Figure 20b).

**Figure 21.** Performance of the sub-cycle calibration method. (a) shows the roll error PSD after sub-cycle cross-calibration at a latitude of  $50^\circ$  for the baseline 10-day sub-cycle orbit (purple) and for the contingency one-day sub-cycle orbit (blue), and the roll error before calibration (green for both simulations). (b) shows the roll error PSD after sub-cycle cross-calibration for the one-day sub-cycle orbit at a latitude of  $50^\circ$  (blue) and at a latitude of  $30^\circ$  (purple), and the roll error before calibration (green for both simulations). (c,d) show in arcsec the RMS of the calibrated roll residuals as a function of latitude (c), and as function of the number of adjoining swaths used concurrently in an inversion performed at  $45^\circ$  (d).



#### 4.5. Summary

##### 4.5.1. Baseline Orbit

There is a wide range of conditions and multiple drivers for the performance of the four calibration methods discussed above. To that extent, Table 1 gives an overview of the calibration problem in a consistent simulation framework and for the baseline orbit.



The upper section of Table 1 describes the basic geometry of each calibration zone and the number of concurrent zones, *i.e.*, whether  $R(t)$  can be estimated independently more than once by each method for a given time step. The table also describes the temporal variability and how much low frequency ocean topography and/or errors are removed with an image-to-image difference.

The rules from Section 3.4 explain the average performance described in the middle section of Table 1. The results are given for three ranges of wavelengths, and sometimes with best/worst case ranges when a given calibration yields geographically varying results. The performance is here given as a ratio between the uncalibrated and calibrated spectra (e.g., a value of 10 means the roll *energy* is divided by 10, *i.e.*, its *amplitude* is divided by a factor of  $\sim 3$ ). All results are based on our “worst-case” scenario where an improvement factor of 10 or more is needed to guarantee that the calibrated roll is within the current error budget for roll.

**Table 1.** Synthesis for baseline orbit (21-day repeat, 10-day sub-cycle).

		Cross-Calibration Method				
		Direct	Crossover (KaRIN only)	Crossover (Nadir)	Sub-Cycle	
Description of calibration zones	How many concurrent zones	1	2 ( $<50^\circ$ ) to 4 ( $>60^\circ$ )	4 ( $<50^\circ$ ) to 8 ( $>60^\circ$ )	2 ( $<60^\circ$ ) or 3 ( $>60^\circ$ )	No benefit from cycle-to cycle difference on this orbit
	Geometry	Entire swath	Diamond	1D Segment	Ribbon	
	Temporal variability	n/a	0 to 10.9 days	0 to 10.9 days	10.9 days	
	Along-track (max)	$>10,000$ km	100 km (near $70^\circ$ ) 800 km (equator)	0–300 km rarely up to 600 km	$>5,000$ km	
	Across-track	Entire swath	up to 130 km (diamond center)	1–8 km (nadir footprint)	1 km ( $26^\circ$ ) to 60 km (high latitudes)	
Order of magnitude of the roll reduction in contingency (PSD uncalibrated / PSD calibrated)	$1 < l < 30$ km	5 to 15	1 to 2	0 to 2	1 to 1.5	n/a
	$30 < l < 150$ km	15 to 7	2 to 5 (best) 1.5 to 4 (worst)	1.5 to 2	1.5 to 2 (best) 1 to 1.5 (worst)	
	$150 < l < 500$ km	7 to 20	5 to 10 (best) 2 to 5 (worst)	1.5 to 2	2 to 20 (best) 1 to 5 (worst)	
	Geographical variability	n/a	Moderate (butterfly deltaT variations)	n/a	High	
Best VS worst zones	Best latitudes	n/a	$20^\circ$ – $45^\circ$ , $55^\circ$ – $65^\circ$ , $>75^\circ$	n/a	$45^\circ$ – $65^\circ$ , $>75^\circ$	n/a
	Worst latitudes	n/a	$<20^\circ$ , $45^\circ$ – $55^\circ$ , $65^\circ$ – $75^\circ$	n/a	$<40^\circ$	
	Void latitudes	n/a	n/a	n/a	$<26^\circ$	
	Reliance on external nadirs	Very High for $l < 100$ km	Moderate	Very high	High (10-day subcycle)	

In general, the direct methods yields better coverage and smaller calibrated roll residuals because the full extent of the swath can be used for calibration (rule O1 > rule O2), but this method makes an extensive use of the external nadir constellation to separate roll from oceanic variability (rule O3) and is also more prone to leakages from oceanic variability on roll (the true SSH signal is more likely to be corrupted if it is misinterpreted as roll) and poor modeling of the true topography spectrum.

The other methods are complementary and provide an additional calibration capability, albeit with complex latitude-varying coverage and performance. The main asset of using three or four methods is to minimize systematic leakages from oceanic variability or measurement errors by maximizing local overlap zones and by minimizing the temporal variability with one-day to three-day image differences.

Previous sections inferred that there is a wide range of geometric and temporal conditions and the geographical variability is summarized at the bottom section of Table 1. Note that the geographical variations of the ocean dynamics are not considered here (our SSH inputs are always normalized to the median ocean variability observed in altimetry maps), nor the geographical variations of the instrumental error (we simulated white noise only, and with a constant standard deviation). To that extent, this table likely omits even more geographical complexity from oceanic and measurement conditions.

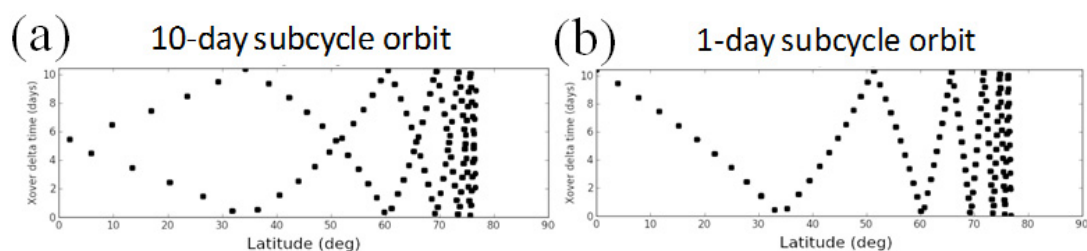
#### 4.5.2. Contingency Orbit

Previous sections highlighted that changing the orbit parameters significantly affects the geometry and/or temporal distribution of the input measurements used for roll calibration. Table 2 provides an overview for the contingency orbit. Various items are highlighted in bold red when there is a notable difference with Table 1 and the baseline orbit.

The baseline and contingency orbits are very similar except in the value of their largest sub-cycle. The sub-cycle method efficiently exploits the one-day sub-cycle of the contingency orbit. This method yields better results and the improvement ranges from a factor of 2 for short wavelengths to 10 or more for longer wavelengths. Most importantly, with the contingency orbit, this method does not require to rely on external nadirs and a  $H_{\text{ref}}$  proxy of  $H_{\text{real}}$ .

To that extent, this orbit is an attractive asset as a self-sufficient (SWOT only) contingency plan for calibration if SWOT processors cannot rely on external missions such as Jason-CS or Sentinel-3.

**Figure 22.** Distribution of the crossover delta-time as a function of latitude for the “nominal” SWOT orbit (a) and for the “backup” SWOT orbit (b). Unit: days.



However there is a major caveat to using this contingency orbit: the KaRIN  $\times$  KaRIN crossover distribution is less attractive. This is illustrated on Figure 22: the butterfly-shaped distribution of crossover time differences ( $\delta t$ ) as function of latitude is replaced by a V-shaped distribution. Although



the baseline orbit ensured that crossovers with short  $\delta t$  were present at all latitude bands (ensuring a stability in the longer wavelengths), this is no longer the case with the contingency orbit where large bands of latitude have no crossover with  $\delta t$  shorter than three to five days. This distribution results in a significant aggregation of the “favorable” and “unfavorable” latitude bands with the contingency orbit.

**Table 2.** Synthesis for contingency orbit (21-day repeat, 1-day sub-cycle).

		Cross-Calibration Method				
		Direct	Crossover (KaRIN only)	Crossover (Nadir)	Sub-Cycle	
Description of calibration zones	How many concurrent zones	1	2 (<50°) to 4 (>60°)	4 (<50°) to 8 (>60°)	2 (<60°) or 3 (>60°)	No benefit from cycle-to cycle difference on this orbit
	Geometry	Entire swath	Diamond	1D Segment	Ribbon	
	Temporal variability	n/a	0 to 10.9 days	0 to 10.9 days	<b>1 day</b>	
	Along-track (max)	>10.000 km	100 km (near 70°) 800 km (equator)	0–300 km rarely up to 600 km	>5.000 km	
	Across-track	Entire swath	up to 130 km (diamond center)	1–8 km (nadir footprint)	1 km (26°) to 60 km (high latitudes)	
Order of magnitude of the roll reduction in contingency (PSD uncalibrated / PSD calibrated)	1 < l < 30 km	5 to 15	1 to 2	1 to 2	<b>1 to 3</b>	n/a
	30 < l < 150 km	15 to 7	2 to 5 (best) 1.5 to 4 (worst)	1.5 to 2	<b>3 to 10 (best)</b> <b>1 to 3 (worst)</b>	
	150 < l < 500 km	7 to 20	5 to 10 (best) 2 to 5 (worst)	1.5 to 2	<b>10 to 100 (best)</b> <b>1 to 5 (worst)</b>	
Best VS worst zones	Geographical variability	n/a	<b>Very high (V-shaped deltaT variations)</b>	n/a	High	n/a
	Best latitudes	n/a	20°–45°, 55°–65°, >75°	n/a	45°–65°, >75°	
	Worst latitudes	n/a	<20°, 45°–55°, 65°–75°	n/a	<40°	
	Void latitudes	n/a	n/a	n/a	<26°	
	Reliance on external nadirs	Very High for 1 < 100 km	<b>High (xover with deltaT &lt; 5d latitude gaps)</b>	Very high	<b>Very low (1-day sub-cycle)</b>	n/a

Furthermore, the situation is very detrimental to latitudes below 20° because they are barely covered by the sub-cycle method, and because the crossover distribution is always larger than 5 to 10 days. In these areas, the direct method and nadir crossovers might be the main asset to mitigate roll.

In other words, the contingency orbit significantly focuses roll calibration on one method and at certain latitudes (e.g., limited observation area to secure a “threshold mission”).

#### 4.5.3. Commissioning/CalVal Orbit

Table 3 gives the same overview for the orbit that would be used for the commissioning (*i.e.*, Calibration/Validation) phase of SWOT. Its exact one-day repeat is very different from the baseline orbit because the spatial coverage is extremely sparse (approximately 5% of the globe is observed) in order to decrease the revisit time to one day.

**Table 3.** Synthesis for commissioning/CalVal orbit (exact 1-day repeat).

			Cross-Calibration Method			
			Direct	Crossover (KaRIN only)	Crossover (Nadir)	Sub-Cycle
Description of calibration zones	How many concurrent zones	1		<b>1 (very sparse coverage)</b>	4 (<50°) to 8 (>60°)	
	Geometry	Entire swath		Diamond	1D Segment	
	Temporal variability	n/a		<b>0 to 1 day</b>	<b>0 to 1 day</b>	Non-existent on this orbit
	Along-track (max)	>10.000 km		100 km (near 70°) 800 km (equator)	0–300 km rarely up to 600 km	
	Across-track	Entire swath		up to 130 km (diamond center)	1–8 km (nadir footprint)	
Order of magnitude of the roll reduction in contingency (PSD uncalibrated / PSD calibrated)	1 < l < 30 km	5 to 15		1 to 2	1 to 2	
	30 < l < 150 km	15 to 7		2 to 5	1.5 to 2	
	150 < l < 500 km	7 to 20		5 to 10	1.5 to 2	n/a
Best VS worst zones	Geographical variability	n/a		<b>Very sparse coverage (&lt;20%)</b>	n/a	
	Best latitudes	n/a		n/a	n/a	
	Worst latitudes	n/a		n/a	n/a	
	Void latitudes	n/a		<b>Very large (&gt;80%)</b>	n/a	
	Reliance on external nadirs	Very High for 1 < 100 km		<b>Very low (always short deltaT)</b>	Very high	n/a

As a result, adjoining swaths no longer overlap, and crossovers become very rare (approximately 10 per revolution, *i.e.*, less than 10% to 20% of a swath is within a crossover zone) but locally very good (less than one day between ascending and descending images).

Most importantly, with this orbit, the collinear method becomes usable thanks to the very short repeat cycle. Its performance as a *correction* was shown to be limited because the roll signature of subsequent cycles cannot be separated; yet this approach remains extremely useful to *quantify* the mean roll error spectrum: because roll is decorrelated within one day, the roll spectrum is directly linked to the estimated spectrum of the cycle-to-cycle difference with a factor of 2. The knowledge gained on actual roll statistics becomes a valuable *a priori* condition for other methods and for the baseline and contingency orbits.

## 5. Discussion and Outlook

### 5.1. End-to-End Performance and the Fusion Mechanism

The roll calibration values shown in this paper are the output of the second processing step in Figure 7 (local roll estimates from a single calibration zone). The rationale for this choice is to investigate the strengths and weaknesses of each local method and to quantify the link between geometry, temporal variability and roll calibration.

However, Table 1 shows that there are sometimes three or more concurrent and independent calibrations of the same KaRIN pixels (e.g., 1 direct, 2 KaRIN crossovers, 3 nadir crossovers, 2 sub-cycles). The purpose of the third processing step of Figure 7 (described by Dibarboure *et al.* [1]) is to perform the fusion of individual roll estimates. This step generally yields a significant improvement upon independent and local roll estimates.

Consequently, end-to-end roll reduction assessments should be carried only *after the fusion step*. The method used by Dibarboure *et al.* [1] to merge roll estimates is an inverse method. But their analysis is limited because it does not take into account the new methods implemented in this paper, nor the more complex input roll scenario (defined as a full spectrum and not just a simple covariance model), let alone the spectral decomposition of the estimation error of each method.

To that extent, the outlook of our study might be to revisit the fusion mechanism from Dibarboure *et al.* [1] in order to account for the improvements we implemented. As far as implementation is concerned, the performance spectra measured in Section 4 might be a major asset for their inverse method scheme: these spectra trivially determine the relative weight of each calibration zone *for each wavelength*.

### 5.2. Methodology Improvements

In this paper, we used the four calibration methods in an independent way for each zone. In practice however, it is possible to solve a larger numerical problem with multiple methods and/or multiple calibration zones. The increased number of observations should improve the calibration (rule O1).

For instance, the calibration performance reported in Section 4 is based only on a single couple of adjoining images (e.g., grey and green in Figure 20a), which limits the overlap zone to 60 km in the cross-track direction.

However, if roll is solved for 3 or more adjoining images (e.g., blue, grey and green in Figure 20a), the overlap zone increases to 120 km for the center image(s), thus, improving the chances to achieve the spectral separation between roll and oceanic variability (yielding better roll observability and less leakage). This is illustrated by the sensitivity tests of Figure 21d: the residual error after calibration decreases by a factor of 2 in amplitude (or 4 in energy like in Table 2) as we increase the number of images used in each inversion. There is however an asymptotical limit of the order of 0.012 arcsec RMS. As far as spectra are concerned (not shown), the gain is observed on high wavenumbers where leakages on roll (from noise or submesoscale) are not spatially correlated and therefore averaged out when multiple swaths are used concurrently. In contrast, the gain is limited on wavelengths longer than 50–100 km because the leakage from swaths  $N - 1$  and  $N$  (Figure 20a, blue and grey), and the leakage from  $N$  and  $N + 1$  (Figure 20a, grey and green) can be correlated.

Extending the calibration zone by using 3+ datasets from KaRIN can be generalized: multiple nadir and KaRIN crossovers could be solved *concurrently* instead of *independently*. Similarly, multiple methods could be solved concurrently (e.g., crossover + direct + sub-cycle), thus maximizing the extent calibration zones, and therefore roll observability.

The major downside of this approach with respect to the simpler strategy of Figure 7 is the increasing size of the matrix with the number of observations  $H_{\text{obs}}$  (not the number of unknowns  $R$ ). For instance, if the basic sub-cycle method (2 swaths) is solved concurrently with the direct method, the  $C_{\text{vv}}$  matrix is approximately five times larger. The matrix size increases much faster than the number of methods or the size of the calibration zone.

### 5.3. Realistic Input Scenario

Our scenarios and results are interesting to understand the roll reduction mechanism, but they are by no means realistic, let alone a true contingency plan if SWOT's roll is larger than its current error budget. To develop a true contingency plan, one would need to answer the following questions and to set up more realistic simulation conditions:

- What part of the oceanic variability should be preserved? Our SSH derived from Klein *et al.* [8] could be improved upon: larger or even global simulation area to account for regional variations in mesoscale and sub-mesoscale (e.g., Western boundary current VS “eddy desert”), or different physics (e.g., internal tides or waves).
- What is a realistic contingency scenario for uncalibrated roll (e.g., what are the residuals if a high-end gyroscope prediction is used)? Our “worst-case” scenario is quite pessimistic because in reality SWOT's roll is unlikely to be *that* high for *all* frequencies. The benefit of using method X or Y and the benefit of the contingency orbit might be substantially changed by the wavelengths where the roll risk is the highest.
- What is the spectrum of the *other* SWOT errors with a cross-track signature that should be accounted for? Should we even try to separate them from roll? For instance: the wet troposphere path delay could result in a topography error with cross-track gradients (e.g., Ubelmann *et al.* [15]). Should roll-oriented simulations assume that the radiometer-based correction has small cross-track residuals and should roll calibration account for those in an empirical way?

- To what extent can a better  $H_{\text{ref}}$  proxy (Equation (2)) be available by SWOT's launch? To illustrate, CNES has recently confirmed that SWOT would feature a Jason-class altimeter (no longer optional) and this sensor would provide an additional nadir measurement (at the same measurement time as KaRIN) able to improve the constellation proxy from Equation (2) (ours is based on Jason-CS and Sentinel-3). New interpolation methods and ocean models with assimilation could also provide a more efficient alternative to our simple AVISO-like map.
- How does the contingency algorithm need to account for land products and coastal transitions? In this paper, we focused on the ocean, following the rationale of [6] for SWOT's baseline error budget. However this might need to be revisited in a contingency plan: if roll is higher than previously anticipated, inland hydrology measurements might require a more sophisticated algorithm sequence than the baseline outlined by Enjolras *et al.* [7]. Similarly, we ignored the case of coastal transitions, whereas the calibration scenes and the method performance is likely to be very different. This might require the development of land-specific roll calibration and a careful hydrology-oriented and coastal-oriented analysis of the interpolation/propagation mechanism discussed in Section 5.1.

## 6. Conclusions

The roll angle would be a major source of topography error for the proposed SWOT mission: to comply with its stringent requirements, the relative antenna positions must be known within a few micro-meters which might be challenging with onboard roll values. To that extent, we carried out simulations as a contingency strategy (to be ready if roll is larger than anticipated) that could be used with ground-based data.

We revisited the empirical calibration algorithms developed by Dibarboure *et al.* [1] with additional solving methods (e.g., based on orbit sub-cycle and collinear overlaps), and better performance assessments based on spectral decompositions. We also quantified the benefits of using external information from traditional nadir altimeters, and various orbits proposed for SWOT.

We used simple assumptions on other errors and the oceanic variability in order to explore the link between the performance of four calibration methods and the attributes of their respective calibration zones: size and geometry (e.g., crossover diamonds), temporal variability (e.g., how many days between two overlapping SWOT images).

In general, the direct method yields better coverage and smaller calibrated roll residuals because the full extent of the swath can be used for calibration, but this method makes an extensive use of the external nadir constellation to separate roll from oceanic variability and it is also more prone to leakages from oceanic variability on roll (the true SSH signal is more likely to be corrupted if it is misinterpreted as roll) and poor modeling of the true topography spectrum.

For SWOT's baseline orbit (21 days repeat and 10.9 days sub-cycle), the three other methods were found to be complementary with the direct method: KaRIN crossovers, external nadir crossovers, and sub-cycle overlaps were shown to provide an additional calibration capability, albeit with complex latitude-varying coverage and performance. The main asset of using all methods concurrently is to minimize systematic leakages from oceanic variability or measurement errors, by maximizing local overlap zones and by minimizing the temporal variability with one-day to three-day image differences.

To that extent, the “contingency orbit” is an attractive risk reduction asset: the one-day sub-cycle overlaps of adjoining swaths provide a good, continuous, and self-sufficient (no need for external nadirs) calibration scheme. The benefit is however essentially located at mid to high-latitudes and substantial only for wavelengths longer than 100 km.

During the commissioning phase, SWOT would be on a one-day exact repeat orbit for which the collinear method is the most attractive to estimate the mean roll spectrum, but not to correct for it because there is an ambiguity between roll values from subsequent cycles.

The complementarities between multiple calibration methods emphasize the need to perform end-to-end simulations where local results are consolidated by a better merging process that exploits the pre-determined spectral accuracy of each method (*i.e.*, the outputs from this paper). Our results also highlight that further methodology improvements might be possible with a hybridization of the roll estimations methods by solving multiple zones and methods in a single and global inversion procedure.

## Acknowledgments

This work was carried in part at Collecte Localisation Satellites (CLS) with support from the Centre National d’Etudes Spatiales (CNES) in the frame of their TOSCA contracts (SWOT mission development) and in part at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (NASA).

## Author Contributions

Gerald Dibarboure developed the roll calibration with four inverse methods, processed input model snapshots and nadir constellation snapshots, and carried out performance estimations in the spectral domain, and wrote the initial manuscript. Clement Ubelmann helped expanding the research focus, prototyped the sub-cycle method, carried out additional simulations and sensitivity tests with the sub-cycle method and contingency orbit, and assisted with manuscript writing.

## Conflicts of Interest

The authors declare no conflict of interest.

## References

1. Dibarboure, G.; Labroue, S.; Ablain, M.; Fjortoft, R.; Mallet, A.; Lambin, J.; Souyris, J.-C. Empirical cross-calibration of coherent SWOT errors using external references and the altimetry constellation. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 2325–2344.
2. Durand, M.; Fu, L.-L.; Lettenmaier, D.; Alsdorf, D.; Rodriguez, E.; Esteban-Fernandez, D. The surface water and ocean topography mission: Observing terrestrial surface water and oceanic submesoscale eddies. *Proc. IEEE* **2010**, *98*, 766–779.
3. Fu, L.L.; Rodriguez, E. High-Resolution Measurement of Ocean Surface Topography by Radar Interferometry for Oceanographic and Geophysical Applications. In *The State of the Planet: Frontiers and Challenges in Geophysics*; IUGG Geophysical Monograph; American Geophysical Union: Washington, DC, USA, 2004; Volume 19, pp. 209–224.

4. Dibarboure, G.; Schaeffer, P.; Escudier, P.; Pujol, M.-I.; Legeais, J.F.; Faugère, Y.; Morrow, R.; Willis, J.K.; Lambin, J.; Berthias, J.-P.; *et al.* Finding desirable orbit options for the “extension of life” phase of Jason-1. *Mar. Geod.* **2012**, *35*, 363–399.
5. Rosen, P.; Hensley, S.; Joughin, I.R.; Li, F.K.; Madsen, S.; Rodriguez, E.; Goldstin, R. Synthetic aperture radar interferometry. *Proc. IEEE* **2000**, *88*, 333–382.
6. Esteban-Fernandez, D. *SWOT Mission Performance and Error Budget*; NASA/JPL Document (Reference: JPL D-79084); Jet Propulsion Laboratory: Pasadena, CA, USA, 2013.
7. Enjolras, V.; Vincent, P.; Souyris, J.-C.; Rodriguez, E.; Phalippou, L.; Cazenave, A. Performances study of interferometric radar altimeters: From the instrument to the global mission definition. *Sensors* **2006**, *6*, 164–192.
8. Klein, P.; Hua, B.L.; Lapeyre, G.; Capet, X.; le Gentil, S.; Sasaki, H. Upper ocean turbulence from high-resolution 3D simulations. *J. Phys. Oceanogr.* **2008**, *38*, 1748–1763.
9. Dibarboure, G.; Pujol, M.-I.; Briol, F.; le Traon, P.Y.; Larnicol, G.; Picot, N.; Mertz, F.; Ablain, M. Jason-2 in DUACS: First tandem results and impact on processing and products. *Mar. Geod.* **2011**, *34*, 214–241.
10. Chelton, D.B.; Schlax, M.G. The accuracies of smoothed sea surface height fields constructed from tandem altimeter datasets. *J. Atmos. Ocean. Technol.* **2003**, *20*, 1276–1302.
11. Bretherton, F.P.; Davis, R.E.; Fandry, C.B. A technique for objective analysis and design of oceanographic experiment applied to MODE-73. *Deep-Sea Res.* **1976**, *23*, 559–582.
12. Le Traon, P.Y.; Faugère, Y.; Hernandez, F.; Dorandeu, J.; Mertz, F.; Ablain, M. Can we merge GEOSAT follow-on with TOPEX/POSEIDON and ERS-2 for an improved description of the ocean circulation? *J. Atmos. Ocean. Technol.* **2003**, *20*, 889–895.
13. Chelton, D.B.; Schlax, M.G.; Samelson, R.M. Global observations of nonlinear mesoscale eddies. *Prog. Oceanogr.* **2011**, *91*, 167–216.
14. Escudier, R.; Bouffard, J.; Pascual, A.; Poulain, P.-M.; Pujol, M.-I. Improvement of coastal and mesoscale observation from space: Application to the Northwestern Mediterranean Sea. *Geophys. Res. Lett.* **2013**, *40*, 2148–2153.
15. Ubelmann, C.; Fu, L.; Brown, S.; Peral, E.; Esteban-Fernandez, D. The effect of atmospheric water vapor content on the performance of future wide-swath ocean altimetry measurement. *J. Atmos. Ocean. Technol.* **2014**, in press.