Relation between deep bioluminescence and oceanographic variables: A statistical analysis using time–frequency decompositions

S. Martini, David Nerini, Christian Tamburini

Relation between deep bioluminescence and oceanographic variables: a statistical analysis using time-frequency decompositions.

Martini S.[a,*], Nerini D.[a,*], Tamburini C.[a,]

[a] *Aix Marseille Université, CNRS, Université de Toulon, IRD, MIO UM 110, 13288, Marseille, France.*

Abstract

We consider the statistical analysis of a 1.7-year high-frequency sampled time series, between 2009 and 2010, recorded at the ANTARES observatory in the deep NW Mediterranean Sea (2,475 m depth). The objective was to estimate relationships between bioluminescence and environmental time series (temperature, salinity and current speed). As this entire dataset is characterized by non-linearity and non-stationarity, two time-frequency decomposition methods (wavelet and Hilbert-Huang) were used. These mathematical methods are dedicated to the analysis of a signal at various time and frequencies scales. This work propose some statistical tools dedicated to the study of relationships between two time series. Our study highlights three events of high bioluminescence activity in March 2009, December 2009 and March 2010. We demonstrate that the two events occurring in March 2009 and 2010 are correlated to the arrival of newly formed deep water masses at frequencies of approximately 4.8×10^{-7} (period of 24.1 days). In contrast, the event in December 2009 is only correlated with cur-

∗. Corresponding author
Email addresses: `severine.martini@club-internet.fr` (Martini S.), `david.nerini@univ-amu.fr` (Nerini D.)

rent speed at frequencies of approximately 1.9×10^{-6} (period of **6.0 days**). The use of both wavelet and Hilbert-Huang transformations has proven to be successful for the analysis of multivariate time series. These methods are well-suited in a context of the increasing number of long time series recorded in oceanography.

1. Introduction

The sampling and the understanding of complex environmental systems aims to detect potential disturbances as a shift from the intrinsic variability of these systems. Marine systems are variable at all time and space scales (Hewitt et al., 2007), and their variability is still poorly understood because of sampling strategy, instrumentation and spatio-temporal heterogeneity challenges. In response to this lack of knowledge, there is a global effort to capitalize oceanographic data and costs for autonomous Eulerian or mobile infrastructures that enable detection of long-term environmental trends as well as special events or perturbations (Favali and Beranzoli, 2006). As an example, OceanSITES, and more generally the Global Ocean Observing System (GOOS), are the major science teams integrating a global network of more than 60 *in situ* observatories acquiring long-term and high-frequency time series over the world (Send and Lankhorst, 2011).

The ANTARES collaboration (Astronomy with a Neutrino Telescope and Abyss environmental RESearch) developed such an infrastructure, in the North-Western Mediterranean Sea (Figure 1), at the end of 2007. This deep-sea-cabled observatory is part of global data networks, such as EMSO, KM3NeT, ESONET and EuroSITES. Initially, the ANTARES observatory was dedicated to the search for high-energy particles, such as neutrinos (Amram et al., 2000; Aguilar et al., 2007; Ageron et al., 2011). About 885 photomultiplier tubes (PMTs) are installed between 2,000 and 2,475 m depth for the purpose of particle physics. All the 12 ANTARES mooring lines are connected, via an electro-optical cable, to a shore station providing real-time acquisition. With the installation of a specific line, namely IL07 (Tamburini et al., 2013), this deep observatory gives the opportunity to record simultaneously

high-frequency oceanographic data such as current speed, salinity and temperature (Figure 2). The IL07 is also equipped with PMTs devoted to the recording of bioluminescence activity. These datasets provide an exceptional way to study the dynamics of the deep ecosystem in real-time and at high-frequency (Craig et al., 2009).

Recently, Tamburini et al. (2013) proposed a descriptive analysis of the ANTARES oceanographic time series gathering both physical and biological variables, with synchronous hydrological records from a surrounding station located in the Gulf of Lion. They highlighted a link between high bioluminescence activity and changes in the properties of deep water temperature and salinity. Such changes attributed to an open-sea convection event renew the deep waters (so-called newly formed deep water) by the fall of upper ocean layer through the water column (Tamburini et al., 2013; Marshall and Schott, 1999; Stabholz et al., 2013; Béthoux et al., 2002). Open-sea convections represent a major vector in fueling the deep-sea ecosystem with nutrients, carbon, oxygen and, potentially, organisms. This might induce higher bioluminescence emission. Indeed, the input of organic matter into the deep water has the potential to fuel the deepsea biological activity. Within this biological stimulation context, bioluminescent bacteria, which are not affected by mechanical stimulation and are able to glow continuously under specific conditions, might be affected by changes in environmental growth conditions. These organisms are potential contributors to high bioluminescence intensity. Moreover, bioluminescent organisms could also be advected from the surface to the deep-sea.

The present study attempts to understand the mechanisms inducing deep-sea bioluminescence activity using temperature, salinity, current speed and bioluminescence time series

4

from the same ANTARES dataset of Tamburini et al. (2013). The search for characteristic scales in time and frequency in this high-frequency sampled dataset is achieved by using statistical methods from signal processing. Whereas most environmental data are non-stationary (Cazelles et al., 2008; Rao and Hsu, 2008; Ghorbani, 2013), most methods used in time-series analyses are based on stationarity assumptions, such as those used in the Fourier decomposition (Frazier, 1999). First, we propose the use of two complementary mathematical methods (Wavelet and Hilbert-Huang transformation) to deal with non-stationarity and to decompose each time series within time and frequency space. Second, we propose some tools to quantify and to illustrate links between bioluminescence time-frequency decomposition and other environmental variable time-frequency decompositions.

2. Dataset for the deep bioluminescence study

The ANTARES site is located 40 km off the French Mediterranean coast (42°48'N, 6°10'E) at a 2,475 m depth (see Figure 1). This work focuses on multivariate time series sampled from the beginning of 2009 until October 2010 for oceanographic variables such as salinity, potential temperature (°C) and current speed ($cm\ s^{-1}$) (Figure 2). Moreover, bioluminescence emission (kHz) was recorded in 2009 and 2010 from the IL07 instrumented line of the ANTARES telescope. The unit of bioluminescence activity (kHz) refers to the photon counting rate per second. We consider the bioluminescence time series as a response variable submitted to changes in environmental characteristics.

In Figure 2, each environmental variable is recorded with a high-frequency sampling rate, involving the presence of a high level of noise. The base-line of each signal is close to a
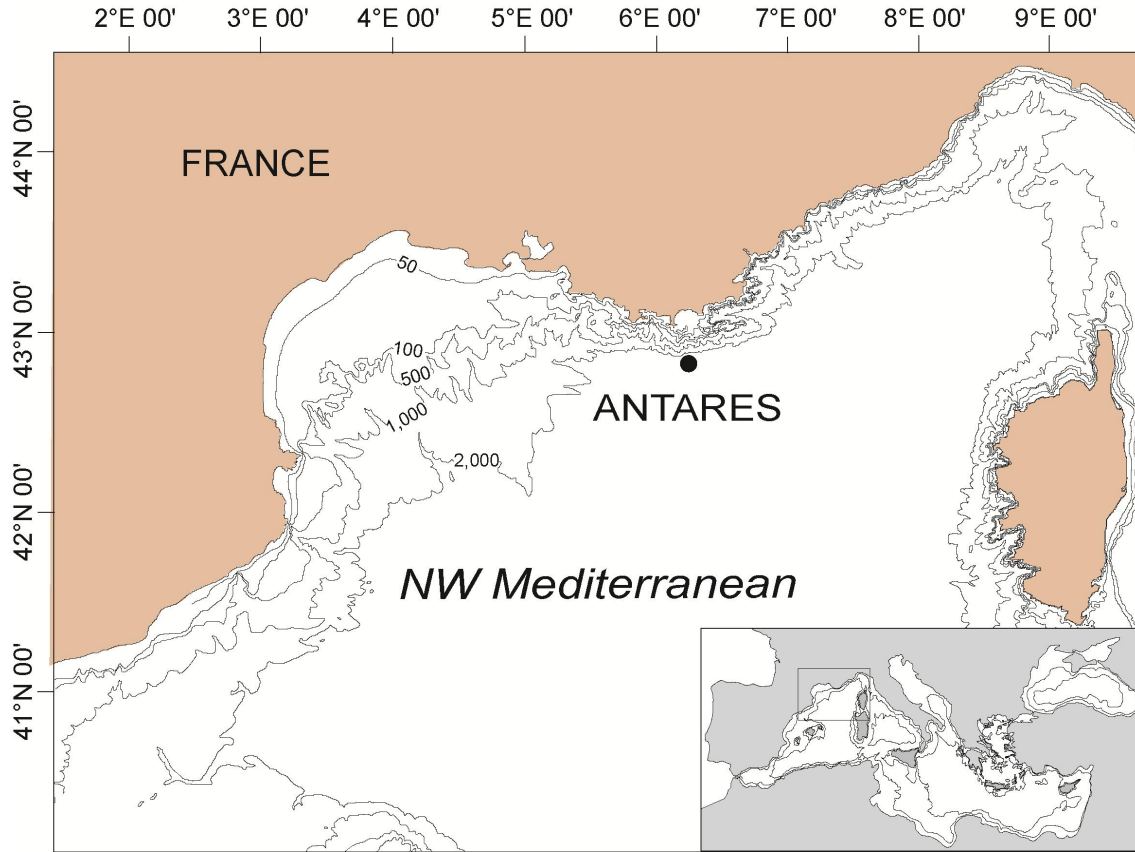
5

FIGURE 1: Map of the North-Western Mediterranean Sea and the ANTARES site (black dot) where the underwater neutrino telescope is immersed at 2,475 m depth. High-frequency sampled time series of bioluminescence, temperature, salinity and current speed were sampled between 2009 and 2010 at this station.

constant background signal and does not display a clear trend when looking at low-frequency variations. However, some breaking sequences appear on each time series, giving their non-linear and intermittent appearance. There is no clear time connection between those sequences from one time series to another. However, they are all characterized by a sudden increase in variability with two main events observed simultaneously occurring in March 2009 and from March to June 2010. If the connection between variables is relatively clear for the event in 2010, it is more difficult to make conclusions for the event in 2009 without a thorough and appropriate analysis.

As mentioned previously, the time series are not stationary. A time series $X(t), t \in \mathbb{R}^+$, is
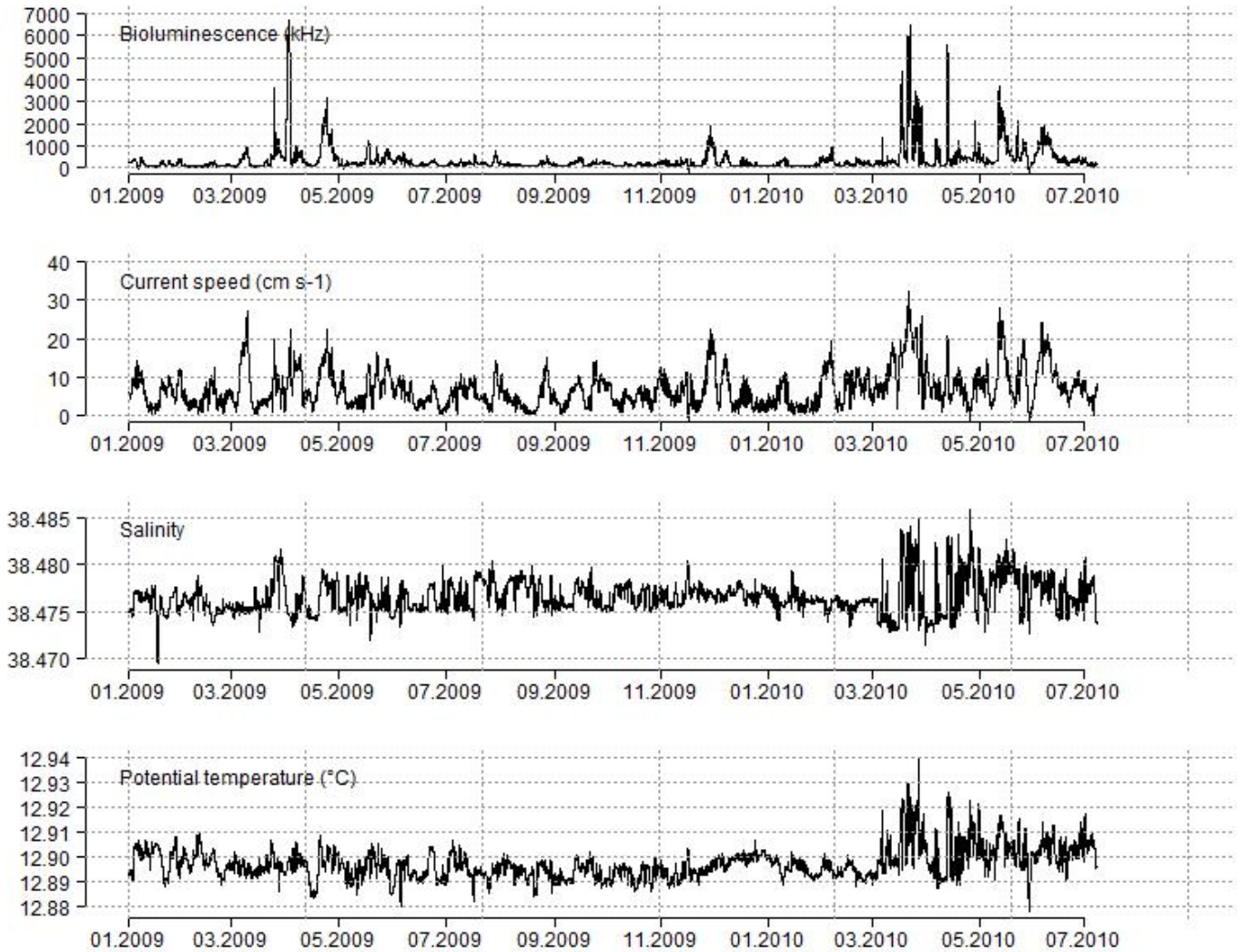
FIGURE 2: Time series recorded in the Mediterranean Sea, at the ANTARES station, between January 2009 and July 2010, for bioluminescence, current speed ($cm\ s^{-1}$), salinity, and temperature (°C). Those high-frequency data, recorded each 1 to 15 minutes, are sampled at a depth between 2,478 m and 2,169 m and uploaded in real-time to the database.

said to be weakly stationary if $X(t)$ owns the same statistical properties (mean and variance) as the shifted time series X_{t+h} for $h > 0$. A Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test, from Kwiatkowski et al. (1992), shows that the data are away from the null-hypothesis of stationarity and that all of these time series are non-stationary (Table 1).

TABLE 1: KPSS-test for testing stationarity of each time series.

Variable	KPPS value	p-value	Interpretation
Bioluminescence (kHz)	8.24	0.01	non-stationary
Current speed $(cm\ s^{-1})$	5.63	0.01	non-stationary
Temperature $(\,^{\circ}\mathrm{C})$	8.50	0.01	non-stationary
Salinity	2.29	0.01	non-stationary

3. Methods dealing with non-stationary time series

Most methods in time series analysis consist of an expansion of the studied signal $X(t)$ into a linear combination of known basis functions. For the Fourier decomposition, the signal is expressed as a linear combination of trigonometric functions:

$$X(t) \ = \ \sum_{j=0}^{\infty} a_j e^{i\omega_j t}. \tag{1}$$

This decomposition provides an analytical expression of the function $X(t)$ with amplitude coefficients a_j giving weight to angular frequencies ω_j. Both coefficients a_j and ω_j are independent of time, providing a global decomposition of the signal in the frequency space.

The Hilbert-Huang transformation (HHT) proposes a generalization of the Fourier decomposition by expanding $X(t)$ such that:

$$X(t) \ = \ \sum_{j=1}^{n} a_j(t) e^{i2\pi \int \omega_j(\tau) d\tau} + R_{n+1}(t) \tag{2}$$

where both amplitude coefficients $a_j(t)$ and angular frequency coefficients $\omega_j(t)$ are functions of time. As further developed, this decomposition relies on a preliminary step to decompose the time series into $n+1$ modes, where $R_{n+1}(t)$ is the trend of the time series.

Another decomposition with time-dependent parameters can be constructed using the continuous wavelet decomposition

$$X(t) = \frac{1}{W_\psi} \int_{-\infty}^{\infty} \int_{0}^{\infty} W_X(a,b)\psi_{a,b}(t)\frac{1}{a^2}dadb \tag{3}$$

In that case, the shape of the basis function, $\psi_{a,b}(t)$, is controlled by a scale parameter a, which can be interpreted as a reciprocal of frequency and a shifting parameter b, both parameters varying continuously over $\mathbb{R}^+$ and $\mathbb{R}$ respectively. The analysis is conducted through the estimation of the bivariate function $W_X(a,b)$, the wavelet transform, whose values are scale and time dependent and on the wavelet function $\psi_{a,b}(t)$. The constant W_ψ is a normalization value depending on the chosen wavelet function.

Once the basis expansion is achieved, each method provides the ability to track, over time, the evolution of the energy distribution.

3.1. Wavelet decomposition

3.1.1. Decomposition into wavelet basis

The continuous wavelet decomposition is a consecutive pass-band filter through time series (Torrence and Compo, 1998; Addison, 2010). The decomposition acts as a linear filter that extracts special features inside the signal through a projection over a wavelet basis $\psi_{a,b}$ giving the coefficient

$$W_X(a,b) = \int_{-\infty}^{+\infty} X(t)\psi_{a,b}^*(t;a,b)dt \tag{4}$$

where $\psi_{a,b}^*$ denotes the complex conjugate of the basis wavelet. The peculiarity of that decomposition is the continuous dependency of the basis functions

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi(\frac{t-b}{a}) \tag{5}$$

regarding scale parameter a and location parameter b. The scale parameter a is inversely proportional to the frequency. The decomposition mainly relies on the choice of a mother basis function. Torrence and Compo (1998) defines some elements to choose the mother wavelet: orthogonality, real or complex function and shape of the mother wavelet related to shape of analyzed time series. In this study, we choose the Morlet wavelet function, commonly used in wavelet analysis. Though other choices can be relevant as well (Ahuja et al., 2005; Cazelles et al., 2008), the wave shape Morlet wavelet fits well with the time series shape observed in this study, which is not the case with Daubechies or Haar wavelets. The Morlet wavelet is defined as:

$$\psi(t) = \frac{1}{\pi^{1/4}}e^{i2\pi f_0 t}e^{-t^2/2} \tag{6}$$

With $2\pi f_0$ being the angular frequency. This periodic wavelet is defined following $\int \psi(t)dt = 0$ and a unit norm $\int |\psi(t)|^2 dt = 1$. The time series is decomposed into coefficients for the real part of the wavelet transform $W_X(a,b)$ at the position b. These coefficients are calculated by continuously moving the wavelet along the signal (modifying the position parameter b) related to a specific scale parameter a. Then, a bi-dimensional surface of $W_X(a,b)$ is constructed and can be displayed as a time-frequency plot.

3.2. Hilbert-Huang transformation (HHT)

Huang et al. (1998) proposed a different method for non-stationary and non-linear signal analysis (Huang et al., 1998; Huang and Wu, 2008; Huang et al., 2009a). The Hilbert-Huang transformation is based on the assumption that a signal $X(t)$ has multi-component meaning and that it can be expanded into a sum of $n+1$ signals such that

$$X(t) = \sum_{j=1}^{n} C_j(t) + R_{n+1}(t), \tag{7}$$

where the C_j are called intrinsic mode functions (IMFs), and R_{n+1} is the trend of $X(t)$. Each IMF is supposed to be almost mono-component *i.e.* composed of only one instantaneous frequency evolving with time in a bounded range. The HHT is then achieved in two main steps: the extraction of IMFs and their decomposition into instantaneous frequencies.

3.2.1. Empirical mode decomposition

To construct the mono-component signals, the n IMFs must fulfill two conditions:

(i) the number of times passing through the origin is identical to the number of extrema or differ by one for the most

(ii) in any time, the mean between the local maximum and the local minimum is zero.

Following these conditions, the decomposition of $X(t)$ into IMFs is based on the following steps:

Step 1 - Identification of the extrema (minimum and maximum) of the signal $X(t)$.

Step 2 - Connection of these extrema by natural cubic spline interpolation for the construction of a upper envelope $e_{max}(t)$. The same process is used for the lower envelope $e_{min}(t)$.

Step 3 - Computation of the average of the two envelopes: $m(t) = [e_{min}(t) + e_{max}(t)]/2$, as

shown in Figure 3.

Step 4 - Computation of the locally centered time series with $d(t) = X(t) - m(t)$.

Step 5 - Steps 1-4 are repeated on $d(t)$ until convergence toward a time series with zero mean $m(t)$ is sufficient. The convergence is controlled by a sifting criterion.

Step 6 - If $d(t)$ has a zero mean, then the properties (i) and (ii) are checked, and the function constitutes an IMF: $C(t)$.

Step 7 - Construction of a new $X(t)$ by subtracting $C(t)$ to previous $X(t)$. Repeat step 1 to step 6. Extract IMFs until rule (i) is broken.

Step 8 - The last remainder constitutes the trend $R_{n+1}(t)$ which is, by construction, a monotonic or a constant function.
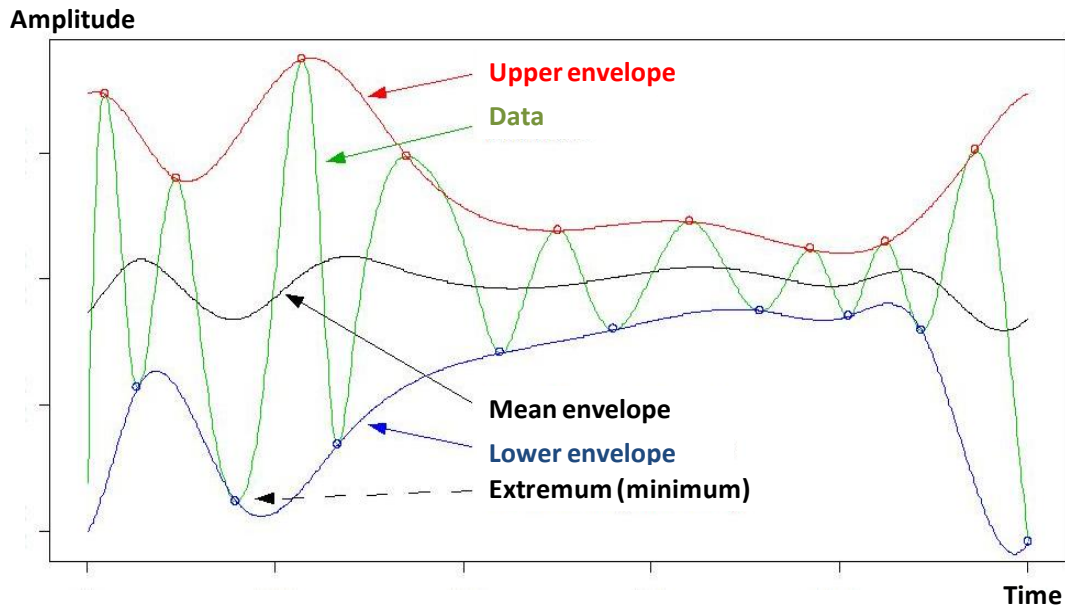

FIGURE 3: Schematic representation of the signal decomposition using the Hilbert-Huang transformation. The upper and lower envelopes are represented by solid lines enclosing the signal. The mean is the centered line, and the empty circles represent the extrema.

3.2.2. Instantaneous frequency using the Hilbert transform

The original signal $X(t)$ is now expressed as a linear combination of IMFs which are supposed to be mono-component. Considering an IMF $C(t)$, the function

$$Z(t) = C(t) + iY(t), \tag{8}$$

is constructed, where $Y(t)$ is the Hilbert transform of the IMF $C(t)$ with:

$$Y(t) = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{C(t')}{t - t'} dt', \tag{9}$$

with P the Cauchy principal value (Paget and Elliott, 1972). The construction of $Z(t)$ makes the decomposition of the IMF possible such that:

$$Z(t) = a(t)e^{i\theta(t)}, \tag{10}$$

where $a(t)$ is the instantaneous amplitude and $\theta(t)$ the instantaneous phase. The parameter $a(t)$ reflects how the energy of the signal varies with time and is given with:

$$a(t) = [C(t)^2 + Y(t)^2]^{1/2}. \tag{11}$$

The instantaneous phase

$$\theta(t) = \arctan\left(\frac{Y(t)}{C(t)}\right) \tag{12}$$

yields the instantaneous frequency $w(t)$ such that:

$$w(t) = \frac{1}{2\pi} \frac{d\theta(t)}{dt}. \tag{13}$$

Applying the instantaneous frequency decomposition for each IMF C_j forms the HHT of the signal $X(t)$. It can be expressed as a linear combination of trigonometric functions with amplitude coefficients $a_j(t)$ and frequency coefficients $w_j(t)$ depending on time:

$$X(t) = \sum_{j=1}^{n} a_j(t)e^{i2\pi \int \omega_j(\tau)d\tau} + R_{n+1}(t) \tag{14}$$

13

3.2.3. Frequency decomposition

In the literature, Rao and Hsu (2008) and Huang et al. (2009a) give an empirical maximal number of IMFs between 6 and 12, depending on the complexity of time series. The IMF decomposition acts as a pass-band filter by separating frequency bands (Flandrin et al., 2004). A way to understand how the empirical decomposition takes place for the HHT method is to study the frequency narrow band for each extracted IMF. A linear relationship, using the least square method, between the logarithm of the IMF average frequencies over the number of IMF is defined empirically following an exponential law such as

$$log(\omega_m(n)) \approx -n \ log(\alpha) \tag{15}$$

with $w_m(n)$ the mean frequency for the IMF n (Huang et al., 2009b). For this log-representation, the relation is linear with a slope coefficient α, and the R^2 coefficient determined to characterize the quality of the fit. The exponential relation with the α coefficient is observed empirically in the literature (Flandrin et al., 2004; Huang and Wu, 2008), but no mathematical assumptions were made at the beginning of the decomposition method (Huang et al., 2009a; Massei and Fournier, 2012). This observation differs from the wavelet decomposition for which a dyadic assumption is made using the set of scales b selected as a functional power of 2:

$$b_i = b_0 \times 2^{i\delta}, i = 0, ..., M \tag{16}$$

with $M = \frac{1}{\delta}log_2\frac{(N\Delta t)}{b_0}$, with N the number of values in the time series, Δt the sampling time-step and δ a scale factor (small values for δ will give a finer definition), b_0 the initial parameter b and b_i the parameter at time i.

3.2.4. Thin-plate spline smoothing

The evolution of instantaneous frequencies and amplitude over time can be represented as a time-frequency amplitude plot. However, a HHT spectrum is discrete and cannot be cross-correlated between two variables. Indeed, the IMFs do not form a common basis where the time series can be expanded. The number of extracted IMFs is different from one time series to another, depending on the properties of the time series themselves. One method to cross time-frequency decompositions is smoothing spectra over both time and frequency to obtain a continuous representation. This is performed with thin plate regression splines (Bookstein, 1989; Mardia et al., 1996; Wood, 2003).

We dispose of n observations $(y_i, \mathbf{x}_i)$ where $\mathbf{x}_i = (\omega_i, t_i) \in \mathbb{R}^2$ is a position in the time-frequency space and y_i, the computed value of the discrete Hilbert spectrum from the IMFs decomposition. We want to estimate the regression surface $H(\mathbf{x})$ as a continuous spectrum such that:

$$y_i = H(\mathbf{x}_i) + \varepsilon_i, \tag{17}$$

where ε_i is a random error term. Thin plate splines can be used to estimate H by finding the function g minimizing the penalized sum of squares

$$\sum_{i=1}^{n} (y_i - g(ni))^2 + \lambda J(g), \tag{18}$$

where $J(g)$ is a penalty function

$$J(g) = \iint_{\mathbb{R}^2} \left(\frac{\partial^2 g}{\partial \omega^2} \right)^2 + 2 \left(\frac{\partial^2 g}{\partial \omega \partial t} \right)^2 + \left(\frac{\partial^2 g}{\partial t^2} \right)^2 d\omega dt, \tag{19}$$

giving a measure of roughness for g and λ, a parameter which controls the trade-off between the fit to data and the smoothness of g.

The value of λ should be determined by cross-validation. The solution of the spline regression problem provides a continuous Hilbert spectrum $H(w,t)$ for a given variable. The Hilbert spectrum gives a measure of amplitude contributing to each frequency, and one can calculate the associated marginal spectrum $h(\omega)$ such that:

$$h(\omega) = \int_0^T H(\omega, t)dt \tag{20}$$

Both wavelet and Hilbert-Huang transformation methods result in a tri-dimensional continuous representation in the time-frequency domain. To quantify the relationships between two series, the following step is to cross spectra by pairs.

4. Bivariate time-frequency cross-analysis

4.1. Wavelet coherence

One of the key issue for the analysis of the ANTARES time series is to compare time-frequency decomposition spectra between variables. This point is well developed for the wavelet method using the cross-wavelet representation, which includes both coherence between coefficients of two different variables and phase delays (Torrence and Compo, 1998; Addison, 2010). We consider two time series X(t) and Y(t) and their wavelet transforms $W_X(a,b)$ and $W_Y(a,b)$, respectively. The coherence is a measure of the squared correlation between wavelets functions of time series for time and frequencies instantaneously. This value allows distinguishing the higher coefficients for both crossed variables, $W_X(a,b)$ and $W_Y(a,b)$ (Torrence and Webster, 1999; Grinsted et al., 2004). The coherency corresponds to the square of the crossed and normalized spectra:

$$C^2(a,b) = \frac{S(W_{XY}(a,b))}{S(|\ W_X(a,b)\ |^2)S(|\ W_Y(a,b))\ |^2}, \forall 0 \leq C^2(a,b) \leq 1. \tag{21}$$

S is a smoothing operator for both the time and scale of the wavelets, and $W_{XY}(a,b)$ the cross-wavelet defined from:

$$W_{XY}(a,b) = W_X(a,b)W_Y^*(a,b). \tag{22}$$

5. Results and discussion for non-stationary time series

Fourier analysis relies on a global expansion of the signal, which can induce spurious harmonic components that cause energy spreading. This misinterpretation especially occurs when dealing with non-linear and non-stationary datasets as shown in Figure 4 A (see introduction of Huang et al. (1998) for more details) compared with the stationary signal decomposed in Figure 4 B. Unlike Fourier decomposition, the continuous wavelet transform and the HHT possess the ability to construct a time-frequency representation as the coefficients of the decomposition are locally dependent on time, which is well adapted to the analysis of non-stationary signals.

5.1. Wavelet decomposition

5.1.1. Wavelet spectral representation

Using the wavelet decomposition, the time-frequency spectrum is plotted for bioluminescence time series (Appendices). For the bioluminescence spectrum (Appendix, Figure A.1 A), three distinct periods (in April-May 2009, December 2009 and March-June 2010) appear with
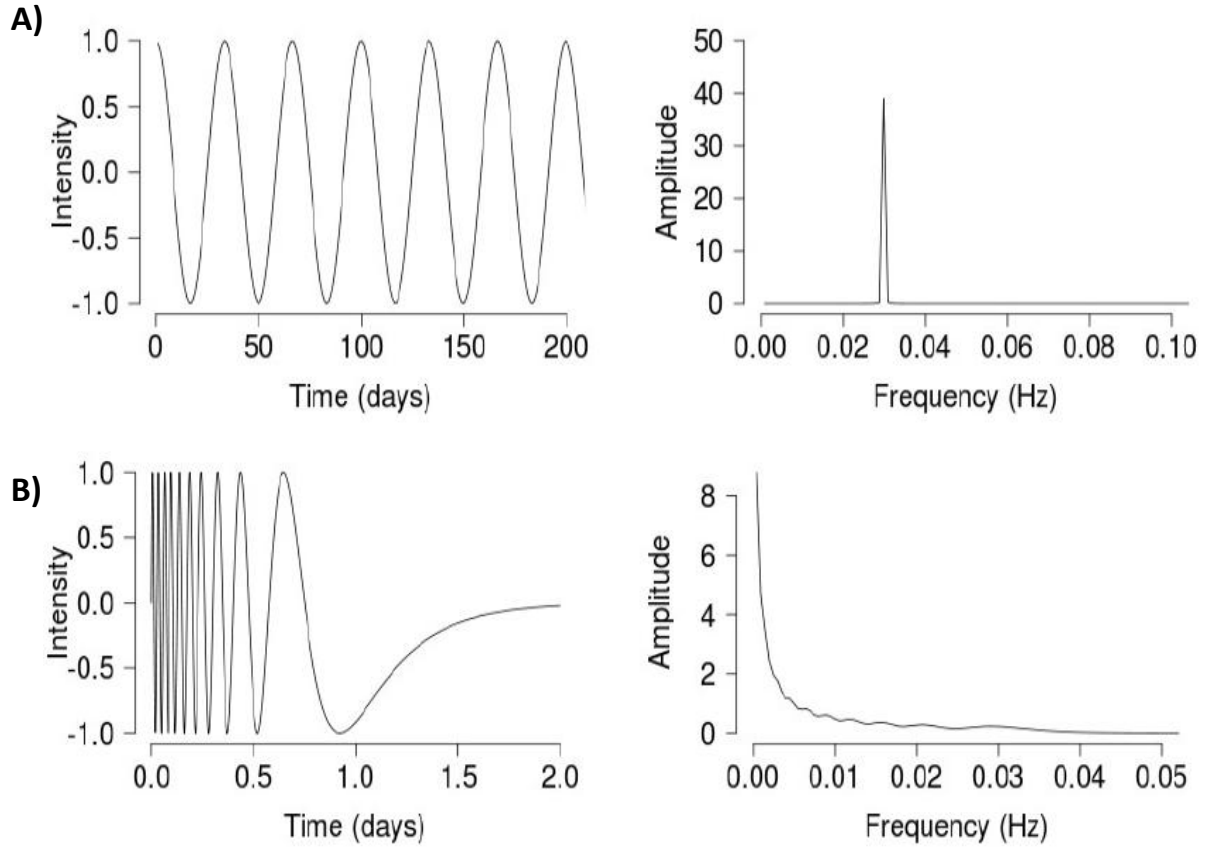
17

FIGURE 4: Two examples of signal decomposition. On the left is the signal represented over time, and on the right is the periodogram over frequencies; A) stationary signal and B) non-stationary signal. The non-stationarity of a signal spreads the energy over frequencies in the periodogram representation.

high (in red) $W_X(a, b)$ wavelet coefficients; see equation (3). The two main periods in April 2009 and April 2010, appear with low excited frequencies, between $4.8 \times 10^{-7} Hz$ (period of 24.1 days) and $3.8 \times 10^{-6} Hz$ (period of 3.0 days). A third one in December 2009 is observed at approximately $1.9 \times 10^{-6} Hz$ (period of 6.0 days). Moreover, frequencies at approximately $2.0 \times 10^{-5} Hz$ (period of 0.6 days or 14 h) are also excited with a high-amplitude coefficient.

For the current speed spectrum (Appendix, Figure A.1 B), lower frequencies at approximately $4.8 \times 10^{-7} Hz$ (period of 24.1 days) appear with high coefficients $W_X(a, b)$ for the whole period. Moreover, a very distinct frequency band is highlighted at $2.0 \times 10^{-5} Hz$ (per-

iod of 0.6 days or 14 h). This last range of excited frequencies is known to be linked to the "internal waves" (Huthnance, 1995). These periodic oscillations generated from the surface and transmitted to the deep-sea have already been observed at periods about 17.6 h at the ANTARES station (Van Haren, 2011). The spectrum shows that if these waves are present during the whole sampled period, their frequency and amplitude coefficients vary over time. As an example, there are higher coefficients in January 2010 than in September 2009.

For salinity (Appendix, Figure A.1 C), two main events are distinct, with very high $W_X(a, b)$ coefficients, in April 2009 and April 2010 with excited frequencies at approximately 9.5×10^{-7} (period of 12.2 days) and 4.8×10^{-7} (period of 24.1 days), respectively.

5.2. Hilbert-Huang transformation

5.2.1. IMFs on ANTARES time series

All of the oceanographic variables recorded at the ANTARES site have been decomposed into IMFs using the HHT. The empirical mode decomposition (EMD) extracts 10 IMFs and a trend for each time series. Figure 5 represents the IMFs for the bioluminescence time series with associated variability intervals. Variability intervals are calculated by generating one hundred decompositions of each time series, varying the sifting criterion to stop the process at more or less 10%.

For the first IMFs (C1 to C5), the variability intervals are very low and cannot be differentiated from the main IMF. They become more important for the last IMFs with, however, similar oscillations over time. By construction, the IMFs are first-order stationary. The first extracted IMFs are high-pass filter, and the last IMFs are low-pass filter. The two biolu-

minescence events in 2009 and 2010 appeared on functions of higher frequencies (Figure 5 from C1 to C7). Once the linear decomposition is achieved, the next step of the Hilbert-Huang transformation is to extract time-dependent frequencies (instantaneous frequencies) from each IMF.

The mean frequency and the standard deviation for the instantaneous frequencies calculated for each IMF from the time series of salinity, bioluminescence, temperature and current speed are represented in Figure 6. The α coefficients from equation (15) estimated for each of the variables are between 1.93 and 1.96, meaning that the mean scale of each mode series is between 1.93 and 1.96 the mean scale of the previous one. All coefficients are significantly different from 2 (not shown), but these values close to 2 indicate a quasi-dyadic filter bank property of the EMD algorithm. This has been found in several other situations (Flandrin et al., 2004; Huang and Wu, 2008; Huang and Schmitt, 2014).

5.2.2. HHT spectral representation

Figures 7 (from A to C) are the continuous spectral representations after smoothing the HHT spectra using thin plate splines for the bioluminescence (Figure 7 A), the current speed (Figure 7 B), and the salinity (Figure 7 C).

For bioluminescence, in Figure 7 A, three events are distinct. Both of them occur from March to May 2009 and from March to June 2010 with high amplitude coefficients and at frequencies between 4.8×10^{-7} (period of 24.1 days) and $1.0 \times 10^{-5} Hz$ (period of 1.2 days). The third event in December 2009 has lower amplitude coefficients and is displayed at 1.9×10^{-6} (period of 6.0 days). The frequency of 2.0×10^{-5} (period of 0.6 days or 14 h) is not excited as it is on Appendix Figure A.1 A.
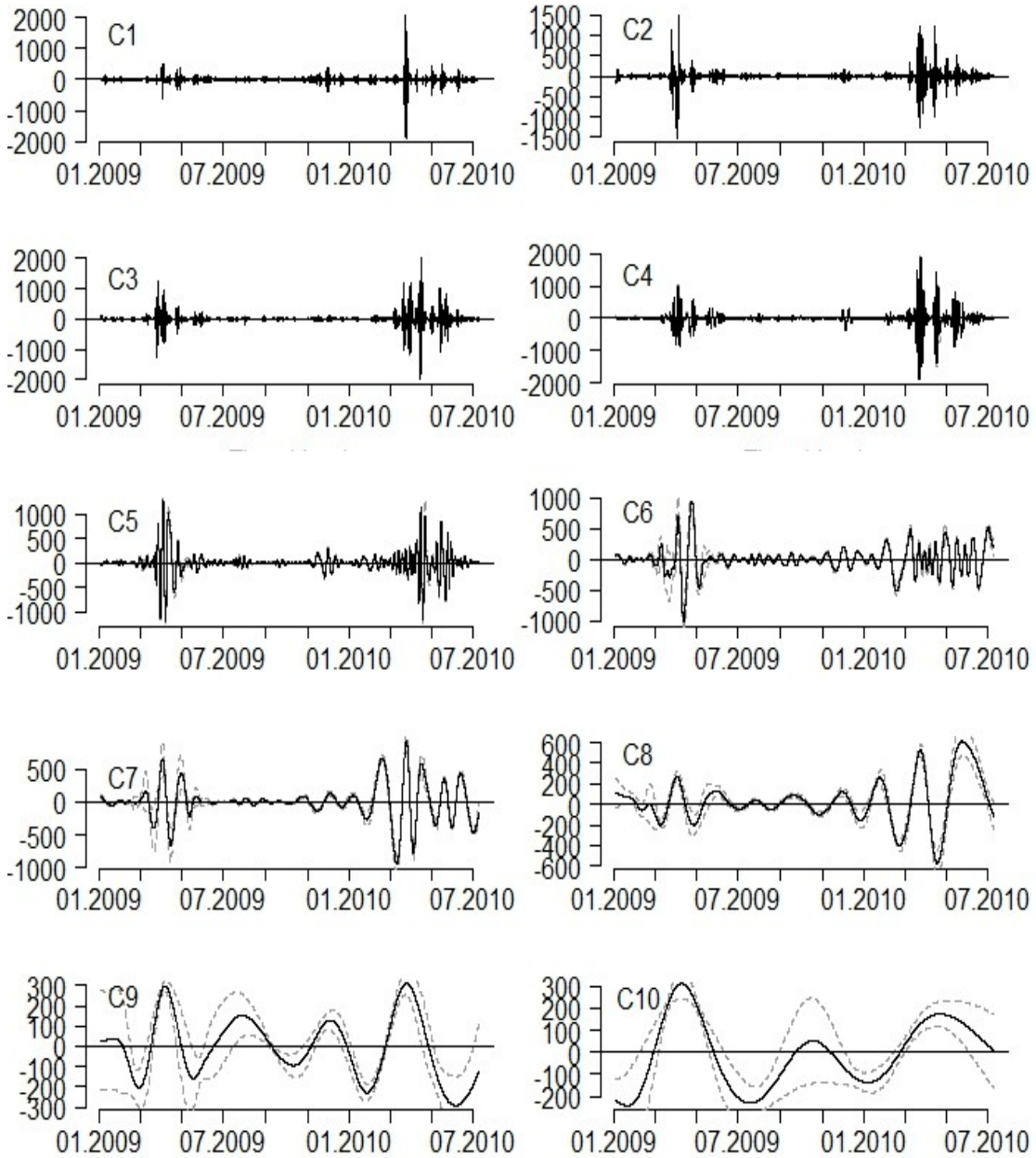
FIGURE 5: Hilbert-Huang transformation for the bioluminescence signal into 10 intrinsic mode functions, from C1 to C10. These intrinsic mode functions are stationary for the first order. In gray, the variability intervals are plotted by moving the threshold defined to stop the sifting process. The trend of the data is not plotted.

For current speed, in Figure 7 B, a large band of frequencies is excited between 4.8×10^{-7} (period of 24.1 days) and $1.0 \times 10^{-5} Hz$ (period of 1.2 days). Another band of excited frequen-
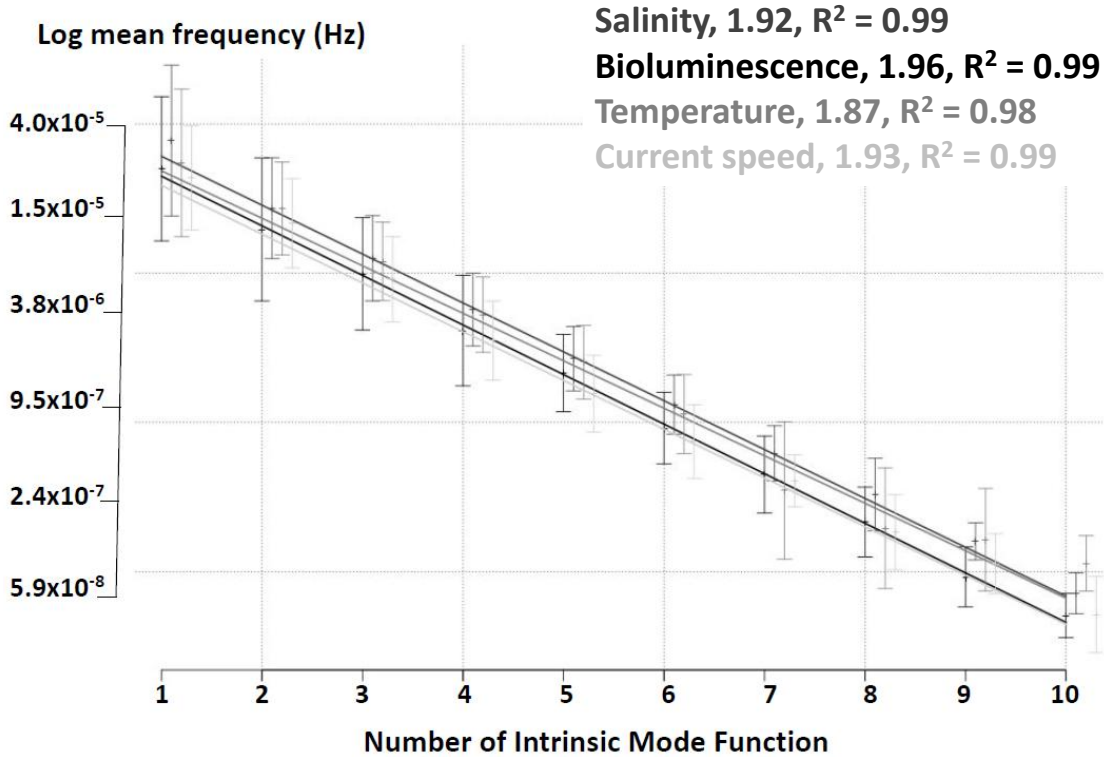
FIGURE 6: Representation of the empirical dyadic decomposition of the HHT. The linear relation is fitted for the logarithm of the mean frequency of each intrinsic mode function, for each variable. Error bars are added for each intrinsic mode function. The slope coefficient is close to 2 but significantly different (not shown) from a dyadic decomposition for all of the time series.

cies is defined at $2.0 \times 10^{-5} Hz$ (period of 0.6 days or 14 h, dotted box). These frequencies are linked to internal waves previously observed and described in Appendix, Figure A.1 B. The internal waves are mainly embedded in the two first IMF in Figure 6.

For salinity, in Figure 7 C, the two main events clearly appear in April 2009 and April 2010 at approximately 9.5×10^{-7} (period of 12.2 days) and at approximately 3.8×10^{-6} (period of 3.0 days).

These spectra provide similar results that corroborate those obtained using the wavelet decomposition in Appendix, Figure A.1.
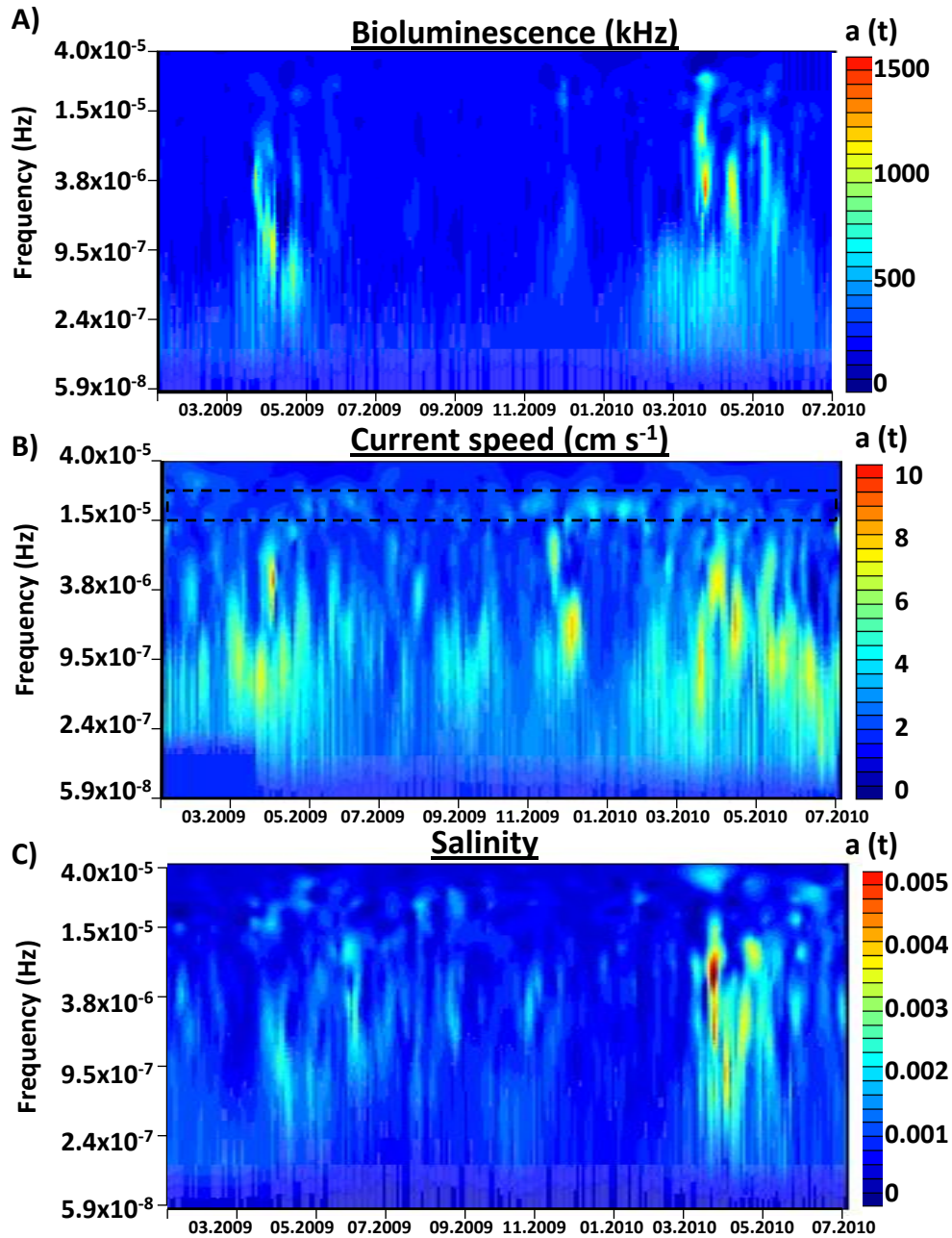
FIGURE 7: Spectral representation using HHT and thin-plate spline smoothing for A) bioluminescence, B) current speed and C) salinity from January 2009 to July 2010. The x-axis is the time. The y-axis the frequency (Hz) and the z-colored-scale is the amplitude coefficient value. In B), the dotted box highlights the frequencies excited by internal waves.

6. Results and discussion for bivariate and cross-analysis

6.1. Bioluminescence mechanically stimulated by current speed

The effects of current speed on bioluminescence activity have been well investigated in the literature (Cussatlegras and Le Gal, 2007, 2005; Fritz et al., 1990). This mechanical stimulation triggers a greater bioluminescent response, enhancing cell membrane excitation in fluid motion (Cussatlegras and Le Gal, 2007; Blaser et al., 2002). However, these studies are all based on dinoflagellates, organisms mainly living in the surface waters. As far as we know, few studies have investigated the role of hydrodynamism on meso- or bathypelagic organisms, such as Hartline et al. (1999) that studied these effects on $Pleuromamma$ $xiphias$, a mesopelagic copepod. Based on ANTARES time series, Tamburini et al. (2013) and Van Haren (2011) show that bioluminescence activity increases with current speed but without clues concerning the classes of organisms impacted by current in this deep environment. This is believed to be due to either impacts of small-sized bioluminescent organisms and particles on the PMTs or the reaction of such organisms to enhanced turbulent motions in the wakes of the PMTs (Priede et al., 2008; Craig et al., 2009).

In Figure 8 A, when crossing the bioluminescence and current speed spectra, from the wavelet decomposition method, coherence coefficients are mostly between 0.6 and 0.95 over a wide range of frequencies. Moreover, there is a high cross coherence coefficient of about 0.85 at frequency approximately $2.0 \times 10^{-5} Hz$ (period of 0.6 days or 14 h), referring to the internal wave frequencies (Huthnance, 1995; Van Haren, 2011). These observations reinforce the well-understood link between current speed and bioluminescence already-known as a mechanical stimulation of bioluminescent organisms by current speed. As no phase delay

is noticed (represented by the right direction of arrows in Figure 8 A), the light emission appears instantaneously to mechanical stimulation.

Thus, bioluminescence is directly linked to current speed at almost all times and frequencies with clear detection of internal waves acting on bioluminescence stimulation with high-amplitude coefficients.

6.2. Links between bioluminescence and new water masses

Tamburini et al. (2013) interpret salinity and temperature time series as proxies of newly formed deep water spreading at the ANTARES site. The authors propose the use of a salinity threshold value to highlight links between newly formed water masses and bioluminescence activity. As shown by these authors, from an ecological point of view, deep-water formations impact bioluminescent organisms in two different ways. On the one hand, these water inputs could carry significant amount of bioluminescent planktonic organisms down from the surface. On the other hand, dissolved organic matter export as well as resuspend particulate organic matter from sediment to the water column to fuel deep-sea bioluminescent activity (Tamburini et al., 2013).

In the present work, Figure 8 B displays results from the wavelet decomposition, crossing bioluminescence and salinity time series. Similar results are obtained crossing bioluminescence and temperature (Figure not shown), and, as mentioned below, both salinity and temperature variables represent proxies of newly formed deep water. Figure 8 B shows two spots of high coherence values ranging from 0.8 to 0.9, from March to May 2009 and March to June 2010. In 2009, high coherence coefficients are highlighted for frequencies of $4.8 \times 10^{-7} Hz$ (period of 24.1 days). In 2010, high correlations are observed for frequencies between 2.4

and $9.5 \times 10^{-7} Hz$ (periods of 48.2 and 12.2 days). This cross-spectrum emphasized a similar range of excited frequencies at $4.8 \times 10^{-7} Hz$ in both March 2009 and 2010 with, in addition, a very high correlation between changes in water mass and bioluminescence activity for both time. Moreover, a delay is observed between water mass and bioluminescence activity (arrows oriented up at $\pi/4$). If no clear explanation can be validated for the low frequency of $4.8 \times 10^{-7} Hz$, then a monthly period of about 24 days can be linked to two phenomena. First, as developed by Tamburini et al. (2013), at the surface, the density of water masses can be increased due to cold and wind during winter. These conditions involve new-water mass formation at the surface, inducing a deep convection in March. The second phenomenon is linked to nutrient regeneration at the surface after the convection event. This involves an increase of chlorophyll-a concentration, at the surface, with a typical annual maximum occurring in spring (Stabholz et al., 2013). Annual variations describe a bloom region in the North-Western Mediterranean Sea around March (D'ortenzio and d'Alcalà, 2009). As a consequence, a maximum of total mass fluxes and high POC fluxes occurred in the deep sea during the spring period, as shown in 2008 and 2009 by Stabholz et al. (2013) in the deep convection area. Such fluxes fuel deep ecosystems, potentially increasing the abundance of bioluminescent organisms and bacteria with a related increase in bioluminescence activity. A potential delay might be due to the biological response sustaining the intense bioluminescence events observed after convection reaching the deep sea. The influence of each of these phenomena must be determined, and deeper observations on the frequency should provide clues on how to interpret the low monthly frequency of $4.8 \times 10^{-7} Hz$ of bioluminescence.

Interestingly, a third bioluminescence event, occurring in December 2009, is observed in the time series (Figure 2) and the spectrogram (Appendix, Figure A.1 A) at frequencies
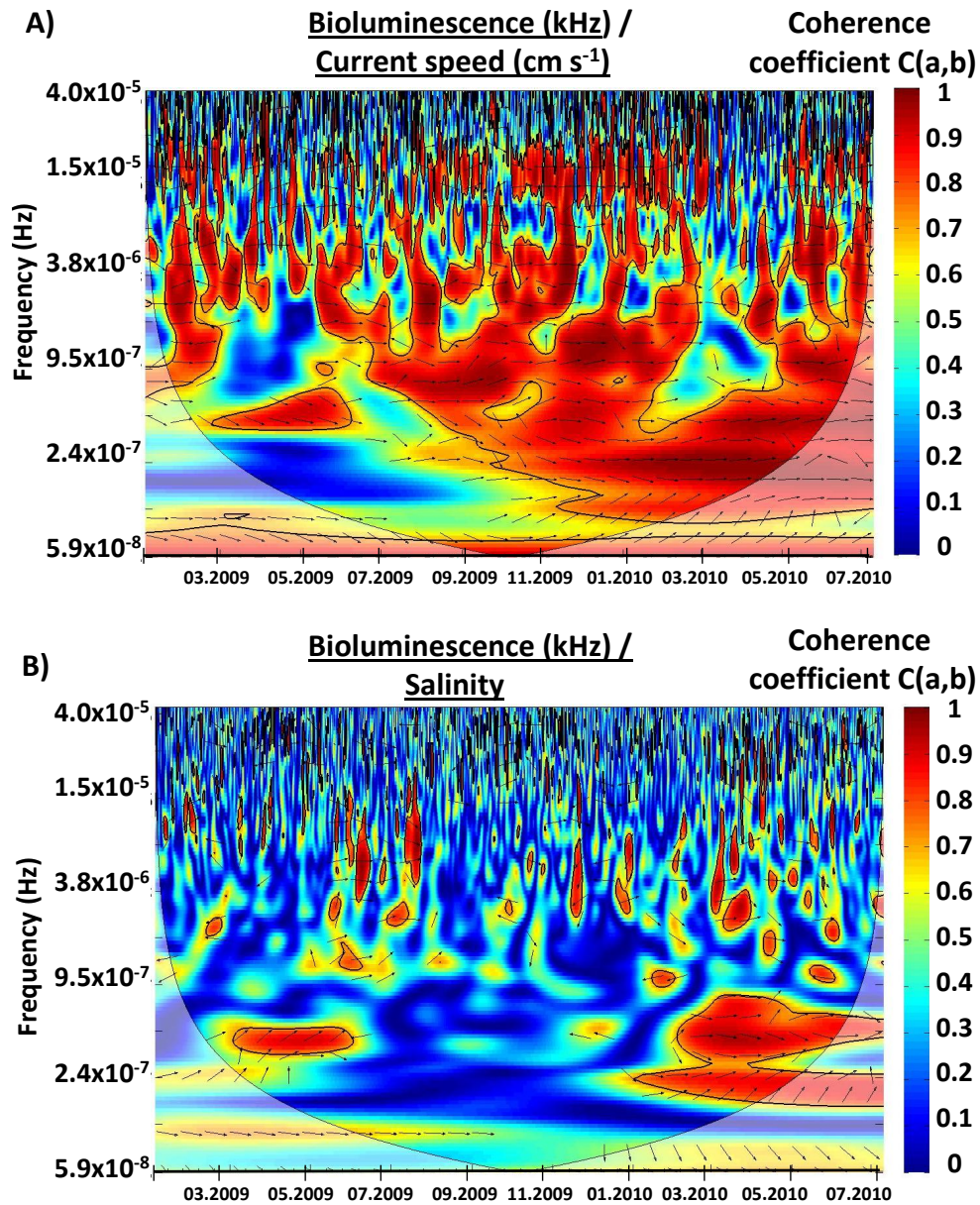
FIGURE 8: Cross spectrogram between bioluminescence and A) current speed and B) salinity, using the wavelet decomposition and coherence measurement. The black arrows represent the phase delay (right arrows mean no phase delay). The black line represents the edge effect for no significant results. The color-scale represents the value for the coherence coefficient from 0 to 1 (blue to red).

between 9.5 and $1.9 \times 10^{-6} Hz$ (period of 6.1 days). Whereas this event is correlated with

current speed in Figure 8 A (coherence coefficient between 0.7 and 0.8), there is no changes

in the water masses (Figure 8 B, coherence values below 0.2). Consequently, in December

2009, current speed generates a mechanical stimulation of bioluminescent organisms and, via its only effect, increases the bioluminescence activity. Figures 8 A and B clearly help to discriminate the events of high bioluminescence intensity affected by newly formed deep water and current speed (April 2009 and 2010) and the events only affected by current speed (December 2009).

To get back to a representation in time and amplitude, we isolate the IMFs referring to the higher correlation coefficients (data not shown). The 3rd IMF for both current speed and bioluminescence are isolated and plotted simultaneously in Figure 9 A. The mean frequency of IMF 3 for bioluminescence is approximately 6.1×10^{-6} (period of 1.9 days), and for current speed, the value is 5.7×10^{-6} (period of 2.0 days). This Figure shows that the high correlation coefficient is mainly due to the two events of high bioluminescence activity with simultaneous oscillation induction. The mean frequency of IMF 5 for salinity is 1.9×10^{-6} (period of 6.1 days) and for IMF 2 for bioluminescence is 1.1×10^{-5} (period of 1.1 days). The use of joint information with critical interpretation is needed with those decomposition methods.

6.3. Two different mathematical methods

From a technical point of view, HHT and the wavelet decompositions provide complementary results. The HHT is based on empirical decomposition, which provides a basis expansion composed with IMFs. This decomposition only relies on the data itself without any assumptions on time series shape. This is not the case for wavelet decomposition, which relies on the choice of the mother wavelet to decompose the signal. Because of the different number of IMFs between variables, time-frequency cross-spectra cannot be defined using HHT, howe-
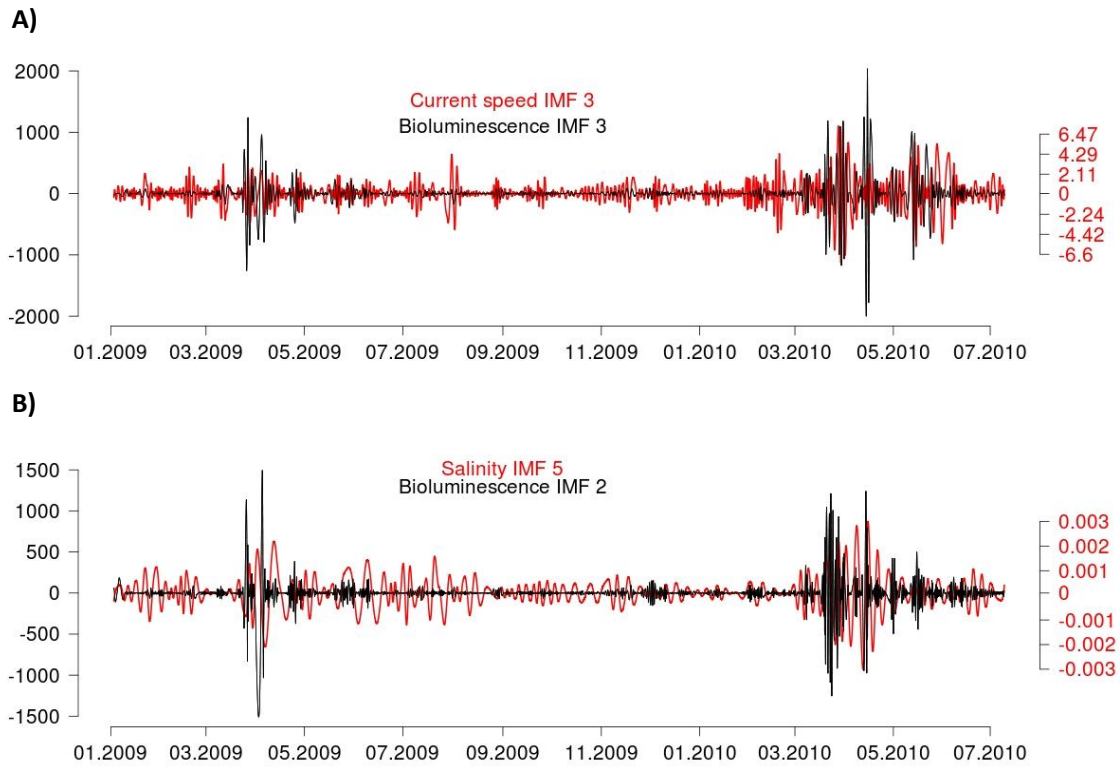
FIGURE 9: Representation of the most correlated intrinsic mode function for A) bioluminescence and current speed, B) bioluminescence and salinity. In A), the mean frequency for bioluminescence IMF 3 is 6.1×10^{-6} (period of 1.9 days), and for the current speed IMF 3 the value is 5.7×10^{-6} (period of 2.0 days). In B), the mean frequency for bioluminescence IMF 2 is 1.1×10^{-5} (period of 1.1 days), and for salinity IMF 5, the value is 1.9×10^{-6} (period of 6.1 days).

ver, the bivariate analysis is more efficient using wavelet decomposition. This method gives phase and delay information for significant cross coefficients improving the ecological interpretation. The use of both wavelet and Hilbert-Huang transform methods not only provides some comparisons in between for further works using time series but also provides a robust interpretation. Moreover, these two methods Further work should also be performed using these two methods with a comparison to the Fourier power spectra to describe theoretical relations with turbulent velocity.

7. Conclusion

There is a clear interest in recording long time series to understand intrinsic variations in an ecosystem. High-frequency sampling helps to detect characteristic scales within a signal. Few time-frequency decomposition methods are adapted to the analysis of non-stationary and non-linear signals and, consequently, to the detection of frequencies excited during unusual events. Both Huang-Hilbert and wavelet decomposition discriminate events due to newly formed deep water spreading and events only due to current speed. Indeed, two different phenomena have been detected involving events of high-intensity bioluminescence. First, bioluminescence is linked to current speed at almost all times and frequencies, without delay and with clear action of internal waves on bioluminescence stimulation. Second, other frequencies with very high coherence coefficients are related to newly formed deep water masses acting on bioluminescence with a small delay and only detected using instantaneous frequencies. It has been demonstrated that such phenomena acts on bioluminescence at lower frequencies. Such periods are close to a month and might be linked to monthly atmospheric phenomena acting at the surface in 2009 and 2010 involving more dense water masses sinking into the deep. Moreover, the spring bloom acting at the surface might also be involved. This period has to be investigated in further studies for more accurate ecological interpretations. Finally, the undeniable advantage of such methods, and the major results in this work, is the possibility of highlighting the correlation between bioluminescence and the arrival of newly formed deep waters at specific time (beginning in March 2009 and March 2010) for frequencies at approximately $4.8 \times 10^{-7} Hz$ (period of 24.1 days). The use of those time-frequency methods applied to signal processing in oceanography for coherent dataset is definitely a method to

better interpret variability, special events and links between variables.

References

Addison, P. S., 2010. The illustrated wavelet transform handbook : introductory theory and applications in science, engineering, medicine and finance. Taylor & Francis, New York.

Ageron, M., Aguilar, J., Al Samarai, I., Albert, A., Ameli, F., André, M., Anghinolfi, M., Anton, G., Anvar, S., Ardid, M., et al., 2011. ANTARES : the first undersea neutrino telescope. Nuclear Instruments and Methods in Physics Research 656 (1), 11–38.

Aguilar, J., Albert, A., Ameli, F., Anghinolfi, M., Anton, G., Anvar, S., Aslanides, E., Aubert, J.-J., Barbarito, E., Basa, S., et al., 2007. The data acquisition system for the ANTARES neutrino telescope. Nuclear Instruments and Methods in Physics A 570 (1), 107–116.

Ahuja, N., Lertrattanapanich, S., Bose, N., 2005. Properties determining choice of mother wavelet. Vision, Image and Signal Processing, IEE Proceedings 152 (5), 659–664.

Amram, P., Anvar, S., Aslanides, E., Azoulay, R., Basa, S., Benhammou, Y., Bernard, F.,

Bertin, V., Billault, M., et al., 2000. Background light in potential sites for the ANTARES undersea neutrino telescope. Astroparticle Physics 13 (2), 127–136.

Béthoux, J., Durieu de Madron, X., Nyffeler, F., Tailliez, D., 2002. Deep water in the western Mediterranean : peculiar 1999 and 2000 characteristics, shelf formation hypothesis, variability since 1970 and geochemical inferences. Journal of Marine Systems 33, 117–131.

Blaser, S., Kurisu, F., Satoh, H., Mino, T., 2002. Hydromechanical stimulation of bioluminescent plankton. Luminescence 17 (6), 370–80.

Bookstein, F. L., 1989. Principal warps : Thin-plate splines and the decomposition of deformations. IEEE 11 (6), 567–585.

Cazelles, B., Chavez, M., Berteaux, D., Ménard, F., Vik, J., Jenouvrier, S., Stenseth, N. C., 2008. Wavelet analysis of ecological time series. Oecologia 156 (2), 287–304.

Craig, J., Jamieson, A. J., Heger, A., Priede, I. G., 2009. Distribution of bioluminescent organisms in the Mediterranean Sea and predicted effects on a deep-sea neutrino telescope. Nuclear Instruments and Methods in Physics Research Section A 602 (1), 224–226.

Cussatlegras, A.-S., Le Gal, P., 2005. Dinoflagellate bioluminescence in response to mechanical stimuli in water flows. Nonlinear Processes in Geophysics 12 (3), 337–343.

Cussatlegras, A.-S., Le Gal, P., 2007. Variability in the bioluminescence response of the dinoflagellate *Pyrocystis lunula*. Journal of Experimental Marine Biology and Ecology 343 (1), 74–81.

D'ortenzio, F., d'Alcalà, M. R., 2009. On the trophic regimes of the mediterranean sea : a satellite analysis. Biogeosciences 6 (2).

Favali, P., Beranzoli, L., 2006. Seafloor observatory science : a review. Annals of Geophysics 49 (2-3).

Flandrin, P., Rilling, G., Goncalves, P., 2004. Empirical mode decomposition as a filter bank. Signal Processing Letters, IEEE 11 (2), 112–114.

Frazier, M. W., 1999. An introduction to wavelets through linear algebra. Springer-Verlag, New York.

Fritz, L., Morse, D., Hastings, J., 1990. The circadian bioluminescence rhythm of gonyaulax is related to daily variations in the number of light-emitting organelles. Journal of Cell Science 95 (2), 321–328.

Ghorbani, M., 2013. Testing the weak stationarity of a spatio-temporal point process. Stochastic Environmental Research and Risk Assessment 27 (2), 517–524.

Grinsted, A., Moore, J. C., Jevrejeva, S., 2004. Application of the cross wavelet transform and wavelet coherence to geophysical time series. Nonlinear Processes in Geophysics 11 (5/6), 561–566.

Hartline, D., Buskey, E., Lenz, P., 1999. Rapid jumps and bioluminescence elicited by controlled hydrodynamic stimuli in a mesopelagic copepod, pleuromamma xiphias. The Biological Bulletin 197 (2), 132–143.

Hewitt, J., Thrush, S., Dayton, P., Bonsdorff, E., 2007. The effect of spatial and temporal

heterogeneity on the design and analysis of empirical studies of scale-dependent systems. The American Naturalist 169 (3), 398–408.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C., Liu, H. H., 1998. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London 454 (1971), 903–995.

Huang, N. E., Wu, Z., 2008. A review on Hilbert-Huang transform : method and its applications to geophysical studies. Review of Geophysics 46 (2007), 1–23.

Huang, N. E., Wu, Z., Long, S. R., Arnold, K. C., Chen, X., Blank, K., 2009a. On instantaneous frequency. Advances in Adaptive Data Analysis 1 (2), 177–229.

Huang, Y., Schmitt, F. G., 2014. Time dependent intrinsic correlation analysis of temperature and dissolved oxygen time series using empirical mode decomposition. Journal of Marine Systems 130, 90–100.

Huang, Y., Schmitt, F. G., Lu, Z., Liu, Y., 2009b. Analysis of daily river flow fluctuations using empirical mode decomposition and arbitrary order hilbert spectral analysis. Journal of Hydrology 373 (1), 103–111.

Huthnance, J. M., 1995. Circulation, exchange and water masses at the ocean margin : the role of physical processes at the shelf edge. Progress in Oceanography 35 (4), 353–431.

Kwiatkowski, D., Phillips, P. C., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. Journal of Econometrics 54 (1-3), 159–178.

Mardia, K., Kent, J., Goodall, C., Little, J., 1996. Kriging and splines with derivative information. Biometrika 83 (1), 207–221.

Marshall, J., Schott, F., 1999. Open-ocean convection : Observations, theory, and models. Reviews of Geophysics 37 (1), 1–64.

Massei, N., Fournier, M., 2012. Assessing the expression of large-scale climatic fluctuations in the hydrological variability of daily Seine river flow (France) between 1950 and 2008 using Hilbert-Huang transform. Journal of Hydrology 448-449, 119–128.

Paget, D., Elliott, D., 1972. An algorithm for the numerical evaluation of certain Cauchy principal value integrals. Numerische Mathematik 19 (5), 373–385.

Priede, I. G., Jamieson, A., Heger, A., Craig, J., Zuur, A. F., 2008. The potential influence of bioluminescence from marine animals on a deep-sea underwater neutrino telescope array in the Mediterranean Sea. Deep Sea Research Part I 55 (11), 1474–1483.

Rao, A. R., Hsu, E.-C., 2008. Hilbert-Huang transform analysis of hydrological and environmental time series. Vol. 60. Springer.

Send, U., Lankhorst, M., 2011. The global component of the US Ocean Observatories Initiative and the global OceanSITES project. In : OCEANS 2011. IEEE, pp. 1–3.

Stabholz, M., Durrieu de Madron, X., Canals, M., Khripounoff, A., Taupier-Letage, I., Testor, P., Heussner, S., Kerhervé, P., Delsaut, N., Houpert, L., et al., 2013. Impact of open-ocean convection on particle fluxes and sediment dynamics in the deep margin of the gulf of lions. Biogeosciences 10 (2), 1097–1116.

Tamburini, C., Canals, M., Durrieu de Madron, X., Houpert, L., Lefevre, D., Martini, S., D'Ortenzio, F., Robert, A., Testor, P., Aguilar, J., et al., 2013. Deep-sea bioluminescence blooms after dense water formation at the ocean surface. PLoS ONE 8 (7), e67523.

Torrence, C., Compo, G., 1998. A practical guide to wavelet analysis. Bulletin of the American Meteorological Society 79 (1), 61–78.

Torrence, C., Webster, P. J., 1999. Interdecadal changes in the ENSO-monsoon system. Journal of Climate 12 (8), 2679–2690.

Van Haren, H., 2011. Meso- and small-scale vertical motions in the deep Western Mediterranean. Nuclear Instruments and Methods in Physics 626-627, S84–S86.

Wood, S. N., 2003. Thin plate regression splines. Journal of the Royal Statistical Society 65 (1), 95–114.
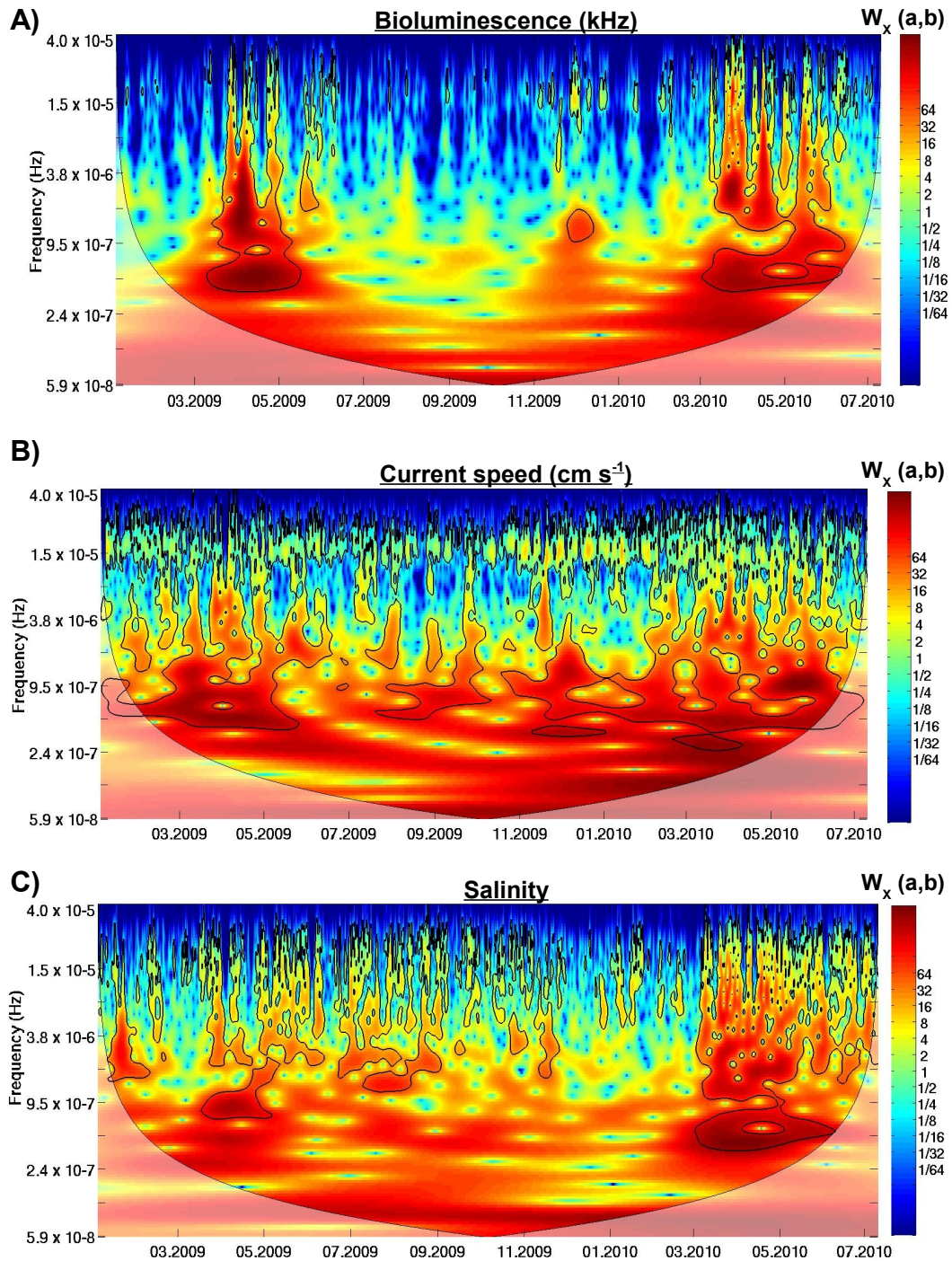
A. Appendix

FIGURE A.1: Time-frequency representation using the wavelet decomposition for A) bioluminescence, B) current speed and C) salinity time series. The color scale for $W_X(a,b)$ is from blue (low coefficient) to red (high coefficient). For the bioluminescence decomposition, three periods with high coefficients are distinct around the dates 04.2009, 12.2009 and 04.2010. The black-contour lines represent the 5% significant level. The shaded part represents the cone of influence where edge effects appear.