Anomaly detection in VHF radar measurements

Véronique Cochin*†, Grégoire Mercier†, René Garello†, Vincent Mariette* and Pierre Broche‡
*ACTIMAR, 24 quai de la douane, 29200 Brest - France
†GET-ENST Bretagne, CNRS 2856 TAMCIC, Technopôle Brest-Iroise CS 83818, 29238 Brest, France
‡LSEET Université de Toulon et du Var, CNRS 6017, BP 132, 83957 La Garde Cedex, France
Email: veronique.cochin@enst-bretagne.fr

Abstract— **The present work develops a new aspect in the use of HF radar as a tool for real time operational oceanography. We propose to implement new algorithms for detecting anomalies (outliers) on VHF radar data. Perspectives of such work are to detect changes in sea surface, such as ship or oil pollution. In this preliminary work, no information are given about the anomalies we are looking for. We propose to study new detection methods based on Support Vector expansion to perform a binary separation between observations. This technique is well adapted to HF radar observations that are never statistically representative but give a good temporal resolution.**

I. INTRODUCTION

Remote sensing by shore based High Frequency (HF) and Very High Frequency (VHF) radars (typically between 5 and 50 MHz) is an important tool to improve monitoring of coastal environment. Since Crombie's study on the backscattering of radio waves by the ocean surface [1], the principles underlying the HF Doppler radar techniques have been investigated theoretically [2] [3]. This concept of measurement received considerable attention in coastal oceanographic experiments [4] [5]. Propagation is mainly done by ground wave propagation and measurements are achieved through the Doppler spectrum of the measured backscatter. HF radars are able to estimate, in quasi real time, sea surface current velocity, wave spectrum and wind direction. Their main advantage is to provide measurements over a large scale, with high spatial (600 meters down to 150 meters) and temporal resolution (one measurement every 30 minutes, down to 10 minutes).

This concept of measurement takes into consideration the physical interpretation of the Doppler spectrum (oceanographic parameters such as sea surface currents, sea states, wind). This makes use of assumptions and theoretical models for extractions of the parameters. Our approach does not take into consideration the physical meaning of the Doppler spectra, but similarities between the spectra. This new aspect in data processing is useful before or in complement of existing algorithm developed to extract ocean parameters. It makes use of temporal resolution in HF radar data.

Anomalies in HF radar Doppler spectrum may be considered as a specific novelty detection (outliers identification) problem. A new detection method, based on Support Vector Machine (SVM) expansion to perform a binary separation between observations is developed. In this paper, we describe implementation of new algorithms for detecting anomalies (outliers) in the Doppler spectrum. The strategy is to use HF radar high temporal resolution, to detect changes in the Doppler spectrum, and to provide an automatic tool for anomaly detection without any physical interpretation of the Doppler spectrum. Perspectives are to develop a new aspect in the use of HF radar measurements as a real time operational oceanography tool, particularly for littoral surveillance (ships, oil spill detection).

To test this new approach, and due to the absence of operational data containing oil spill, we used measurements from the VHF *Courants Océaniques de Surface MEsurés par Radar* (COSMER) system acquired during the program *Evaluation et Prévision des conditions d'Environnement Littorales* (EPEL), supported by the French naval hydrographic and oceanographic service (*Service Hydrographique et Océanographique de la Marine*, SHOM) [6] [7]. A description of the VHF radar measurement principle is done in section II. In section III, we present the approach of SVM algorithm and kernel-based detection. Then, in section IV, we focus on the algorithm applied in case of anomaly in the Doppler spectrum such as ship echo's used as an example in this preliminary work for training the selected kernel.

II. VHF RADAR MEASUREMENTS

A number of important properties of the ocean wave spectrum can be obtained by looking at the Doppler spectrum of the backscattered radar HF signal. Ground wave radar energy is backscattered from the moving ocean by surface gravity ocean waves. First order scatter from specific spectral components of the ocean wave field produces the dominant contribution. These components, termed "Bragg waves", have a wavelength of one half the radar wavelength λ_{radar} and move radially toward or away from the radar station ("Bragg resonant scattering" principle). The echo Doppler spectrum (fig.1) consists of two distinct peaks, resulting from the wave components, symmetrically positioned about the radar frequency. These peaks are displaced according to the phase velocity of the ocean waves:

$$v_v = \sqrt{\frac{g}{k}}, \tag{1}$$

where $k = 2\pi/\lambda$ is the wavenumber for the ocean waves and g is the acceleration due to gravity.

In the absence of ocean currents, the Doppler frequency shift always occurs at a known position that depends only on the radar transmitter frequency. This is termed "Bragg frequency":

$$f_B = \sqrt{\frac{g\,k_{\text{radar}}}{2\pi^2}} \tag{2}$$

where k_{radar} is radar wave vector pointing in backscattered sea surface.

If a surface current is present, these peaks are shifted in frequency from their theoretical position (f_B) by an amount Δf. This shift is related to the radial component of the effective surface current Δv:

$$\Delta v = \frac{\lambda_{\text{radar}}}{2} \Delta f \tag{3}$$

where a positive (respectively negative) value for the velocity means that the target is departing from (respectively approaching) the radar station. The ratio between the intensities of the two first order echoes depends strongly on the direction of the wind over the ocean. So, the radial current velocity, extracted from the Doppler frequency of the strongest first order echo, is usually taken as the actual radial current velocity. The surface vector current is estimated by combining the components along two radar beams coming from two radar stations.

The rest of the power spectrum comprises a continuum, referred to as the second order part of the spectrum and a noise floor. The second order sea echoes carry the information on ocean waves and sea state. They are due to higher order scatter, the greater part of which arises from second order interaction between two ocean waves components constituting the total ocean wave field. Several approaches have been developed to provide a theoretical formulation for relating the Doppler spectrum of the backscatter cross section to the complete ocean wave directional spectrum. Inversion methods of the nonlinear integral equation are developed to estimate ocean wave spectra [4] [8].

This concept of measurement has been used for many years. In this study, we used VHF COSMER data. This system has been developed by the *Laboratoire de Sondages Electromagnétiques de l'Environnement Terrestre* (LSEET, France) [6]. It includes two pulsed Doppler radars (respectively 45MHz and 47.8MHz) to map coastal sea surface currents over a large area, of about 25km by 25km. During EPEL program, this system has been deployed in operational mode with a radial resolution of 600m and a beam width of 13° (linear phased array of 8 antennas) [7]. Radar measurements were acquired continuously during 28 days, providing one set of derived parameters (namely an average spectrum) every 30 minutes. The radar data obtained after real time signal processing consist of the 256 points Doppler spectra (Fast Fourier Transform (FFT) algorithm [9]). These spectra are given for all ranges and azimuth, so called *radar images*. The database contains 2688 radar images, issued from the two radar stations. For this study, the dimensions of the radar images are range, azimuth and spectral points (typically 41x8 pixels).

III. Anomaly detection algorithm

Kernel technique, such as SVM is among the most powerful nonparametric tool for classification, regression and novelty (outliers) detection. Complete mathematical formulation can be found in [10] [11].

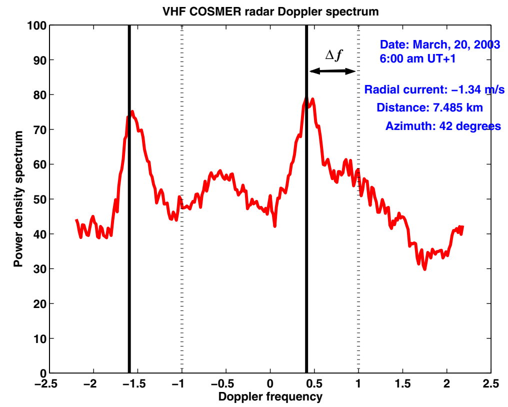

Fig. 1. Typical radar sea echo spectrum from VHF COSMER radar (EPEL program) showing the Doppler-shifted peaks away from the theoretical position of the radar Bragg peaks.

A. Principle

One basic approach to novelty detection is based on estimating the support of the data, the domain in which most of the data are. Samples that fall outside of the domain are declared to be outliers. Therefore, the problem is considered to be one of "single class problem": either data values belong to the data domain, else they are outliers. Since the data are unlabeled, this technique is an "unsupervised learning". Thus, the geometric approach tries to estimate a function f, which is supposed to be positive when the data are in the region or negative on the complement.

Suppose that a set of training vectors $X = x_1, .., x_m$ is available in the input space, denoted χ , where m is the number of observations. Each x_i is supposed to be "support vector": it will be used as a support to determine a separating hyperplane characterized by the normal vector w. Then, the detection function is defined by comparing the data to the separating plan such that

$$\begin{aligned} f(x) &= < w, x > -\rho, \\ &\begin{cases} f(x) > 0 & \text{if } x \text{ is similar to } \chi, \\ f(x) < 0 & \text{if } x \text{ is not similar to } \chi, \\ & \quad \text{then considered as an outlier.} \end{cases} \end{aligned} \tag{4}$$

where w is the normal vector, ρ is an offset and $< \cdot, \cdot >$ is the *dot product* (also referred to as *inner product* or *scalar product*). The hyperplane is defined by its parameters:

$$< w, x > -\rho = 0 \tag{5}$$

B. Kernel-based detection

The strategy of the parametric approach is to map the data into the feature space corresponding to the kernel $k(x, x_i)$, and to separate them from the origin with a maximum range. This method is used to simulate a non-linear projection of the data from a higher dimension space (feature space) where linear separability may be achieved. Computation can be done without an explicit form of the projection, but only with the kernel corresponding to the dot product between projections

[12] [13]. Thus the decision function is transformed into a kernel expansion:

$$f(x, \alpha) = \sum_i \alpha_i k(x, x_i) - \rho \qquad (6)$$

To define the hyperplane, Lagrangian multipliers α are introduced, such that they are non-zeros only for the support vectors. We obtain the dual problem:

$$\text{minimize } \frac{1}{2} \sum_{ij} \alpha_i \alpha_j k(x_i, x_j)$$

$$\text{subject to } 0 \leqslant \alpha_i \leqslant \frac{1}{\nu m}, \quad \sum_i \alpha_i = 1$$

where ν acts as a parameter linked to the selectivity of "abnormal" detection.

The value $f(x)$ is determined by evaluating in which side of the hyperplane a new observation x falls in the feature space.

C. Kernel for Doppler spectrum

Every kernel that satisfies Mercer's conditions [10] can be considered as an eligible kernel. Examples of classical kernels are given in [11]. To study similarities between the spectra, we used Spectral Angle (SA) kernel, which is defined in order to measure the spectral difference (or similarity in the spectrum shape) between x and y. Spectral Angle $\theta(x, y)$ for x, $y \in \chi$ is defined by:

$$\theta(x, y) = \arccos \frac{< x, y >}{\|x\| \quad \|y\|} \qquad (7)$$

Then, the spectral kernel is defined in a similar way as the well known Gaussian Radial Basis Function (RBF) kernel by:

$$k(x, y) = e^{-\theta(x,y)} \qquad (8)$$

This Spectral Angle kernel is robust to differences of the overall energy and has already been applied in hyperspectral data for classification purpose [14].

IV. Illustration with Ship Echoes

The target is modeled with its Radar Cross Section (RCS) and Doppler shift, caused by the radial component of its velocity. Large ships with significant RCS can exhibit strong returns that are clearly apparent within the Doppler spectrum [15]. Unlike the position of the first-order ocean echo within the Doppler spectrum (Bragg frequency), the position of the ship echo's peak is proportional to the radar carrier frequency.

In this preliminary work, no information is given about the anomalies we are looking for. Nevertheless, in order to test the training of the kernel, we chose a situation where the Doppler spectrum contained a peak due to ship. Ship detection requires specific processing on the data [16] [17], particularly when the ship echo's peak is close to the Bragg peak. The purpose here is to focus on an "abnormal" spectrum, containing a peak due to ship echo. An example of ship echo is given in fig.2. To illustrate preliminary result in novelty detection, fig.3 shows the position of the ship echo's detected by the algorithm. It is represented on a sea surface current map, acquired during EPEL program by VHF COSMER radar.

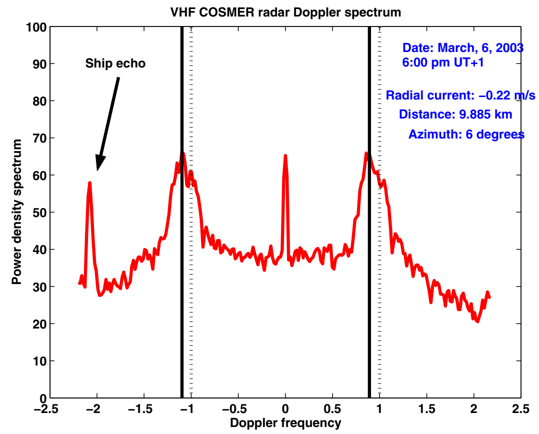

Fig. 2. Radar sea echo spectrum from VHF COSMER radar (EPEL program) showing the Doppler-shifted peaks away from the theoretical position of the radar Bragg peaks and showing peak due to ship echo (on the left).

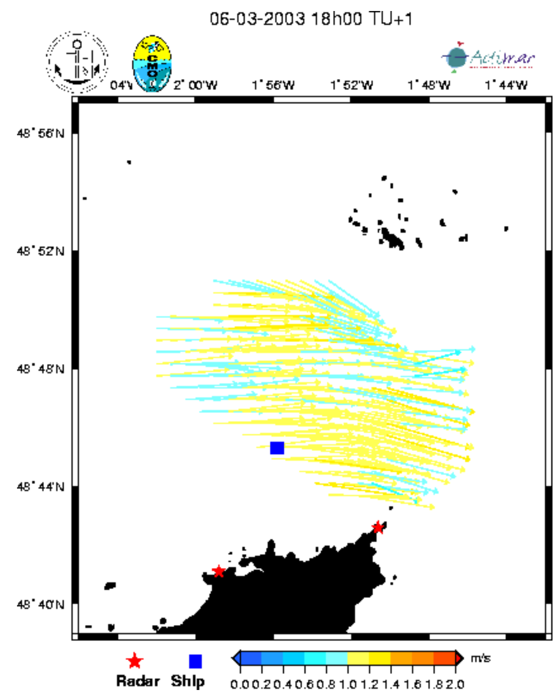

Fig. 3. Illustration of ship echo detected as an outlier by the algorithm in the radar image, reported on sea surface current map (VHF COSMER radar, EPEL program).

V. Conclusion and Perspectives

The objective of this study is to provide a real time tool in operational coastal monitoring. In this preliminary work, we have validated the use of kernel-based anomaly detection to process in Doppler frequency spectrum with no physical meaning. Works still have to be done in order to detect oil spill in terms of anomalies in the spectrum. This way, oil spill should be considered as anomaly. In this case, the difficulty remains in choosing the appropriate kernel, which can characterize changes that can occur in the ocean wave directional spectrum or in the Doppler spectrum. Perspectives are to find the appropriate kernel, which can integrate the effects induced by the presence of oil spill.

ACKNOWLEDGMENT

The authors gratefully acknowledge SHOM and particularly Fabrice Ardhuin. We also wish to thank LSEET group and A. Coat (Actimar) for support in collecting the VHF COSMER data.

REFERENCES

[1] D. D. Crombie, "Doppler spectrum of sea echo at 13.56 Mc/s," *Nature*, vol. 175, no. 4459, pp. 681–682, April 1955.

[2] D. E. Barrick, "First-order theory and analysis of MF/HF/VHF scatter from the sea," *IEEE Trans. Antennas Propagat.*, vol. 20, no. 1, pp. 2–10, January 1972.

[3] D. E. Barrick, M. W. Evans, and B. Weber, "Ocean surface currents mapped by radar," *Science*, vol. 198, pp. 138–144, 1977.

[4] L. R. Wyatt, "Limits to the inversion of HF radar backscatter for ocean wave measurement," *Journal of Atmospheric and Oceanic Technology*, vol. 17, pp. 1651–1665, December 2000.

[5] K. Gurgel, H. Essen, and S. Kingsley, "High-frequency radars: physical limitations and recent developments," *Coastal Engineering*, vol. 37, no. 3-4, pp. 201–218, August 1999.

[6] P. Broche, J. C. Crochet, J. C. de Maistre, and P. Forget, "VHF radar for ocean surface current and sea state remote sensing," *Radio Science*, vol. 22, pp. 69–75, 1987.

[7] V. Cochin, V. Mariette, A. Coat, P. Broche, J. C. de Maistre, Y. Barbin, J. Gaggelli, G. Mercier, and R. Garello, "Tidal analysis and currents mapping using VHF radar and ADCP measurements in the Normand Breton Gulf (France)," in *OCEANOPS 2004 proceedings*, 2004.

[8] B. J. Lipa and D. E. Barrick, "Extraction of sea state from HF radar sea echo: Mathematical theory and modeling." *Radio Science*, vol. 21, no. 1, pp. 81–100, 1986.

[9] D. E. Barrick, "Accuracy of parameter extraction from sample averaged sea echo Doppler spectra," *IEEE Trans. Antennas Propagat.*, vol. 28, no. 1, pp. 1–11, 1980.

[10] V. Vapnik, *Statistical Learning Theory*, Wiley, Ed., New York, 1998, forthcoming.

[11] B. Schölkopf and A. Smola, *Learning with Kernels*, MIT press, Ed., Cambridge, MA, 2002.

[12] B. Schölkopf, R. Williamson, A. Smola, J. Shawe-Taylor, and J. C. Platt, "Support vector method for novelty detection," *Advances in Neural Information Processing Systems*, vol. 12, pp. 582–588, 2000.

[13] B. Schölkopf, J. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the support of a high-dimensional distribution," *Neural Computation*, vol. 13, no. 7, pp. 1443–1471, 2001.

[14] G. Mercier and M. Lennon, "Support vector machines for hyperspectral image classification with spectral-based kernels," in *Proc. of the IEEE Geoscience and Remote Sensing Symp. (IGARSS 2003)*, vol. 1, no. 21-25, 2003, pp. 288–290.

[15] A. M. Ponsford, L. Sevgi, and H. C. Chan, "An integrated maritime surveillance system based on surface wave HF radars, Part II: operational status and system performance," *IEEE Trans. Antennas Propagat.*, vol. 43, no. 5, pp. 52–63, 2001.

[16] R. Khan, B. Gamberg, D. Power, J. Walsh, B. Dawe, W. Pearson, and D. Millan, "Target detection and tracking with a high frequency ground wave radar," *IEEE J. Oceanic Eng.*, vol. 19, no. 4, pp. 540–548, 1994.

[17] D. M. Fernandez, J. F. Vesecky, C. C. Teague, J. D. Paduan, and K. E. Laws, "Ship detection with high-frequency phased-array and direction-finding radar systems," in *Proc. of the IEEE Geoscience and Remote Sensing Symp. (IGARSS 1998)*, vol. 1, 1998, pp. 204–206.