IMPROVING THE ESTIMATION OF THE SEA LEVEL ANOMALY SLOPE

C. Mailhes[(1)(4)], *O. Besson*[(2)(4)], *A. Guillot*[(3)] *and S. Le Gac*[(3)]

[(1)] University of Toulouse/INP-ENSEEIHT/IRIT, 2 Rue Charles Camichel, Toulouse, France (corinne.mailhes@tesa.prd.fr)
[(2)] University of Toulouse/ISAE-Supaero, 10 av. Edouard Belin, Toulouse, France (olivier.besson@isae.fr)
[(3)] Centre National d'Etudes Spatiales, CNES, Toulouse, France ([amandine.guillot,sophie.legac]@cnes.fr)
[(4)] Cooperative research laboratory TéSA, 7 Boulevard de la Gare, Toulouse, France

ABSTRACT

Satellite altimeters provide sea level measurements along satellite track. A mean profile based on the measurements averaged over a time period is then subtracted to estimate the sea level anomaly (SLA). In the spectral domain, SLA is characterized by a power spectral density of the form $f^{-\alpha}$ where the slope α is a parameter of great interest for ocean monitoring. However, this information lies in a narrow frequency band, located at very low frequencies, which calls for some specific spectral analysis methods. This paper studies a new parametric method based on an autoregressive model combined with a warping of the frequency scale (denoted as ARWARP). A statistical validation is proposed on simulated SLA signals, showing the performance of slope estimation using this ARWARP spectral estimator, compared to classical Fourier-based methods. Application to Sentinel-3 real data highlights the main advantage of the ARWARP model, making possible SLA slope estimation on a short signal segment, i.e., with a high spatial resolution.

Index Terms— Sea level anomaly (SLA), slope estimation, AR model, frequency warping, spectral analysis.

1. INTRODUCTION

Satellite radar altimetry provides data that can be used for ocean monitoring. Wavenumber spectra of along-track sea level anomalies (SLA) are widely used to analyze different quantities such as the energy cascades between large-scale, meoscale and submeoscale dynamics. As awaited from the quasigeostrophic theory [1], the power spectrum density (PSD) of SLA signals exhibits a decrease of the type $f^{-\alpha}$, starting from some minimal frequency f_1. The parameter α is of great interest in the altimetry community since it characterizes ocean dynamics [2–5]. The SLA signal is usually corrupted by additive noise, classically assumed to be white and Gaussian. Figure 1 displays the estimated PSD of real data from the Sentinel-3 satellite as well as its expected shape (properly fitted): as can be seen, the informative part of the slope area lies in a very small part, around $[0.001, 0.01]$ (equivalent to $30 - 320$ km) of the total normalized frequency interval $[0, 0.5[$, located at very low frequencies, making a log-log representation necessary, while the "noise floor" is clearly visible (horizontal red line σ^2).

Estimation of α can be carried out by a linear regression applied on a cumulated weighted periodogram (Welch periodogram), as shown in Fig. 1 [1–4]. Cumulating several periodograms is necessary to reduce the variance of the estimated spectrum and perform a reliable slope estimation. However, this raises several questions. One of these questions is the stationarity property: how can we guarantee that SLA is stationary (i.e., with the same slope α) along the different SLA segments used in the Welch periodogram? When using a Welch periodogram, the averaged wavenumber spectra are related to Earth footprints that can be distant from some hundreds of kilometers, which questions the stationarity of these spectra. The second question is linked to the inevitable choice of the weighting window in the periodogram: it is well-known that the periodogram is affected by a convolutive bias related to the equivalent spectral window [6]. This affects the slope of the SLA spectrum and thus its estimation highly depends on the window choice.

PSD estimation is an old problem with well-known solutions such as Fourier-based methods. However, the shape of the SLA PSD and the localization of the spectral bands of interest in a very low frequency range calls for specific processing. The aim of this paper is to investigate a new spectral estimation method making possible SLA slope estimation using short signal segments, i.e., with a high spatial resolution and avoiding the critical problem of window choice.

The paper is organized as follows. Section 2 introduces the proposed spectral estimator referred to as ARWARP. A statistical analysis of slope estimation using ARWARP is conducted on simulated SLA signals in Section 3. An application to real data from the Sentinel-3 satellite is presented in Section 4. Conclusions are finally reported in Section 5.

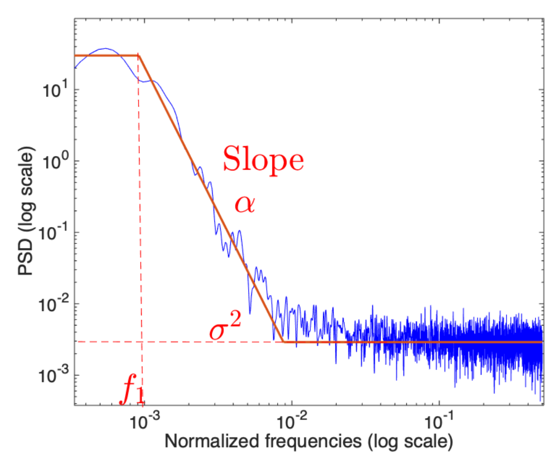

Fig. 1: Estimated Sentinel-3 SLA PSD (in blue) estimated by Welch periodogram [6] on 8 segments of $N = 3000$ samples from Agulhas Current SLA, using a 10% Tukey window applied after detrending, zero-padding by a factor 3, compared with the expected shape in red.

2. THE PROPOSED SPECTRAL ESTIMATOR

2.1. Parametric spectral analysis

Parametric spectral analysis is an effective alternative to non-parametric Fourier-based analysis [6]. It relies on a PSD model depending on a parameter vector $\boldsymbol{\theta}$ so that the PSD estimation $S_x(f;\boldsymbol{\theta})$ amounts to that of $\boldsymbol{\theta}$, yielding an estimate $S_x(f;\hat{\boldsymbol{\theta}})$ where $\hat{\boldsymbol{\theta}}$ is an estimate of $\boldsymbol{\theta}$. Two main benefits have been advocated to support this parametric approach. First, with possibly a small number of parameters describing the PSD, accurate estimation can be conducted with a low number of samples. For SLA signals, this would allow us to estimate the slope on a small ocean area, with a few SLA samples (in the case of Sentinel-3, the distance between two consecutive samples is $319m$ along the satellite track). Another advantage is that estimates of the form $S_x(f;\hat{\boldsymbol{\theta}})$ exhibit less variance than Fourier-based estimates, leading to smoother PSD, which will facilitate a reliable slope estimation in the case of SLA signals.

A very popular and understood PSD parametric model is the AR (AutoRegressive) one, due to the fact that obtaining the AR parameters reduces to solving a linear least-squares problem, for which computationally efficient algorithms have been proposed [6]. Moreover, an AR model is of interest for a large class of signals since it consists in modeling a signal $x(n)$ as a linear combination of its past samples with an additive component representing the unexpected part of the signal

$$x(n) = -\sum_{\ell=1}^{p} a_\ell x(n-\ell) + e(n) \tag{1}$$

where p is the AR model order, a_ℓ is the ℓth AR coefficient and $e(n)$ is the model error (namely the linear prediction error (LPE)). The fitting of such a model to a signal leads to the following spectral estimator of $x(n)$

$$S_{\text{AR}}(f;[a_1 \ldots a_p\ \sigma_e^2]) = \frac{\sigma_e^2}{\left|1+\sum_{\ell=1}^{p} a_\ell e^{-i2\pi\ell f}\right|^2} \tag{2}$$

where σ_e^2 is the LPE power.

In the case of SLA signals, our experience is that AR modeling can work fine, provided that a sufficient AR model order (large number of parameters) is used. However, this is not fully satisfactory since we loose the interest of a model with few parameters and we do not take into account the problem specificities, namely that the spectral part of interest lies in very low frequencies while the rest of the frequency band contains mostly white noise.

2.2. Proposed pre-processing: warping

In order to account for the frequency distribution of the signal power, we use the basic idea of a non-uniform spectral representation, with a view to emphasize the lower part of the spectrum compared to the high frequency part. This idea of using an unequal resolution related to the frequency is an old one, which goes back to the seventies [7]. It has been extensively used for audio applications [8] where it is sometimes referred to as "frequency warping". The basic idea is to obtain a transformed sequence $y(n)$, which corresponds to an expansion over a set of orthogonal sequences $\psi_k(n)$, i.e.,

$$x(n) = \sum_{k=-\infty}^{+\infty} y(k)\psi_k(n) \tag{3}$$

where the functions $\psi_k(.)$ should be chosen so that the Fourier transforms of $x(n)$ and $y(n)$ are related to one another by a function $W(f)$ such that

$$Y\left[W(f)\right] = X(f). \tag{4}$$

Hence a conventional Fourier transform of $y(n)$ over equally spaced frequencies yields a non-equally spaced frequency analysis for $x(n)$. A significant advantage of this technique is that it can be implemented very easily from digital filters. Various choices exist for the sequences $\psi_k(n)$ which result in different non-linear functions $W(f)$. Since they exhibit good properties and are widely used in warping methods, this paper uses Laguerre functions [9], leading to

$$W(f) = f + \frac{1}{\pi}\arctan\left(\frac{b\sin(2\pi f)}{1-b\cos(2\pi f)}\right) \tag{5}$$

where the parameter $b \in [-1,1]$ impacts the shape of the function $W(f)$. In our application, one wishes to dilate low frequencies while compressing high frequencies, leading to the constraint $b > 0$.

2.3. The proposed spectral analysis: ARWARP

For SLA signal analysis, we propose to use frequency warping as a pre-processing step, which enhances the low-frequency components before an AR spectral analysis. A linear regression is finally conducted on the resulting AR estimator allowing the slope α to be estimated, as illustrated in Fig. 2. The warping pre-processing might be combined with any spectral analysis method. However, based on the benefits of parametric methods detailed above, AR modeling has been preferred for SLA analysis. The pair (frequency warping, AR modeling) will be referred to as ARWARP in the sequel.

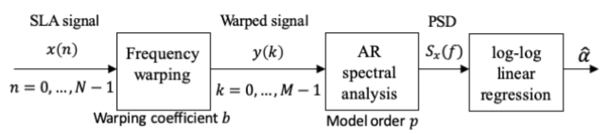

Fig. 2: ARWARP: proposed SLA processing.

ARWARP requires the tuning of three parameters: the warping coefficient b, the number of warped samples M and the AR model order p. The selected values of b and M usually result from some trade-off. More precisely, the warping coefficient b is directly linked to a so-called turning point frequency f_w [8]

$$b = \cos\left(2\pi f_w\right). \tag{6}$$

For $b > 0$, the PSD is sampled with higher resolution at frequencies lower than f_w, and lower resolution at frequencies higher than f_w. For SLA signals, f_w should correspond to the end of the slope region in a log-log scale (also corresponding to the beginning of the noise floor), i.e., $f_w \simeq 0.01$ in normalized frequencies, as observed in Fig. 1, which corresponds to $b = 0.99$.

The input SLA signal being of finite length (N samples), the warped sequence should be of infinite size [7]. However, from a practical point of view, only M samples of the warped sequence are computed. The influence of the warping sequence truncation has been studied in [9]: taking into account the total group propagation time of the Laguerre warping system, the minimum number of warped samples allowing a quasi-reversible transformation is defined by

$$M = N\frac{1+|b|}{1-|b|}. \tag{7}$$

Note that for $N = 3000$ and $b = 0.99$, we obtain $M = 597000$ which induces a high computational cost in the warping step. In order to reduce the value of M, we have chosen $b = 0.9$, corresponding to a value of M more than ten times lower, i.e., $M = 57000$ and a turning point frequency of $f_w = 0.07$, which is acceptable in view of Fig. 1.

Finally, the model order p needs to be adjusted. One might think of using classical AR model order criteria such as Akaike or minimum description length (MDL) [6]. However, these criteria are more adapted to line spectra, which is not the case for SLA signals. The model order p has to be low enough to guarantee a "smooth" spectral behavior. A reasonable choice is $p \in \{5, ..., 9\}$.

Once the ARWARP parameters have been set, one can compute the ARWARP spectral estimator as follows

$$S_x(f) = \frac{\sigma_e^2}{\left|1 + \sum_{\ell=1}^{p} a_\ell e^{-i2\pi\ell W(f)}\right|^2} |\Lambda_0(f)|^2, \quad (8)$$

where $a_\ell, \ell = 1, ..., p$ are the AR coefficients estimated using any linear prediction algorithm applied to the warped sequence $y(k), k = 0, ..., M-1$, σ_e^2 is the LPE power and $\Lambda_0(f)$ is the lowpass filter (in the case of $b > 0$) corresponding to the zero-order Laguerre sequence [9]. Once the PSD (8) has been computed, we propose to estimate the slope α using a classical linear regression on a log-log spectral representation. The full estimation strategy is summarized in Fig. 2.

3. VALIDATION ON SIMULATED SIGNALS

The first validation step is to compare the performance of the proposed ARWARP algorithm to conventional Fourier-based methods on simulated SLA signals with a known slope α, as explained in the next sections.

3.1. Simulation model

A simplified model of SLA PSD is considered

$$S_x(f) = \sigma^2 + S_\alpha(f) = \sigma^2 + \begin{cases} C f_1^{-\alpha} & 0 < f < f_1 \\ C f^{-\alpha} & f \geq f_1 \end{cases} \quad (9)$$

where σ^2 is the white noise power, and where the PSD at very low frequencies is fixed by continuity arguments (note that the SLA PSD is unknown for $f < f_1$, or at least not easy to characterize as f_1 is very low). This model is represented by the red curve in Fig 1.

Simulated SLA signals were generated as Gaussian vectors whose correlation function is computed as the inverse transform of $S_\alpha(f)$ in (9), except for the zero-lag where the noise power σ^2 was added (additive white noise). The corresponding Matlab code for the generation of SLA signals is available in [10].

3.2. Statistical analysis

In this section, we compare the performance of estimators of α obtained after performing linear regression on both a classical periodogram and an ARWARP spectral estimator. Hence, 1000 Monte-Carlo simulations were run for a SLA signal of $N = 3000$ samples with PSD (9), the frequency f_1 corresponding to a distance $d_1 = 319$ km (i.e., a normalized frequency $f_1 = 10^{-3}$) and a white noise level fixed to $\sigma^2 = 0.003$ (to be coherent with Sentinel-3 real data). Fourier-based spectrum estimation was conducted by using a periodogram with a 10% Tukey window applied after detrending and zero-padding by a factor 3. For the ARWARP model, the warping parameter was $b = 0.9$ and different values of the model order were considered, i.e., $p \in \{5, 7, 9\}$. In order to estimate the spectral slope, linear regression was performed on a frequency range corresponding to the two dashed green vertical lines in Fig. 3. Figure 4

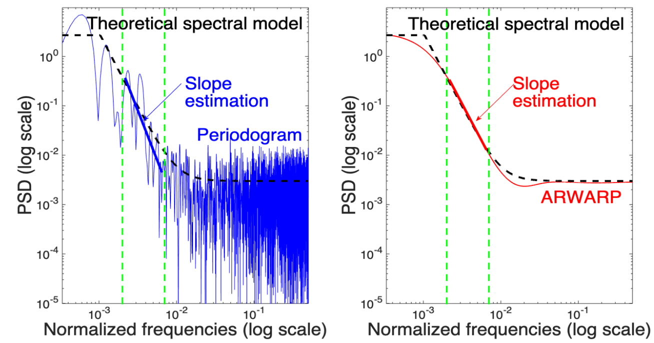

Fig. 3: Slope estimation by linear regression on estimated PSD of a simulated SLA signal ($N = 3000$, $\alpha = 3$). (Left) PSD estimated using a periodogram (detrending, 10% Tukey window, zero-padding by a factor 3). (Right) PSD estimated using the ARWARP method ($p = 5$, $b = 0.9$, $M = 57000$).

displays the boxplots of the bias between the estimated and theoretical values of α versus different values of α, while the corresponding values of the mean-square error (MSE) are reported in Table 1 .

α	2	2.5	3	3.5	4
Periodogram MSE	0.76	0.75	0.68	0.57	0.54
ARWARP(5) MSE	0.20	0.23	0.16	0.12	0.23
ARWARP(7) MSE	0.23	0.21	0.13	0.10	0.27
ARWARP(9) MSE	0.27	0.19	0.15	0.25	0.54

Table 1: MSE of slope estimates after computing a linear regression on estimated PSDs using periodogram and ARWARP models ($N = 3000$).

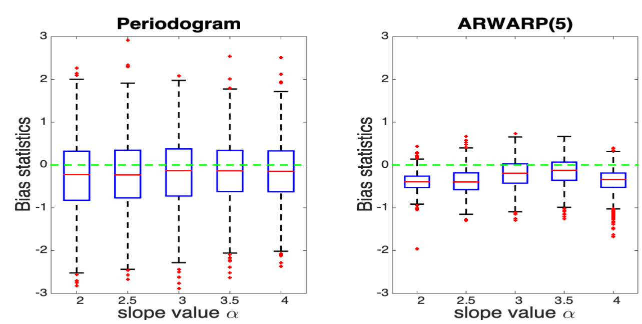

Fig. 4: Statistics (boxplots) of slope bias on 1000 sequences of simulated SLA signals ($N = 3000$): estimation using periodogram (left) and ARWARP ($p = 5$, right). In each boxplot, the central mark is the median, the edges are the 25th and 75th percentiles, the whiskers extend to the most extreme datapoints the algorithm considers to be not outliers, and the outliers are plotted individually in red.

From a spectral point of view, Fig. 3 highlights the interest of the ARWARP model, which provides a spectral estimator well-fitted to SLA signals, allowing a PSD estimation with a small number of samples, i.e., with a high spatial resolution, which is not possible with Fourier-based methods.

Figure 4 shows that the variance of the slope estimates is much smaller when linear regression is applied to the ARWARP model compared to the periodogram. These results show that the estimation of the SLA slope is possible with a high spatial resolution using the ARWARP method, while Fourier-based methods yield less accurate estimates. Table 1 confirms these results quantitatively and shows the impact of the model order p: a good compromise seems to be $p = 5$ for all slope values.

4. VALIDATION ON REAL SIGNALS

4.1. SLA around the Equator

For a second validation step, we apply the proposed ARWARP method on Sentinel-3 real data measured around the Equator in an area where the slope is known to have low values and to be mostly stationary. Figure 5 presents on the left the result of slope estimation on 52 segments of $N = 3000$ samples with periodogram in blue and ARWARP in red. As expected, slope estimation via periodogram gives spurious results with an unacceptable variance along the different segments, while ARWARP estimates yield quite constant values of the slope, as expected from a physical point of view. No ground truth exists for these real data. However, the results associated with the ARWARP model seeem more reliable, with values around $\alpha = 1.45$. Note that a classical Welch (cumulated) periodogram on these 52 segments gives an estimation of the slope $\widehat{\alpha} = 1.51$ (green line). Once again, compared to a Welch periodogram applied on the 52 segments, the ARWARP model allows PSD and slope estimations on each individual segment, as illustrated in the right figure.

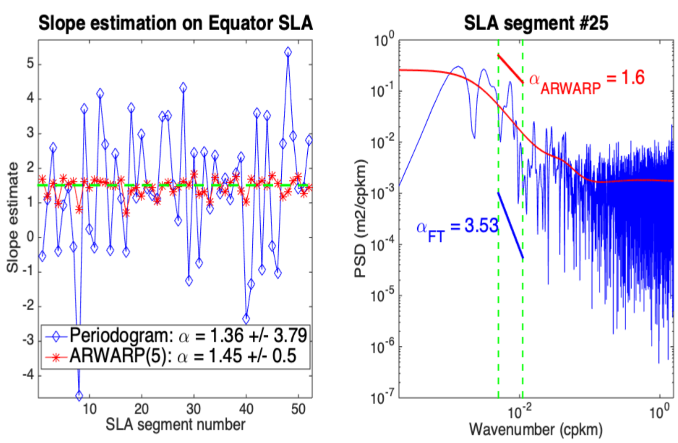

Fig. 5: (Left) slope estimates on segments around the Equator ($N = 3000$), using periodogram (blue), ARWARP ($p = 5$, red) and Welch (cumulated on the 52 segments, green). The mean value of the estimated slopes is also reported with an error interval of +/- twice the standard deviation. (Right) PSD and slope estimates for segment #25.

4.2. SLA in the Agulhas Current

This last section considers SLA segments from the Agulhas Current, which is an area where higher and different slope values are expected. Figure 6 (left) displays slope estimates obtained on different SLA segments ($N = 3000$): the periodogram obviously gives spurious and unreliable results, while ARWARP estimates are more coherent. Figure 6 (right) shows examples of PSD and slope estimations where the periodogram fails to give a good slope estimate, while the ARWARP PSD remains "smooth" guaranteeing a more accurate slope estimation.

5. CONCLUSIONS

This paper proposed a new method to estimate the slope of SLA signals based on a combination of frequency warping, AR modeling and linear regression. This method showed improved results compared to Fourier-based strategies for simulated SLA signals. This improvement has also been observed on real Sentinel-3 data. The proposed ARWARP model makes PSD and slope estimations possible on a

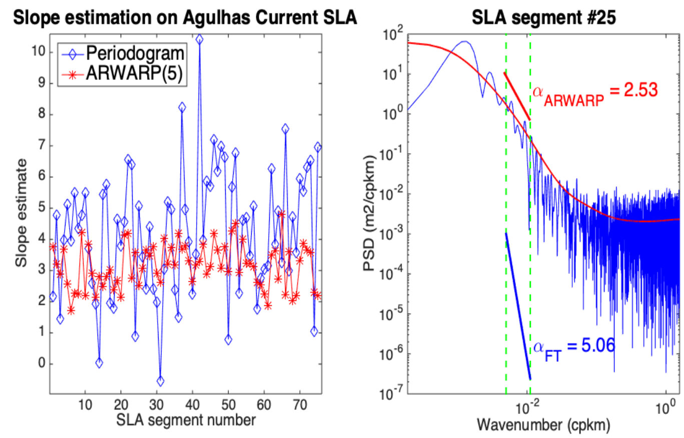

Fig. 6: (Left) slope estimation on segments ($N = 3000$) of SLA in the Agulhas Current area, using periodogram (blue) and ARWARP (order $p = 5$, red). (Right) PSD and slope estimations on segment #25.

short SLA segment, i.e., allows spatial resolution of the estimates to be improved. Future works include validation on more real data, which is encouraged using the Matlab code available in [10].

6. REFERENCES

[1] P. Y. Le Traon, P. Klein, B. L. Hua, and G. Dibarboure, "Do altimeter wavenumber spectra agree with the interior or surface quasigeostrophic theory?," *J. of Phys. Oceanograph.*, vol. 38, pp. 1137–1142, May 2008.

[2] C. Dufau, M. Orsztynowicz, G. Dibarboure, R. Morrow, and P. Y. Le Traon, "Meoscale resolution capability of altimetry: Present and future," *J. of Geophys. Res. Oceans*, vol. 121, pp. 4910–4927, 2016.

[3] J. G. Richman, B. K. Arbic, J. F. Shriver, E. J. Metzger, and A. J. Wallcraft, "Inferring dynamics from the wavenumber spectra of an eddying global ocean model with embedded tides," *J. of Geophys. Res.*, vol. 117, pp. C12012, 2012.

[4] Y. Xu and L. L. Fu, "Global variability of the wavenumber spectrum of oceanic mesocale turbulence," *J. of Phys. Oceanograph.*, vol. 41, pp. 802–809, 2011.

[5] C. B. Rocha, T. K. Chereskin, S. T. Gille, and D. Menemenlis, "Meoscale to submesoscale wavenumber spectra in Drake passage," *J. of Phys. Oceanograph.*, vol. 46, pp. 601–620, 2016.

[6] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*, Pearson Prentice Hall, Upper Saddle River, NJ, 2005.

[7] A. Oppenheim, D. Johnson, and K. Steiglitz, "Computation of spectra with unequal resolution using the Fourier transform," *Proc. of the IEEE*, vol. 59, no. 2, pp. 299–301, Feb. 1971.

[8] A. Härmä, M. Karjalainen, L. Savioja, V. Välimäki, U. K. Laine, and J. Huopaniemi, "Frequency-warped signal processing for audio applications," *J. Audio Engineering Society*, vol. 48, no. 11, pp. 1011–1031, 2000.

[9] G. Evangelista and S. Cavaliere, "Frequency-warped filter banks and wavelet transforms: a discrete-time approach via Laguerre expansion," *IEEE Tr. on Signal Process.*, vol. 46, no. 10, pp. 2638–2650, Oct. 1998.

[10] C. Mailhes and O. Besson, *Matlab code for SLA analysis*, Jan. 2020, https://www.tesa.prd.fr/software-tesa.p56.html.