A New Way to Model Nonstationary Sea Clutter

Jing Hu, *Member, IEEE*, Wen-wen Tung, and Jianbo Gao, *Member, IEEE*

***Abstract*—Sea clutter refers to the radar backscatter from a patch of ocean surface. To properly characterize radar clutter returns, a lot of effort has been made to fit various distributions to the observed amplitude data of sea clutter. However, the fitting of real sea clutter data using those distributions is not satisfactory. This may be due to the fact that sea clutter data is highly nonstationary. This nonstationarity motivates us to perform distributional analysis on the data obtained by differentiating the amplitude data of sea clutter. By systematically analyzing differentiated data of 280 sea clutter time series measured under various sea and weather conditions, we show that the Tsallis distribution fits sea clutter data much better than commonly used distributions for sea clutter such as the K distribution. We also find that the parameters from the Tsallis distribution are more effective than the ones from the K distribution for detecting low observable targets within sea clutter.**

***Index Terms*—K distribution, sea clutter, target detection, Tsallis distribution.**

I. Introduction

SEA clutter refers to the radar backscatter from a patch of sea surface illuminated by a radar pulse. Statistical models properly characterizing radar clutter processes are critical for designing optimum algorithms for detecting targets within radar clutter as well as for performance evaluation of those detectors. In modern radar systems operating at low grazing angles with high resolution capabilities, sea clutter has often been observed to be highly non-Gaussian [1]–[5] and, hence, is very challenging to model. A lot of effort has been made to fit various distributions to the observed amplitude data of sea clutter, including Weibull [1], log-normal [2], K [3], [4], and compound-Gaussian [5] distributions. However, the fitting of those distributions to real sea clutter data is not satisfactory, and quite often using parameters estimated from those distributions is not very effective for distinguishing sea clutter data with targets from those without targets [6].

The limitation of the aforementioned distribution models may be due to the fact that sea clutter data are highly nonstationary. As we shall explain in Section II, one good way of transforming nonstationary processes to stationary ones is through differentiation [7]. Therefore, we perform distributional analysis on the data obtained by differentiating the amplitude data of sea clutter. Specifically, we shall employ Tsallis distribution [8], [9], which is thought to provide a foundation for stable laws [10]. Earlier, by analyzing a few sea clutter datasets, we have shown that Tsallis distribution is a promising model for sea clutter [11]. Here, we shall systematically analyze the differentiated data of 280 sea clutter time series measured under various sea and weather conditions, and we shall show that the differentiated data can be well fitted by the Tsallis distribution. Furthermore, we shall quantitatively evaluate how much better the Tsallis distribution is over the commonly used distribution for sea clutter modeling, the K distribution, in terms of distribution fitting as well as for detecting small targets within sea clutter.

Manuscript received July 01, 2008; revised October 14, 2008. Current version published January 16, 2009. This work is supported in part by the NSF under Grant no. CMMI-0825311. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Deniz Erdogmus.

J. Hu and J. Gao are with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611 USA (e-mail: jinghu@ufl.edu; gao@ece.ufl.edu).

W.-W. Tung is with the Department of Earth and Atmospheric Sciences, Purdue University, West Lafayette, IN 47907 USA (e-mail: wwtung@purdue.edu).

Digital Object Identifier 10.1109/LSP.2008.2009844

II. Complexities of Sea Clutter Data

A. Data

We have obtained 14 sea clutter measurements from a website maintained by Prof. S. Haykin: http://soma.ece.mcmaster.ca/ipix/dartmouth/datasets.html. The measurement was made using the McMaster IPIX radar at Dartmouth, Nova Scotia, Canada. The radar was mounted in a fixed position on land 25–30 m above sea level, with the operating (carrier) frequency 9.39 GHz (and hence a wavelength of about 3 cm). It was operated at low grazing angles, with the antenna dwelling in a fixed direction, illuminating a patch of ocean surface. The measurements were performed with the wave height in the ocean varying from 0.8 m to 3.8 m (with peak height up to 5.5 m) and the wind conditions varying from still to 60 km/hr (with gusts up to 90 km/hr). For each measurement, 14 areas, called antenna footprints or range bins, were scanned. Their centers were depicted as $B_1, B_2, \ldots, B_{14}$ in Fig. 1. The distance between two adjacent range bins was 15 m. One or a few range bins (say, B_{i-1}, B_i, and B_{i+1}) hit a target, which was a spherical block of styrofoam of diameter 1 m, wrapped with wire mesh. The locations of the three targets were specified by their azimuthal angle and distance to the radar. They were (128 degree, 2660 m), (130 degree, 5525 m), and (170 degree, 2655 m), respectively. The range bin where the target is strongest is labeled as the primary target bin. Due to drift of the target, bins adjacent to the primary target bin may also hit the target. They are called secondary target bins. For each range bin, there were 2^{17} complex numbers, sampled with a frequency of 1000 Hz.

We analyze amplitude data of two polarizations, horizontal transmission, horizontal reception (HH) and vertical transmission, vertical reception (VV). Careful examination of the amplitude data indicates that four datasets were severely affected by clipping. We discard those four datasets, and analyze the remaining ten measurements, which contain 280 sea clutter time series.

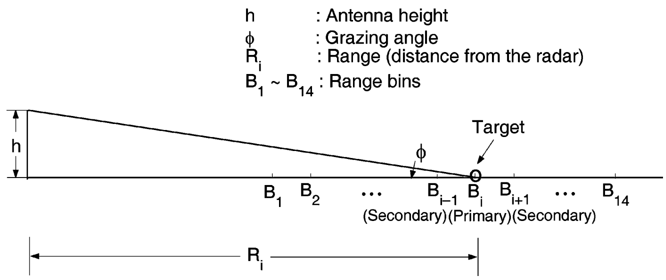

Fig. 1. Schematic showing how the sea clutter data were collected.

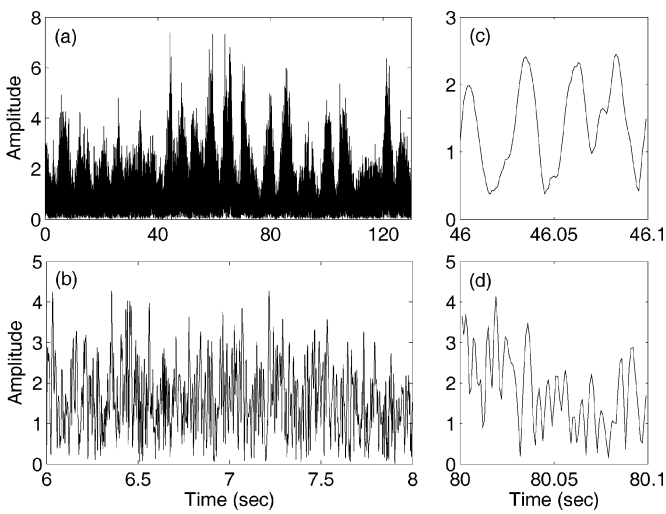

Fig. 2. Example of the sea clutter amplitude data. (a) Entire data, (b) a 2-s segment, and (c, d) two 0.1-s segments extracted from the data.

B. Long-Range Correlations and Nonstationarity in Sea Clutter

There are two sources of complexity in sea clutter, the rough sea surface, sometimes oscillatory and sometimes turbulent, and the multipath propagation of radar backscatter from the sea surface. To appreciate the complexity, especially the multiscale nature of sea clutter, in Fig. 2, we have plotted a typical sea clutter signal on three different time scales, 130 s, 2 s, and 0.1 s. It is clear that the signal is not purely random, since the waveform can be fairly smooth on short time scales [Fig. 2(c)]. However, the signal is *doubly* nonstationary: 1) the randomness of the signal changes dramatically over time [Fig. 2(a) and (b)]; 2) the frequency of the signal changes from time to time [Fig. 2(c) and (d)], indicating the ever changing nature of the wave pattern on the sea surface.

By carefully carrying out fractal and multifractal analysis of sea clutter [6], [12], [13], we have found two important time scales, one about 0.01 s, the other a few seconds (more precisely, with the sampling frequency being 1000 Hz, they correspond to about 2^4 samples and 2^{12} samples, respectively). In between, sea clutter is a type of $1/f^\alpha$ processes with $1 < \alpha < 3$. The defining property of such processes is that their variances increase with time t as $t^{\alpha-1}$ [13]. Therefore, sea clutter data are nonstationary on time scales greater than about 0.01 s but smaller than a few seconds.

The most well-known example of ideal $1/f^\alpha$ $(1 < \alpha < 3)$ processes is the standard Brownian motion (Bm) process, with $\alpha = 2$. The generalization of Bm processes is called fractional Brownian motion (fBm), which are Gaussian process with mean 0, stationary increments, variance

$$E[(B_H(t))^2] = t^{2H} \tag{1}$$

and covariance

$$E[B_H(s)B_H(t)] = \frac{1}{2}\left\{s^{2H} + t^{2H} - |s-t|^{2H}\right\}. \tag{2}$$

Differentiation of fBm yields stationary fractional Gaussian noise (fGn), whose autocorrelation function $R(k)$ decays in a power-law manner, $R(k) \sim k^{2H-2}$, as $k \to \infty$, where $0 < H = (\alpha-1)/2 < 1$ is called the Hurst parameter. Depending on whether H is smaller than, equal to, or larger than $1/2$, the fGn noise is said to have anti-persistent correlation, short-range correlation, and persistent long-range correlation [10], [13].

The above discussions make it clear that one effective means of transforming the nonstationary sea clutter data to stationary ones is through differentiation. Note that differentiation can also transform a nonstationary process with trend in the mean or periodic components to a quasi-stationary process [7], and it is at the foundation of the popular unit root tests for detecting nonstationarity in autoregressive (AR)-based time series modeling, including the augmented Dickey–Fuller test (ADF) [14], [15].

It should be emphasized, however, that the random walk (or diffusion) model employed here is fundamentally different from the particular first-order AR model of the type

$$y(n+1) = y(n) + \eta(n) \tag{3}$$

where $\eta(n), n \geq 0$ are a series of uncorrelated noise, and $y(0)$ typically may not be zero. This is because our random walk model allows long-range correlations in the noise process. Consequentially, AR-based popular tests for nonstationarity, such as unit root tests and ADF test [14], [15], cannot be straightforwardly used to test nonstationarity in sea clutter (recalling the long-range correlations on time scales from about 2^4 samples to 2^{12} samples in sea clutter, for an AR model to be effective, it may have to involve about 4000 time lags; therefore, is very costly computationally).

III. Distributional Analysis of Sea Clutter

Denote the sea clutter amplitude data by $y(n), n = 1, 2, \ldots$. The differentiated data of sea clutter is denoted as

$$x(n) = y(n+1) - y(n), \quad n = 1, 2, \ldots. \tag{4}$$

We fit the differentiated data of sea clutter by the Tsallis distribution [8], [9]. The latter is expressed as

$$p(x) = \frac{1}{Z_q}[1 + \beta(q-1)x^2]^{1/(1-q)} \tag{5}$$

where $1 < q < 3$, and Z_q is a normalization constant. The parameter β is related to the variance of x, and the parameter q quantifies the departure of $p(x)$ from the Gaussian distribution. When $q \to 1$, $p(x)$ reduces to the Gaussian distribution; when $q = 2$, $p(x)$ is the Cauchy distribution. When $5/3 < q < 3$, the distribution is heavy-tailed, $P(X \geq x) \sim x^{-b}$, where $0 < b = 2/(q-1) - 1 < 2$. The most well-known example of heavy-tailed distribution is the Pareto distribution, defined by $P(X \geq x) = (c/x)^b$, where $x \geq c > 0$. Using the definition for the nth moment

$$\overline{x^n} = \int x^n f(x) dx = -\int x^n dP(X \geq x)$$

where $f(x)$ denotes probability density function (PDF), it is easy to verify that when $0 < b < 2$, the second and all higher order moments for Pareto and heavy-tailed distributions are infinite (when $0 < b \leq 1$, the mean is also infinite).

The Tsallis distribution has been considered to provide a statistical mechanical foundation for α-stable laws [8], [9]. A random variable X is called (strictly) stable if the distribution for $\sum_{i=1}^{n} X_i$ is the same as that for $n^{1/\alpha} X$

$$\sum_{i=1}^{n} X_i \stackrel{d}{=} n^{1/\alpha} X \tag{6}$$

where $\stackrel{d}{=}$ denotes equal in distribution, and $X_1, X_2, \ldots$ are independent random variables, each having the same distribution as X. From (6), one readily finds that $n\mathrm{Var}X = n^{2/\alpha}\mathrm{Var}X$. One can then prove that $0 < \alpha \leq 2$ [10]. When $\alpha = 2$, the corresponding stable law is the Gaussian distribution. By the central limit theorem, it can be considered as an "attractor": so long as independent identically distributed (iid) random variables $X_1, X_2, \ldots, X_n$ have finite variance, their summation converges to the Gaussian distribution. When $0 < \alpha < 2$, an α-stable law is characterized by a heavy tail with infinite variance. Summation of iid random variables with infinite variance will converge to an α-stable law. This is the generalized central limit theorem [10]. Note that besides the Gaussian distribution, there are only two more α-stable laws, corresponding to $\alpha = 1/2$ and 1, that have closed form PDFs. They are the Levy and Cauchy distributions, respectively [10]. Although most α-stable laws do not have closed-form formulas for the PDFs, they can be conveniently represented by characteristic functions. For more details, we refer to [10, Chap. 7].

We have fitted the Tsallis distribution to the differentiated data of all the 280 sea clutter data measured under various sea and weather conditions. Fig. 3(a) and (b) shows typical results of using the Tsallis distribution to fit the differentiated data of sea clutter data without and with target, respectively. We observe that the Tsallis distribution fits the differentiated data very well.

To better appreciate how effective the Tsallis distribution is for modeling sea clutter, we compare the fitting by the Tsallis distribution with other commonly used distributions for sea clutter. Conventionally, distributional analysis of sea clutter is performed on the original amplitude data, not on the differentiated data. Since the K distribution is one of the best distributions found for sea clutter, for illustrative purpose, we show the fitting results for the K distribution. Indeed, we have found that other commonly used distributions, such as Weibull, log-normal, and compound-Gaussian distributions, do not yield better fitting than the K distribution. The K distribution is expressed as [16]

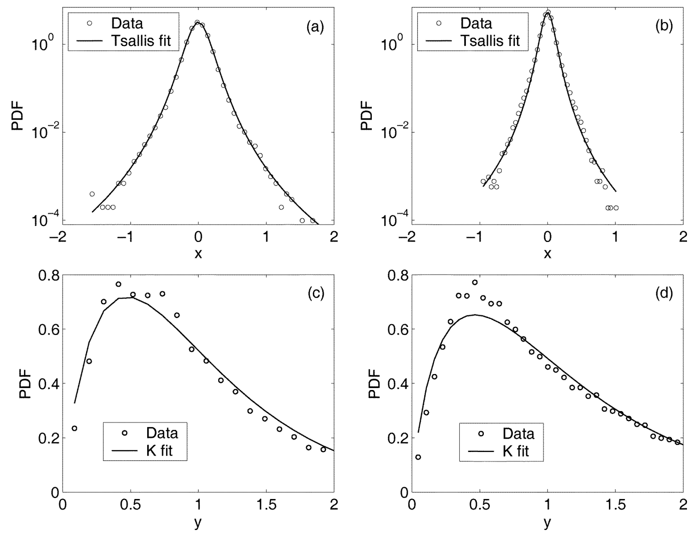

Fig. 3. Representative results of using (a,b) Tsallis distribution to fit the differentiated data and (c,d) K distribution to fit the amplitude data of sea clutter. (a,c) are for the sea clutter data without target, while (b,d) are for the data with target. Circles and solid lines denote the raw and fitted probability density functions, respectively. The estimated parameters are (a) $(q, \beta) = (1.34, 43.14)$, (b) $(q, \beta) = (1.51, 147.06)$, (c) $(\nu, \mu) = (1.30, 0.86)$, and (d) $(\nu, \mu) = (1.20, 0.96)$.

$$f(x) = \frac{\sqrt{2\nu}}{\sqrt{\mu}\Gamma(\nu)2^{\nu-1}} \left(\sqrt{\frac{2\nu}{\mu}}x\right)^{\nu} K_{\nu-1}\left(\sqrt{\frac{2\nu}{\mu}}x\right), \quad x \geq 0 \tag{7}$$

where ν and μ are parameters (μ is in fact equal to half of the second moment), $\Gamma(\nu)$ is the usual gamma function, and $K_{\nu-1}$ is the modified Bessel function of the third kind of order $\nu - 1$. Fig. 3(c) and (d) shows the results of using the K distribution to fit the same amplitude data that were used to generate Fig. 3(a) and (b). We observe that the fitting to these data using the K distribution is worse than that using the Tsallis distribution.

To systematically compare the fitting using the Tsallis and K distributions, we have quantified the goodness-of-fit by the Kolmogorov–Smirnov (K-S) test, which can be expressed as [17]

$$D_e = \max_{1 \leq i \leq N} |F(x_i) - F_t(x_i)| \tag{8}$$

where $F(x_i)$ and $F_t(x_i)$ are the empirical and theoretical cumulative distribution function (CDF), respectively, and N is the total number of data points, determined after the sample data set has been sorted into an ascending order without duplicates. When D_e is smaller than certain threshold value, the data are often said to have passed the goodness-of-fit test, and they can be considered to arise from the specified distribution. Indeed,

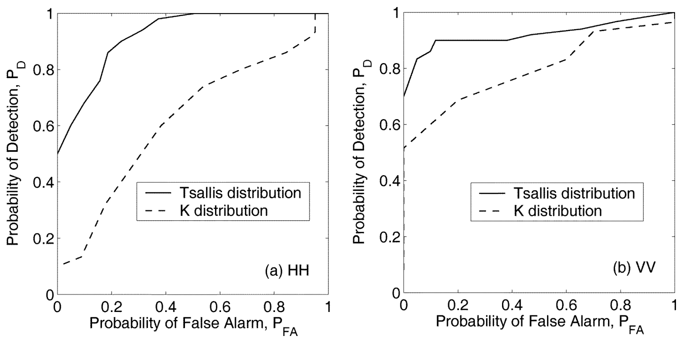

Fig. 4. Receiver operation characteristic curves of the parameter β from the Tsallis distribution (solid lines) and of the parameter ν from the $\mathbb{K}$ distribution (dash lines) for detecting target among (a) HH and (b) VV datasets.

we have found that D_e is much smaller for Tsallis distribution than for the $\mathbb{K}$ distribution.

To further appreciate the advantage of fitting the sea clutter data by the Tsallis distribution, we now consider target detection within sea clutter by using the parameters of the Tsallis and $\mathbb{K}$ distributions. For illustrative purpose, we consider the parameter β from the Tsallis distribution and the parameter ν from the $\mathbb{K}$ distribution. Fig. 4(a) and (b) shows the receiver operation characteristic (ROC) curves for detecting target among the HH and VV datasets, respectively, where the solid curves are for the Tsallis distribution, and the dash ones are for the $\mathbb{K}$ distribution. In a binary detection problem (i.e., whether a range bin hit a target or not), if one decides one hypothesis (say, H_1, with a target) against the other (say, H_0, without a target) based on whether the parameter (β from the Tsallis distribution and μ from the $\mathbb{K}$ distribution) is larger or smaller than a threshold value, one obtains the correct probability of detection and probability of false alarm. A ROC curve corresponds to a curve when the threshold is continuously varied. One ROC curve is better than the other if one is on top of the other. Since the ROC curve for the Tsallis distribution is always on top of that for the $\mathbb{K}$ distribution, we can conclude that β from the Tsallis distribution is more accurate than ν from the $\mathbb{K}$ distribution in separating sea clutter data with and without target.

IV. Concluding Remarks

We have proposed to use the Tsallis distribution to model the nonstationary sea clutter amplitude data. By systematically analyzing the differentiated data of 280 sea clutter time series measured under various sea and weather conditions, we have shown that the Tsallis distribution fits sea clutter data much better than commonly used distributions for sea clutter, such as the $\mathbb{K}$ distribution. In particular, the parameter q estimated from the Tsallis distribution for all the sea clutter data ranges from 1.10 to 1.82 and, hence, is quite different from 1. This signifies that the differentiated data of sea clutter is non-Gaussian.

To understand why the Tsallis distribution provides a better fit to the sea clutter data, we reemphasize two sources of complexity for sea clutter: the roughness of sea surface, largely due to wave-turbulence interactions on the surface and ocean sprays, and the multipath propagation. Either source of complexity suggests that sea clutter may be considered as a superposition of signals massively reflected from ocean surface. Therefore, the central limit theorem or the generalized central limit theorem has to play a crucial role. Consequentially, the Tsallis distribution can be expected to fit sea clutter data well.

Finally, we note that the information provided by the Tsallis distribution analysis and the multifractal analysis [6] is complementary, since distributional analysis does not consider correlations, while multifractal analysis considers correlations but not the distributions. Therefore, it may be interesting to integrate the two methods to improve accuracy for target detection within sea clutter.

Acknowledgment

The authors would like to thank three anonymous reviewers and the associate editor for many stimulating comments.

References

[1] F. A. Fay, J. Clarke, and R. S. Peters, "Weibull distribution applied to sea-clutter," in *Proc. IEE Conf. Radar'77*, London, U.K., 1977, pp. 101–103.
[2] F. E. Nathanson, *Radar Design Principles*. New York: McGraw-Hill, 1969, pp. 254–256.
[3] E. Jakeman and P. N. Pusey, "A model for non-Rayleigh sea echo," *IEEE Trans. Antennas Propag.*, vol. AP-24, no. 6, pp. 806–814, Nov. 1976.
[4] S. Sayama and M. Sekine, "Log-normal, log-Weibull and $\mathbb{K}$-distributed sea clutter," *IEICE Trans. Commun.*, vol. E85-B, pp. 1375–1381, 2002.
[5] F. Gini, A. Farina, and M. Montanari, "Vector subspace detection in compound-Gaussian clutter, Part II: Performance analysis," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 38, no. 4, pp. 1312–1323, Oct. 2002.
[6] J. Hu, W. W. Tung, and J. B. Gao, "Detection of low observable targets within sea clutter by structure function based multifractal analysis," *IEEE Trans. Antennas Propag.*, vol. 54, no. 1, pp. 136–143, Jan. 2006.
[7] K. S. Shanmugan and A. M. Breipohl, *Random Signals: Detection, Estimation and Data Analysis*. New York: Wiley, 1988, pp. 587–590.
[8] C. Tsallis, S. V. F. Levy, A. M. C. Souza, and R. Maynard, "Statistical-mechanical foundation of the ubiquity of Levy distributions in nature," *Phys. Rev. Lett.*, vol. 75, pp. 3589–3593, 1995.
[9] C. Beck, G. S. Lewis, and H. L. Swinney, "Measuring nonextensitivity parameters in a turbulent Couette-Taylor flow," *Phys. Rev. E*, vol. 63, p. 035303(R), 2001.
[10] J. B. Gao, Y. H. Cao, W. W. Tung, and J. Hu, *Multiscale Analysis of Complex Time Series — Integration of Chaos and Random Fractal Theory, and Beyond*. New York: Wiley, 2007.
[11] J. Hu and J. B. Gao, "Modeling sea clutter as a nonstationary and nonextensive random process," in *Proc. IEEE 2006 Int. Radar Conf.*, Verona, NY, Apr. 24–27, 2006.
[12] J. Hu, J. B. Gao, F. L. Posner, Y. Zheng, and W. W. Tung, "Target detection within sea clutter: A comparative study by fractal scaling analyses," *Fractals*, vol. 14, pp. 187–204, 2006.
[13] J. B. Gao, J. Hu, W. W. Tung, Y. H. Cao, N. Sarshar, and V. P. Roychowdhury, "Assessment of long range correlation in time series: How to avoid pitfalls," *Phys. Rev. E*, vol. 73, p. 016117, 2006.
[14] G. S. Maddala and S. W. Wu, "A comparative study of unit root tests with panel data and a new simple test," *Oxford Bull. Econ. Statist.* (special issue), pp. 631–652, 1999.
[15] E. Said and D. A. Dickey, "Testing for unit roots in autoregressive moving average models of unknown order," *Biometrika*, vol. 71, pp. 599–607, 1984.
[16] A. Stuart and J. K. Ord, *Kendall's Advanced Theory of Statistics*. London, U.K.: Griffin, 1987, vol. 1.
[17] Y. H. Zhou, J. B. Gao, K. D. White, I. Merk, and K. Yao, "Perceptual dominance time distributions in multistable visual perception," *Biol. Cybern.*, vol. 90, pp. 256–263, 2004.