Statistical Analyses of Measured Radar Ground Clutter Data

J. B. BILLINGSLEY, Member, IEEE
MIT Lincoln Laboratory

A. FARINA, Senior Member, IEEE
Alenia Marconi Systems

F. GINI, Member, IEEE

M. V. GRECO

L. VERRAZZANI
University of Pisa

The performance of ground-based surveillance radars strongly depends on the distribution and spectral characteristics of ground clutter. To design signal processing algorithms that exploit the knowledge of clutter characteristics, a preliminary statistical analysis of ground-clutter data is necessary. We report the results of a statistical analysis of X-band ground-clutter data from the MIT Lincoln Laboratory Phase One program. Data non-Gaussianity of the in-phase and quadrature components was revealed, first by means of histogram and moments analysis, and then by means of a Gaussianity test based on cumulants of order higher than the second; to this purpose parametric autoregressive (AR) modeling of the clutter process was developed. The test is computationally attractive and has constant false alarm rate (CFAR). Incoherent analysis has also been carried out by checking the fitting to Rayleigh, Weibull, log-normal, and K-distribution models. Finally, a new modified Kolmogorov–Smirnoff (KS) goodness-of-fit test is proposed; this modified test guarantees good fitting in the distribution tails, which is of fundamental importance for a correct design of CFAR processors.

Manuscript received August 1, 1997; revised March 10 and July 9, 1998.

IEEE Log No. T-AES/35/2/04309.

Authors' addresses: J. B. Billingsley, Lincoln Laboratory, Massachusetts Institute of Technology, 244 Wood St., Lexington, MA 02420-9108, E-mail: (billingsley@ll.mit.edu); A. Farina, System Analysis & Scientific Calculus Group-System Area, Alenia Marconi Systems, Via Tiburtina Km 12.4, 00131 Rome, Italy, E-mail: (farina@alenia.finmeccanica.it); F. Gini, M. V. Greco and L. Verrazzani, Dept. of Information Engineering, University of Pisa, Via Diotisalvi 2, 56126 Pisa, Italy, E-mail: (gini@iet.unipi.it, greco@iet.unipi.it).

I. INTRODUCTION

For many years, in radars with low resolution capabilities, the clutter echoes were considered as having a Gaussian probability density function (pdf); see for example [16, sect. 10.2]. In modern radar systems, operating at low grazing angles or with resolution capabilities high enough to resolve the structure of sea or ground surfaces, the statistics of the clutter have been observed to deviate from Gaussianity [1, 2, 7, 8, 12, 19, 20]. The clutter is spikier than if it were Gaussian, and the spikes are processed by the radar detector as targets, with increased false-alarm rate (FAR). The understanding of clutter behavior and the modeling of the non-Gaussian clutter, both in the spectrum and in the distribution, are problems of fundamental interest in the radar community for successful radar design and performance prediction. The spectral characteristics of clutter determine, for instance, the design of moving-target-indicator (MTI) processors [5, 10], and the amplitude statistics of clutter affect the design of constant false-alarm rate (CFAR) processors (see e.g., [16, p. 306; 27]).

In particular, the amplitude statistics resulting from the cell-to-cell spatial variation of clutter (the main subject of this paper) are highly important in clutter modeling for determining signal-to-clutter power ratios at inputs to clutter cancellation processors, so as to determine whether or not detection and tracking can occur in the clutter residues left after clutter cancellation.

This paper reports statistical analyses performed on experimental X-band ground-clutter data in open agricultural terrain. The data were recorded at the Wolseley, Saskatchewan, site with the MIT Lincoln Laboratory Phase One radar.

Our main aim is to test the theoretical models so far proposed in the literature, and to verify the deviations of the recorded X-band data from Gaussianity.

In the rest of this section, the experimental set-up is briefly described. The next section reports the analysis of the in-phase (I) and quadrature (Q) components in terms of histograms, and third- and fourth-order zero-lag cumulants (called, respectively, *skewness* and *kurtosis*). Azimuth and range spectral analysis is the topic of Section III. In parallel to classical nonparametric FFT-based (fast Fourier transform) estimation, parametric estimation based on autoregressive (AR) modeling is also carried out because it is needed for the novel Gaussianity test employed in Section IV. This test, which is based on cumulants of order higher than the second, applies to stochastic processes irrespective of the temporal or spatial nature of the data; it was proposed and applied to seismic data by Giannakis and Tsatsanis in [14]. We have applied it for the first time to radar

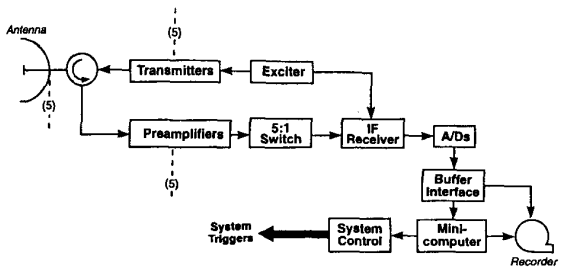


Fig. 1(a). Phase One instrument block diagram.

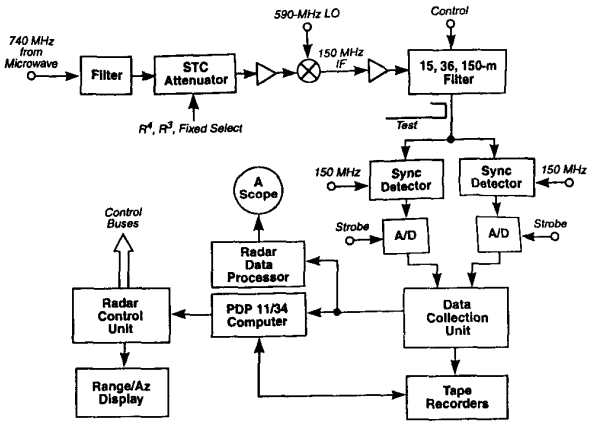


Fig. 1(b). Phase One receiver and signal processor.

clutter data in order to test for the Gaussianity of the I and Q samples. The results of this test are in perfect agreement with those presented in Section II. In Section V, the statistical analysis of the ground clutter is completed with the incoherent analysis of the data amplitude. Incoherent analysis has also been carried out by checking the fitting to Rayleigh, Weibull, log-normal, and K-distribution models. In particular, a modified version of the Kolmogorov–Smirnoff (KS) test is proposed that overcomes the limitation of classical KS and chi-squared tests when applied to radar clutter data. As already noted in [3 and 7], the classical KS test is not able to distinguish among the Rayleigh, Weibull, log-normal and K-distribution models when we are interested in obtaining a good fit to the distribution tails for correct design of CFAR processors. Some conclusions are reported in Section VI.

A. Measurement Instrumentation

The Phase One radar could operate in any one of five different radar bands (VHF, UHF, L-, S-, and X-bands). The results of this work are for ground clutter data recorded at X-band. Results in all bands are given in [2–4]. The overall radar system block diagram is shown in Fig. 1(a). The system exciter supplied all transmit and receive local oscillator (LO) frequencies and provided the basic timing reference for the system. The basic frequency reference for the exciter was a HP8662A synthesizer signal generator, which had sufficient stability to support an overall clutter improvement factor of 60 dB. There were five transmitters in the system. The X-band transmitter had a traveling wave tube (TWT) as the final high-power output stage. The signals from each of the five high-power transmitters were fed through their respective circulators to transmission lines. The X-band signal was transmitted to its antenna in a separate waveguide. The waveguide run consisted of 3.66 m sections that were connected together during tower erection.

The signals received from the antennas were fed to a separate preamplifier for each frequency band. The first IF for all frequency bands was 740 MHz. The first IF entered into a common receiver used for all frequencies.

The diagram of the receiver and signal processor is shown in Fig. 1(b). The receiver IF gain could be varied dynamically according to R^3 or R^4 sensitivity time control (STC) function. Both STC functions provided 40 dB attenuation at 1 Km. In addition, the preamplifiers (shown in Fig. 1(a)) could be bypassed, and fixed attenuation could be switched into the system to avoid system saturation by large target returns.

The 740 MHz IF signal was mixed to 150 MHz, where the matched filtering was accomplished. Additional IF amplification was also provided. The 150 MHz IF signal was converted to I and Q signals at baseband where the 13 bit analog-to-digital (A/D) conversion was performed. The Data Collection Unit provided real-time buffering of the I and Q data and routed the clutter data to the PDP-11/34 computer or tape recorder. The Radar Data Processor provided a real-time A-scope display as shown. Also provided was a nonreal-time range/azimuth display. The Radar Control Unit consisted of timing and data control logic that accepted commands from the computer and converted them to radar control signals.

Data collection was controlled automatically by the computer. The system operator input parameters via keyboard or floppy disk, which established the range and azimuth limits, frequency, scan rate, waveform, sampling rate, and polarization to be used for each experiment.

The instrument maintained coherence and stability sufficient for 60 dB, two-pulse-canceller clutter attenuation in post-processing; it had uncoded pulsed waveforms, with two pulse lengths available in each band to provide high and low range resolutions. Polarization was selectable as vertical or horizontal, with transmit and receive antennas always copolarized, i.e., the cross-polarized component in the radar return signal could not be received. Frequency, polarization, and pulse length as well as spatial extent in range and

TABLE I
Parameters of Phase One Radar

Frequency Band (MHz)	VHF 165	UHF 435	L-Band 1230	S-Band 3240	X-Band 9200
Polarization (TX/RX)	VV or HH				
Range Resolution	150, 36, 15 m				
Azimuth Resolution	13°	5°	3°	1°	1°
Peak Power	10 KW (50 KW at X-Band)				
Antenna Control	Step or Scan through Azimuth Sector				
Tower Height	18.3 or 30.5 m				
10 Km Sensitivity	normalized RCS = −60 dB				
Amount of Data	25 Tapes/Site				
Acquisition Time	2 Weeks/Site				

azimuth of the recording window, number of pulses, and pulse-repetition rate were selectable by computer console for each recorded clutter experiment.

The instrument was self-contained and mobile on truck platforms. Antennas were mounted on erectable towers and had relatively wide elevation beams that were fixed horizontally at 0° depression angle. The radar was internally calibrated for every clutter measurement and externally calibrated at almost every site, using standard-gain antennas and corner reflectors mounted on portable towers. The characteristics of the Phase One radar are reported in Table I.

B. Data Acquisition

The analyzed clutter data files were recorded at X-band at Wolseley, Saskatchewan, located on the Canadian prairie at a latitude of 50.36° N and a longitude of 103.15° W. The illuminated area was covered by agricultural crops (83%), deciduous trees (11%), lakes (4%), and rural farm buildings (2%). The terrain was a sequence of gentle slopes ($< 1°, < 2°$) with a relief of 25–150 ft.

The analyzed X-band (9.2 GHz) data are divided in two different sets. In the first one (N007001.35) the polarization is HH, in the second one (N007001.34) it is VV. Each set contains four range intervals recorded in scan mode, and each range interval contains 316 range cells. The scan velocity was 2°/s, i.e., about 2 beams/s. The emitted pulse repetition frequency (PRF) was 500 Hz, but only 1 out of 2 pulse returns was recorded. The pulses were further coherently integrated in groups of 16, so the effective PRF of the data from this experiment is 15.625 Hz.

The data were stored in a 316×703 matrix, each row for a fixed range, each column for a fixed azimuth. For each integrated pulse, 316 range samples are provided at 10 MHz sampling rate. The data were collected one range interval after another in "windshield-wiper" mode. In these Wolseley data, for the first range interval the antenna scanned 270° to 360° (from the north) while 316 range samples were collected covering about 1 to 5.7 km for each of 703 pulses. For the second interval, the antenna scanned back from 360° to 270° while 316 range samples were now collected covering about 5.7 to 10.5 km. For the third range interval, the antenna scanned 270° to 360° while data from about 10.5 to 15.2 km were collected, and for the fourth range interval the antenna scanned 360° to 270° while data from about 15.2 to 20.0 km were collected.

The radar depression angle decreased from about 1.5° to 0.7° across the first range interval, and further decreased to $\sim 0.5°$ in the second interval, to $\sim 0.3°$ in the third interval and to $\sim 0.2°$ in the fourth range interval. The azimuth beamwidth of the antenna was 0.018 rad at HH polarization and 0.019 rad about 1° (i.e., at VV polarization). The nominal azimuth sampling interval was $2/15.625 = 0.128°$/pulse; thus for each range cell the radar recorded seven to eight azimuth samples. The range resolution was 15 m (i.e., pulse length = 100 ns, which matches the 10 MHz sampling rate). The return from each pulse was provided in I and Q format calibrated in units of (radar cross section)$^{1/2}$ [2].

The analysis was performed on the data after normalizing them with respect to the square root of the cell area. This normalization is helpful in comparing results for statistical populations extended in range, as with the Wolseley data. The cell area is $A = D\,\Delta D\,\Delta\theta$, where D is the distance of the illuminated area from the radar, ΔD is the range resolution and $\Delta\theta$ is the azimuth resolution. The normalization makes the amplitude values of clutter returns independent of the distance.

In Fig. 2(a) the 2D clutter map relative to the first range interval of the VV polarized data is reported. The data are plotted in logarithmic scale to span the grey-scale code (0–255). The black areas of the image indicate regions of high radar reflectivity, usually from discrete vertical clutter sources, such as buildings, fence lines, trees, and bushes, aligned along roads and field borders; the white areas indicate regions of relatively low reflectivity, such as field surfaces. On the x-axis the 316 range samples are reported from 1 to 5.7 km, whereas on the y-axis the 703 azimuth samples are from 270° to 360°. Fig. 2(b) shows the same data of Fig. 2(a), but in a 3D format to highlight the presence of spikes, well evident in this range interval.

Wolseley is a generally open farmland site of low relief. Good visibility of the land surface is provided from the radar antenna of all four range intervals for which clutter data are reduced and results are discussed in this work. At open farmland sites like Wolseley, spatial clutter statistics are dominated by spatially localized, fixed, discrete scatterers that comprise all the vertical objects that occur on the landscape. These include the following: isolated trees and small clusters of trees; agricultural field boundaries and the vertical objects along them such

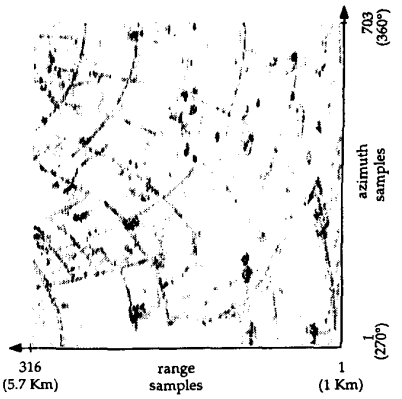


Fig. 2(a). 2D clutter map of 1st range interval, VV polarization.

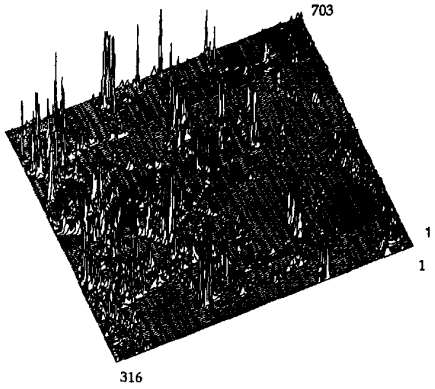


Fig. 2(b). 3D clutter map of 1st range interval, VV polarization.

as fences and higher uncultivated vegetation (tall grass, shrubs); roads and the vertical objects along them such as utility poles and wires; banks of streams and creeks; complexes of farm buildings and farm machinery; other cultural artifacts such as water towers and radio towers; and, also, locally high points in the microtopography itself. Such discrete objects cause strong spikes of clutter distributed randomly over the agricultural field surfaces, which themselves constitute a relatively weak, area-extensive, backscattering medium. These spikes of clutter are observed in Fig. 2, where the curvilinear patterns indicate field and road boundaries on a predominantly north–south, east–west grid. Such spikes are of extremely wide variation in amplitude and result in long tails in empirical clutter spatial amplitude distribution applicable to such terrain.

We performed the same analysis that is described in Section IV on temporal data from a different Phase One X-band file relative to range intervals containing windblown trees in contrast to open farmland. A sample of the results is reported in the Appendix in order not to disrupt the flow of the paper. The analysis shows 1) how different spatial and temporal clutter statistics can be from one another, and 2) how the procedures used can generate wholly different results

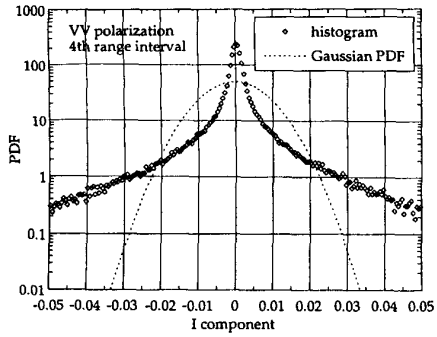


Fig. 3. Histogram of I component, 4th range interval, VV polarization.

and find Gaussian statistics when they exist in the clutter.

II. ANALYSIS OF I AND Q CLUTTER COMPONENTS

The first step of the analysis was to check whether the I and Q components have Gaussian pdfs. To this purpose we drew the histograms of the I and Q components, for each range interval and for each polarization (VV and HH). The histogram (also called the empirical pdf) was compared with the Gaussian pdf having the same mean and variance of the data. The dc offset of each channel was estimated from each of the four range intervals of $316 \times 703 = 222{,}148$ samples, and then subtracted from the data.

This analysis, performed on each range interval, has shown that I and Q pdfs deviate considerably from Gaussianity; the clutter amplitude is therefore not Rayleigh distributed. This deviation is well evident in Fig. 3, where the histogram of the I component for the fourth range interval (VV polarization) is compared with the Gaussian pdf having the same variance and zero-mean.

These results were confirmed by estimating the *skewness* and *kurtosis*, defined, respectively, as

$$\gamma_3^Z \triangleq \frac{E\{(Z-\mu_Z)^3\}}{E\{(Z-\mu_Z)^2\}^{3/2}}, \qquad \gamma_4^Z \triangleq \frac{E\{(Z-\mu_Z)^4\}}{E\{(Z-\mu_Z)^2\}^2} - 3 \tag{1}$$

where $\mu_Z \triangleq E\{Z\}$ is the mean value of the random variable Z. The skewness characterizes the degree of asymmetry of a distribution around its mean value. A positive value of skewness corresponds to a distribution with an asymmetric tail extending on the right of the mean. A negative value of the skewness corresponds to a distribution with an asymmetric tail extending on the left. The kurtosis measures the relative peakedness or flatness of a distribution.

For a Gaussian pdf these two parameters are identically zero, so they are a measure of the deviation from Gaussianity. The estimates of the skewness and kurtosis from two range intervals are reported in Table II. The asymmetry is not significant for either

TABLE II
Standard Deviation, Skewness, Kurtosis

1st Range Interval	HH Polarization I	HH Polarization Q	VV Polarization I	VV Polarization Q
stand. dev.	0.00931	0.00924	0.0110	0.0107
skewness	−0.22284	0.18859	−0.1199	0.2341
kurtosis	260.929	257.923	277.728	271.549
4th Range Interval	**HH Polarization I**	**HH Polarization Q**	**VV Polarization I**	**VV Polarization Q**
stand. dev.	0.0081	0.0086	0.0111	0.0111
skewness	0.0918	0.0522	0.078	−0.173
kurtosis	115.346	111.309	131.982	133.314

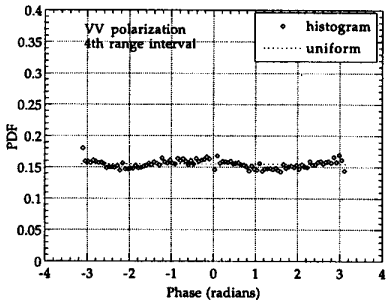


Fig. 4. Phase histogram, 4th range interval, VV polarization.

polarization; $\gamma_3^Z \cong 0$, except for estimation errors. On the contrary, the peakedness is considerable (see also Fig. 3), particularly for the first range interval (e.g., $\gamma_4^Z \cong 261$ for HH data). The results for the I and Q are very similar, as expected because they should have the same statistics.

We also analyzed the phase statistics, i.e., the distribution of $\varphi \triangleq \text{arctg}(Z_Q/Z_I)$. It is generally accepted that the phase statistics of ground clutter are uniform. This uniformity assumption seems to be correct because the phase φ is associated only with the absolute distance d of the ground clutter scatterers from the radar, and with the operational wavelength λ ($\varphi = 4\pi d/\lambda$). It is reasonable to expect that nature does not favor one location of scatterers over another with respect to the radar. However, if care is not taken in the estimation of the phase statistics from I and Q samples, strange results can occur. For example, any dc offsets on the I and Q components will force the point (0,0) to fall away from the center of the square A/D quantization grid causing certain phases to be favored.

Another problem is due to quantization effects. Because the quantization of the A/D converters is independent of the amplitude of the I and Q signals, smaller amplitudes will result in larger quantizations in phase, and larger amplitudes will have a much finer phase quantization. In Fig. 4 a typical phase histogram from VV clutter data is reported. In this case, the phase distribution is almost uniform.

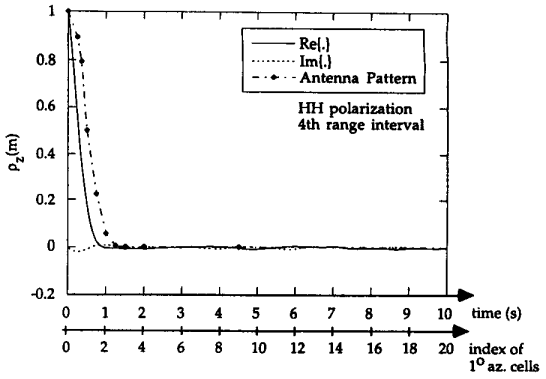


Fig. 5. Azimuthal correlation coefficient, 4th range interval, HH polarization.

III. SPECTRAL ANALYSIS

Spectral properties may have a significant effect on the performance of processing algorithms aimed at suppressing clutter for signal detection, especially in terms of FAR behavior [10, 27]. Hence, to characterize the coherent clutter process further we investigated the spectral properties of the disturbance. In this section we present the results of the analysis performed on the VV and HH data.

A. Azimuth Spectral Analysis

To estimate the azimuth power spectral density (PSD) of the clutter we first used the nonparametric approach based on the Fourier transform (FT) of the autocorrelation sequence. The autocorrelation sequence was estimated from the data without making any assumption, other than the stationarity, about the structure of the clutter process. The sample estimator processed $M = 316$ records $\{\mathbf{z}_k\}_{k=1}^{M}$, one for each range cell, of $N = 703$ complex samples according to the following algorithm [17, ch. 9]:

$$\hat{R}_z(m) = \frac{1}{NM}\sum_{k=1}^{M}\sum_{n=0}^{N-m-1} z_k(n)z_k^*(n+m)$$

$$= 2\hat{R}_{z_I}(m) + j2\hat{R}_{z_I z_Q}(m) \qquad (2)$$

where $z_I(n)$ and $z_Q(n)$ are the in-phase and quadrature components and $z(n) = z_I(n) + jz_Q(n)$ is the complex envelope of the observed signal. In Fig. 5 the correlation coefficient, $\rho_z(m) \triangleq \hat{R}_z(m)/\hat{R}_z(0)$ is reported for the fourth range interval. It is observed that the signal decorrelates to 0.29 in one 1° beamwidth, and to ~ 0 in four beamwidths. It is worth observing that the imaginary part of $\rho_z(m)$, i.e., the cross-correlation coefficient $\hat{R}_{z_I z_Q}(m)/\hat{R}_z(0)$, is approximately zero. The nonparametric PSD of Fig. 6 was obtained by zero padding $\rho_z(m)$ to 1024 points and then Fourier transforming via FFT. The spectral content ($\sim \pm 3$ dB)

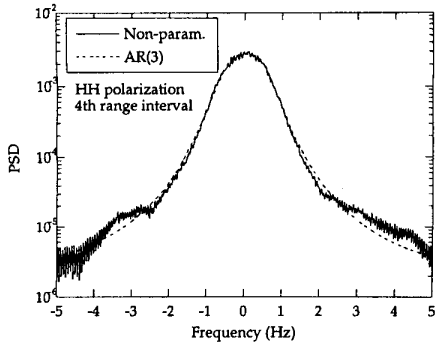


Fig. 6. Azimuthal PSD, 4th range interval, HH polarization.

in Fig. 6 largely falls within 2 Hz (i.e. ±1 Hz), which matches the decorrelation time shown in Fig. 5 to be ~0.5 s.

If a clutter scatterer is an absolutely motionless, large object like a water tower or a feed-storage silo, the spectral density of the fixed-beam temporal return is simply proportional to the absolute square of the FT of the pulse shape, of very narrow spectral width about zero Doppler, where the narrow width is inversely proportional to the interval of observation. We call this a dc return and in fact use such water tower experiments to check system performance in the field. More generally, however, when we look at long sequences of X-band pulses from farmland clutter cells, we almost always see a strong dc component in the spectrum of the return at relatively high levels of spectral power, but spectral broadening due to windblown vegetation at lower levels of spectral power. Such intrinsic-motion spectral spreading is important in quantifying MTI performance against small targets. Quantification of intrinsic-motion spectral spreading is not the subject of this work, but is taken up elsewhere using long-time-dwell experiments with a stationary antenna beam [4], as opposed to the slow-azimuth-scan, short-time-dwell experiments analyzed here to quantify the cell-to-cell spatial statistics of clutter. In these latter experiments, the scanning motion of the antenna results in relative radial velocities of the clutter scatterers with respect to the antenna. The spectral broadening shown in Fig. 6 is due to this antenna radial motion. This broadening, which in totality lies within the interval bounded by ±5 Hz, is wide compared with the width of the dc return (~0.015 Hz), but narrow compared with the intrinsic-motion spectrum (typically, within ±60 Hz at X-band in farmland). If the intrinsic motion spectral broadening could be seen at very low levels (i.e., at levels 60 to 80 dB below the zero-Doppler peak in Fig. 6, it would be observed to be symmetric about zero Doppler because the back-and-forth motion of the vegetation causes symmetric positive and negative Doppler components in the spectrum. The spectrum of Fig. 6 is observed to be symmetric about zero because the rotating antenna also has symmetric components of radial motion towards and away from the clutter cell. In either case, one would not expect asymmetric Doppler shifting in azimuthal samples.

As a comparison, in Fig. 5 the beam pattern of the antenna is also displayed. Fig. 5 shows that, although the azimuthal decorrelation is mainly caused by the antenna pattern, the experimental data cannot be simply predicted from the pattern.

In Fig. 6 we also represent the PSD obtained by a parametric approach based on an AR modeling assumption.[1] The characteristic parameters of the AR model were estimated by the method of linear least squares (LS) [26, p. 564]. The parameters of an AR model of order P, synthetically AR(P), are related to the PSD by the following expression [26, p. 505]:

$$S_z(e^{j2\pi f}) = \frac{|b_0|^2}{|1 + a_1 e^{-j2\pi f} + \cdots + a_p e^{-j2\pi Pf}|^2}. \quad (3)$$

The coefficients $\{a_i\}_{i=0}^{P}$ were estimated from the estimates of the autocorrelation sequence by solving (in the LS sense) the overdetermined Yule–Walker equation:

$$\mathbf{R}_E \mathbf{a} \overset{\text{LS}}{=} \mathbf{0} \quad (4)$$

where

$$\mathbf{R}_E \triangleq \begin{bmatrix} \hat{R}_z(1) & \hat{R}_z(0) & \cdots & \hat{R}_z(-P+1) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{R}_z(L) & \hat{R}_z(L-1) & \cdots & \hat{R}_z(L-P) \end{bmatrix}, \quad \text{and}$$

$$\mathbf{a} \triangleq \begin{bmatrix} 1 \\ \hat{a}_1 \\ \vdots \\ \hat{a}_P \end{bmatrix} \quad (5)$$

with $L > P$. $\mathbf{R}_E$ is an $L \times P$ matrix. Once the $\{a_i\}_{i=1}^{P}$ are estimated, we obtain an estimate of $|b_0|^2$ as $|\hat{b}_0|^2 = 1 + \sum_{m=1}^{P} \hat{R}_z(m)\hat{a}_m$ [26, p. 508]. The L-P extra equations provided in (5) would be redundant in the ideal case. Owing to estimation errors, the extra

[1]The Wierstrass Theorem asserts that continuous PSD can be approximated arbitrarily closely by a rational PSD, i.e., by an autoregressive moving average (ARMA) model, provided that the order is sufficiently large [24, ch. 3], that is, the rational PSDs form a dense set in the class of all continuous spectra. This observation partly motivates the large interest of the radar community on AR models (see e.g., [15, ch. 2] and references therein). The rationale for adopting AR models for the radar echoes is to have a highly parametrized model with a minimum number of parameters that can be easily estimated. Once the model is determined for the data description, we can formulate the target detection problem (which is not tackled in this work) as testing a hypothesis about the model parameters, for which the likelihood test can be constructed; this technique is known as *model-based detection* [15, ch. 2]. Moreover, AR modeling is also necessary here for finding the minimum number of lags necessary to implement the cumulant-based Gaussianity test described in Section IV.

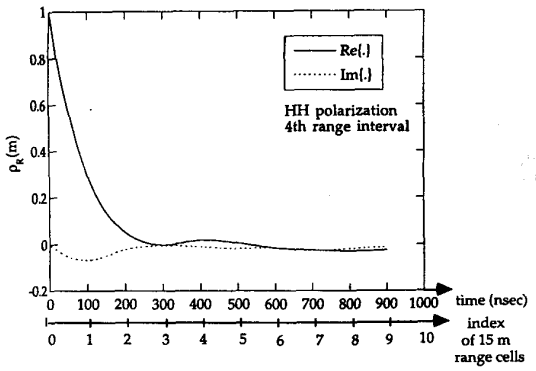


Fig. 7. Range correlation coefficient, 4th range interval, HH polarization.

equations provide information that can lead to a better estimate of the model parameters.

To choose the order P of the AR model, we tried *Akaike's information-theoretic criteria* (AIC) and *Rissanen's minimum description length* (MDL), but neither of these procedures was useful for selecting the model order. AIC and MDL may not be appropriate for this problem because the collected clutter data were only scanned one "look" (or one snapshot), and there is not enough input to use the MDL or the AIC procedure to come up with the order of the AR model. Therefore we fixed $L = 10$ and tried to fit the nonparametrically estimated PSD with AR models of increasing order, starting from $P = 1$. The fit starts to be fairly good for $P = 3$, as shown in Fig. 6 for HH polarization. The results for the other intervals and the other polarization are very similar.

B. Range Spectral Analysis

To estimate the range PSD we performed only a nonparametric analysis. We calculated the correlation for each azimuth cell, and then we averaged the 703 estimates. The correlation coefficient $\rho_R(m)$ is plotted in Fig. 7. Comparing this figure with Fig. 5 we observe that the two decorrelation times are quite different. Along the azimuth direction the coefficient reduces to 0.1 in a few seconds, whereas, along the range, we have the same decreasing in few hundreds of nanoseconds. These apparently large differences are merely the result of the different time-sampling frequencies utilized in range (10 MHz) and in azimuth (15.625 Hz). The PSD has been obtained by zero padding to 1024 points and then Fourier transforming via FFT. The plot is not reported here because it is almost flat, in fact, the correlation time of ~ 100 ns ($\sim$ one pulse length) provides frequency content over the complete ± 5 MHz Nyquist frequency range provided.

The assumption usually made in adaptive radar detection (see, e.g., [15, ch. 3]) of independence of the data from different range cells seems to be reasonable in the Wolseley data.

The important thing to observe in considering Figs. 5–7 is that, owing to the heterogeneity of the spatial scattering ensemble in open farmland terrain (strong discrete sources dispersed over a weakly scattering medium), the returned signal from the scanning antenna largely decorrelates from one spatial cell to the next, whether the variation is in the range direction or in the azimuth direction. Consider first the azimuth variation results of Figs. 5 and 6. The azimuth extent of the spatial resolution cell is determined by the beamwidth of $\sim 1°$. The scan rate is $2°/s$, so the expectation is that the returned signal would decorrelate in ~ 1 beamwidth or ~ 0.5 s. Fig. 5 shows the azimuthal correlation coefficient to drop to 0.29 in 0.5 s (one beamwidth), but to take ~ 2 s (~ 4 beamwidths) to decorrelate to zero. This largely meets our expectation of decorrelating in one azimuthal interval, given that the azimuthal cell specified is the 3 dB beamwidth, with resultant beam overlap between 1° cells (see Fig. 5). The azimuth sampling rate used in data acquisition was 15.625 samples (pulses) per second. That is, in azimuth the sampling time (0.064 s, that is, the reciprocal of 15.62 Hz) is much less than the cell size (~ 0.5 s). The azimuthal PSD in Fig. 6 shows that the correlation time of ~ 0.5 s, i.e., ~ 1 beamwidth, matches the frequency content of $\sim \pm 1$ Hz.

Consider again the range-variation results shown in Fig. 7. The range extent of the spatial resolution cell is determined by the 3 dB pulse length, which is specified to be 100 ns. This is matched to the range sampling rate of 10 MHz (i.e., in range, the sampling interval equals the cell size). Therefore, the expectation is that in range the returned signal would largely decorrelate, sample-to-sample, in ~ 100 ns. Fig. 7 shows the range correlation coefficient to drop to 0.3 in 100 ns (one pulse length), but it takes 300 ns (3 pulse lengths) to decorrelate to zero. Thus the range results of Fig. 7 very closely match the azimuthal results of Fig. 5, in terms of equivalent cell-to-cell decorrelation in azimuth and range.

IV. GAUSSIANITY TEST

A. Mathematical Background

Now we test for the Gaussianity of I and Q components by means of a cumulant-based test. This test is particularly suitable for parametric *autoregressive moving average* (ARMA) models because it employs the minimal number of cumulant lags that uniquely characterize ARMA processes [14]. It is well known that for a zero-mean Gaussian process the cumulants of order $k \geq 3$ are identically

zero [13, 21]. Thus, they can be used to quantify deviations from Gaussianity.

The third- and fourth-order cumulants of a zero-mean, stationary real random process[2] $z(n)$ are defined, respectively, as

$$c_3^z(l_1,l_2) \triangleq E\{z(n)z(n+l_1)z(n+l_2)\} \tag{6}$$

and

$$\begin{aligned} c_4^z(l_1,l_2,l_3) \triangleq & E\{z(n)z(n+l_1)z(n+l_2)z(n+l_3)\} \\ & - c_2^z(l_1)c_2^z(l_2-l_3) - c_2^z(l_2)c_2^z(l_3-l_1) \\ & - c_2^z(l_3)c_2^z(l_2-l_1). \end{aligned} \tag{7}$$

For our zero-mean process, the autocovariance sequence $c_2^z(l)$ can be consistently estimated using the autocorrelation sample estimator:

$$\hat{c}_2^z(l) = \frac{1}{N}\sum_{n=0}^{N-1-l} z(n)z(n+l), \qquad \text{for} \quad l \geq 0 \tag{8}$$

and, similarly, the sample estimates of the third- and fourth-order cumulants are given by

$$\hat{c}_3^z(l_1,l_2) = \frac{1}{N}\sum_{n=0}^{N-1-l_1} z(n)z(n+l_1)z(n+l_2), \qquad \text{for} \quad (l_1,l_2) \in I_3^N \tag{9}$$

and

$$\begin{aligned} \hat{c}_4^z(l_1,l_2,l_3) = & \frac{1}{N}\sum_{n=0}^{N-1-l_1} z(n)z(n+l_1)z(n+l_2)z(n+l_3) \\ & - \hat{c}_2^z(l_1)\hat{c}_2^z(l_2-l_3) - \hat{c}_2^z(l_2)\hat{c}_2^z(l_3-l_1) \\ & - \hat{c}_2^z(l_3)\hat{c}_2^z(l_2-l_1), \\ & \text{for} \quad (l_1,l_2,l_3) \in I_4^N \end{aligned} \tag{10}$$

where N is the sample size. Note that the sample estimates in (8), (9), and (10) need to be computed only over lags in the nonredundant cumulant regions $\{I_k^N;\ k = 2,3,4\}$, defined as [13]:

$$I_k^N \triangleq \{0 \leq l_{k-1} \leq \cdots \leq l_1 < N\}. \tag{11}$$

The remaining lags can be deduced by using the symmetries present in cumulants [13, 14]. If the process is ARMA(P,Q), the number of cumulant lags that characterize the model uniquely is finite. These lags belong to the region:

$$\begin{aligned} I_k(Q+3P) \triangleq & \{|l_1|,|l_2| \leq Q+3P, -P \leq l_i \leq 2P, \\ & i = 3,\ldots,k-1\}. \end{aligned} \tag{12}$$

So, the Gaussianity test we applied to the clutter data employs a finite number (N_c) of lags of c_3^z and c_4^z. If the clutter process is AR(P), the number of $\mathbf{c}_3^z$ lags characterizing the process is $N_c = 3P(3P+1)/2$ [14]. We collected these lags into two $N_c \times 1$ vectors, $\mathbf{c}_3^z$ and $\mathbf{c}_4^z$, respectively. The Gaussianity tests are based on the asymptotic Gaussianity of $\mathbf{c}_3^z$ and $\mathbf{c}_4^z$, and zero mean value. For stationary processes with absolutely summable cumulants [14]:

If

$$\sum_{l_1,\ldots,l_{k-1}=-\infty}^{\infty} |l_i|\,|c_k^z(l_1,\ldots,l_{k-1})| < \infty, \qquad \forall \quad k,i = 1,\ldots,k-1 \tag{13}$$

Then

$$\sqrt{N}(\hat{\mathbf{c}}_3^z - \mathbf{c}_3^z) \underset{N\to\infty}{\overset{\text{distr.}}{\approx}} \mathcal{N}^{(r)}(\mathbf{0},\mathbf{S}_c) \tag{14}$$

where $\mathbf{c}_3^z$ is the theoretical third-order cumulant vector, $\mathbf{c}_3^z = \lim_{N\to\infty} E\{\hat{\mathbf{c}}_3^z\}$, and $\mathbf{S}_c$ is the asymptotic covariance matrix of $\hat{\mathbf{c}}_3^z$:

$$\mathbf{S}_c \triangleq \lim_{N\to\infty} E\{(\hat{\mathbf{c}}_3^z - \mathbf{c}_3^z)(\hat{\mathbf{c}}_3^z - \mathbf{c}_3^z)^{\mathrm{T}}\}N \tag{15}$$

where T stands for transpose; $\mathbf{S}_c$ is an $N_c \times N_c$ dimensional matrix. If $P_c(l,j)$ denotes the $\text{cov}\{\hat{\mathbf{c}}_3^z(l_1,l_2),\hat{\mathbf{c}}_3^z(j_1,j_2)\}$, then the covariance matrix $\mathbf{P}_c \cong N^{-1}\mathbf{S}_c$, for N large enough.

The covariance matrices of $\mathbf{c}_3^z$ and $\mathbf{c}_4^z$ must be estimated from the data. By using independent records of the process, e.g., those coming from different range cells, the covariance matrix of $\mathbf{c}_3^z$ is estimated as

$$\begin{aligned} \hat{P}_c(l,j) = & \frac{1}{R}\sum_{r=1}^{R}[\hat{c}_3^{z(r)}(l_1,l_2) - \bar{c}_3^z(l_1,l_2)] \\ & \times [\hat{c}_3^{z(r)}(j_1,j_2) - \bar{c}_3^z(j_1,j_2)] \end{aligned} \tag{16}$$

where $\bar{c}_3^z(l_1,l_2) \triangleq (1/R)\sum_{r=1}^{R} c_3^{z(r)}(l_1,l_2)$ and R is the number of available records in range. l_1,l_2 are the indexes relative to the lags of the cumulants, l is the index relative to the position held of each cumulant lag in the cumulant vector. The meaning of j_1, j_2 and j is the same. The covariance matrix of $\mathbf{c}_4^z$ is estimated in a similar way.

The Gaussianity test that uses the third-order cumulant is formulated as the following binary hypothesis testing problem:

$$\begin{cases} \mathrm{H}_0 : \hat{\mathbf{c}}_3^z \approx \mathcal{N}^{(r)}(\mathbf{0},N^{-1}\mathbf{S}_c) \\ \mathrm{H}_1 : \hat{\mathbf{c}}_3^z \approx \mathcal{N}^{(r)}(\mathbf{c}_3^z,N^{-1}\mathbf{S}_c) \qquad \text{with} \quad \mathbf{c}_3^z \neq \mathbf{0}. \end{cases} \tag{17}$$

The test based on the fourth-order cumulant is formulated in a similar way. Recalling that kth-order cumulants of Gaussian processes vanish for $k \geq 3$, it is clear that $\mathbf{c}_3^z \neq \mathbf{0}$ implies non-Gaussianity (hypothesis H_1). The Gaussianity test can be addressed using a chi-squared test based on the statistic $d_{G,3}$ defined as follows:

$$d_{G,3} \triangleq N(\hat{\mathbf{c}}_3^z)^{\mathrm{T}}\hat{\mathbf{S}}_c^{-1}\hat{\mathbf{c}}_3^z \tag{18}$$

[2]The Gaussianity test has been applied to the I and Q components; to simplify the notation in this section we use indifferently $z(n)$ in place of $z_{\mathrm{I}}(n)$ and $z_{\mathrm{Q}}(n)$.

where $\hat{\mathbf{S}}_c^{-1}$ is the inverse of the matrix $\hat{\mathbf{S}}_c = N\hat{\mathbf{P}}_c$ (or the pseudoinverse if $\hat{\mathbf{S}}_c$ is rank deficient).

In fact, in [14] it was demonstrated that under H_0 ($\mathbf{c}_3^z = \mathbf{0}$) $d_{G,3}$ converges in distribution to a central chi-squared pdf with Nc degrees of freedom, i.e.:

$$d_{G,3} \underset{N\to\infty}{\overset{\text{distr.}}{\approx}} \chi^2_{N_c}. \tag{19}$$

Therefore, for an α-level of significance, the test in (18) reduces to a chi-squared test:

$$d_{G,3} \underset{H_0}{\overset{H_1}{\gtrless}} t_G = \chi^2_{N_c}(\alpha). \tag{20}$$

The threshold t_G is found using χ^2 tables, after fixing the probability of false alarm:

$$P_{FA} = \alpha \le \Pr\{d_{G,3} \ge \chi^2_{N_c} \mid H_0\}. \tag{21}$$

B. Results on $d_{G,3}$ and $d_{G,4}$

As discussed in Section III, the azimuthal spectrum of the clutter well fits an AR(3) model, both for HH and VV data. We applied the Gaussianity test on the I and Q components of the clutter data assuming $P = 3$ and $Q = 0$, so obtaining $N_c = 45$. We used 316 records (one for each range cell) of 713 azimuthal samples. We repeated our analysis for all the four range intervals.

The $\mathbf{c}_3^z$ lags in $I_3(3P)$ needed for the test were estimated for each row using 703 samples for each range cell as follows:

$$\hat{c}_3^z(l_1,l_2) = \frac{1}{703}\sum_{n=0}^{702-l_1} z(n)z(n+l_1)z(n+l_2), \quad (l_1,l_2) \in I_3(3P). \tag{22}$$

Each row was considered as an independent record, so

$$\hat{P}_c(l,j) = \frac{1}{316}\sum_{r=1}^{316}[\hat{c}_3^{z(r)}(l_1,l_2) - \bar{c}_3^z(l_1,l_2)] \times [\hat{c}_3^{z(r)}(j_1,j_2) - \bar{c}_3^z(j_1,j_2)] \tag{23}$$

where $\bar{c}_3^z(l_1,l_2) = \frac{1}{316}\sum_{r=1}^{316} c_3^{z(r)}(l_1,l_2)$.

The α-level of significance being fixed at 2%, the threshold t_G for a χ^2 of order $N_c = 45$ is almost equal to 55 [25, p. 1237]. Fig. 8, referring to the I component $z_I(n)$, depicts 316 values of the statistic $d_{G,3}$ and the threshold value for the fourth range interval and HH polarization. It is evident that the threshold stays above the statistic in almost every realization. A similar result was obtained for the Q component. This does not necessarily mean that the I and Q components of the land clutter are Gaussian: it is possible that $E\{\hat{c}_3^z\} = 0$ even if $z_I(n)$ (and $z_Q(n)$) is non-Gaussian, it is sufficient that the pdf of $z_I(n)$ (and $z_Q(n)$) is symmetrically distributed around zero.

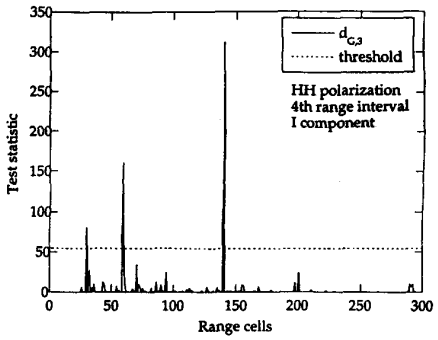


Fig. 8. $d_{G,3}$ statistic for I component (z_I), 4th range interval, HH polarization.

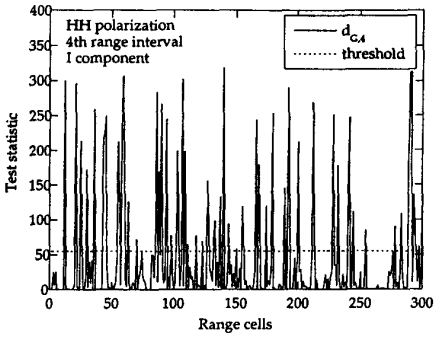


Fig. 9. $d_{G,4}$ statistic for I component (z_I), 4th range interval, HH polarization.

Thus, because H_0 is accepted in the $d_{G,3}$ level, the test was continued by considering also the fourth-order statistic:

$$d_{G,4} \triangleq N(\hat{\mathbf{c}}_4^z)^T \hat{\mathbf{S}}_c^{-1} \hat{\mathbf{c}}_4^z. \tag{24}$$

Fig. 9 depicts the 316 values of the statistic $d_{G,4}$ (I component), relative to the 316 range cells of the fourth range interval, and the threshold. The results relative to the Q component are very similar. Because in this case the threshold is often overcome, the I and Q components are non-Gaussian.

It is worth observing that the order of the AR process can be influenced by the antenna motion that spreads the PSD of the clutter. This means that, without the motion, the order of the AR model could result higher. This is not a problem. In fact, N_c is the minimum number of cumulant lags to check for the Gaussianity, but it needs only a cumulant lag different from zero to verify the non-Gaussianity. So, if the actual order of the AR process had been higher, for example $P = 4$ and so $N_c = 78$, the results would have been the same: the data are non-Gaussian. Summarizing, underestimating the order P does not affect the final result of the test when the data are non-Gaussian.

V. AMPLITUDE PDF ANALYSIS

To complete our study, we performed an incoherent analysis looking for the best fit of the clutter amplitude pdf. Many distributions have been proposed in the literature to model the amplitude pdf of spiky clutter, see for example [1, 2, 6–8, 12, 19, 20, 22]. In this work, we compare the empirical pdf with the K, log-normal (LN) and Weibull pdfs having the same first- and second-order moments, and with the Rayleigh pdf having the same variance. Herein we report the expressions of these pdfs and their moments (see [19, 25, p. 237, and 25, p. 182]); in the following, Z denotes the amplitude, i.e., $Z = |z_I + jz_Q| = \sqrt{z_I^2 + z_Q^2}$.

A. K-Distribution

pdf:

$$p_Z(z) = \frac{\sqrt{2\nu}}{\sqrt{\mu}\Gamma(\nu)2^{\nu-1}}\left(\sqrt{\frac{2\nu}{\mu}}z\right)^{\nu} K_{\nu-1}\left(\sqrt{\frac{2\nu}{\mu}}z\right)u(z) \tag{25}$$

moments:

$$E\{Z^n\} = \frac{(2\mu)^{n/2}\Gamma(\nu + n/2)\Gamma(n/2+1)}{\nu^{n/2}\Gamma(\nu)}, \qquad n = 1,2,\ldots \tag{26}$$

where $u(\cdot)$ is the unit step function, $\Gamma(\cdot)$ is the gamma function, $K_{\nu-1}(\cdot)$ is the modified Bessel function of the third kind of order $\nu - 1$, ν is the *shape parameter*, and μ is the *scale parameter* of the K-distribution.

B. Log-Normal Distribution

pdf:

$$p_Z(z) = \frac{\delta}{z\sqrt{2\pi}}\exp(-\tfrac{1}{2}(\vartheta + \delta\ln(z))^2)u(z) \tag{27}$$

moments:

$$E\{Z^n\} = \exp\left[\frac{n}{\delta}\left(\frac{n}{2\delta} - \vartheta\right)\right], \qquad n = 1,2,\ldots \tag{28}$$

where δ is the *shape parameter* and ϑ is the *scale parameter* of the log-normal distribution.

C. Weibull Distribution

pdf:

$$p_Z(z) = \frac{c}{b}\left(\frac{z}{b}\right)^{c-1}\exp[-(z/b)^c]u(z) \tag{29}$$

moments:

$$E\{Z^n\} = b^n\Gamma(n/c+1), \qquad n = 1,2,\ldots \tag{30}$$

where b is the *scale parameter* and c is the *shape parameter*. The Rayleigh distribution is a particular case of the Weibull distribution for $c = 2$.

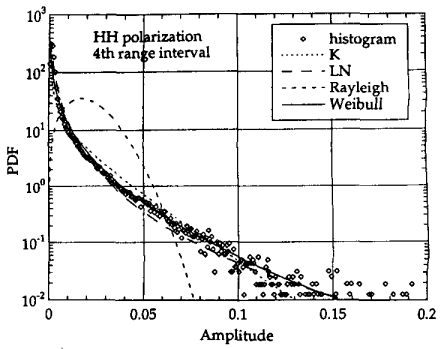


Fig. 10. Amplitude histogram, 4th range interval, HH polarization.

The data seem best to fit the Weibull distribution for the first and second range intervals, while for the third and fourth range intervals the data show a behavior that is intermediate between Weibull and log-normal. The results of histogram analysis for the fourth range cell are reported in Fig. 10 on a log-scale. The same results were obtained for the VV data of the Wolseley site. The characteristic parameters of the theoretical pdfs were estimated by the *method of moments* (MoM) [17, ch. 9], exploiting the following relationships:

$$E\{Z^2\} = 2\mu, \qquad \frac{E\{Z^2\}}{E^2\{Z\}} = \frac{4\nu\Gamma^2(\nu)}{\pi\Gamma^2(\nu+0.5)} \tag{31a}$$

$$E\{Z\} = \exp\left(\frac{1}{2\delta^2} - \frac{\vartheta}{\delta}\right), \qquad \frac{E\{Z^2\}}{E^2\{Z\}} = \exp(1/\delta^2) \tag{31b}$$

$$E\{Z\} = b\Gamma(1/c+1), \qquad \frac{E\{Z^2\}}{E^2\{Z\}} = \frac{\Gamma(2/c+1)}{\Gamma^2(1/c+1)} \tag{31c}$$

for the K, the log-normal, and Weibull distribution, respectively. Table III lists the estimated parameters for each range interval and for both the polarizations.

We observe that the values of the estimated scale and shape parameters for the four different range cells are of the same order of magnitude, but are obviously different for different pdfs.

Then we estimated normalized moments, defined as

$$m_Z(n) \triangleq \frac{E\{Z^n\}}{E^n\{Z\}}. \tag{32}$$

The estimates for the VV and HH polarizations show good agreement with the Weibull distribution for the first and second range intervals, and an intermediate behavior between the Weibull and the log-normal distribution for the third and fourth range intervals. In particular, $\{\hat{m}_Z(n)\}_{n=2}^{6}$ are displayed in Fig. 11 for the fourth range interval and HH polarization.

TABLE III
Scale and Shape Estimated Parameters for K, Log-Normal, Weibull PDFs

Range Interval	K Distribution ν	μ	Log-Normal Distr. δ	ϑ	Weibull Distr. c	b
1st HH	3.68E-2	8.46E-5	0.633	4.33	0.39	1.01E-3
1st VV	4.43E-2	1.25E-4	0.655	4.24	0.41	1.56E-3
2nd HH	5.12E-2	1.58E-4	0.673	4.19	0.43	2.09E-3
2nd VV	4.48E-2	3.02E-4	0.656	3.96	0.41	2.46E-3
3rd HH	5.32E-2	1.85E-4	0.68	4.16	0.43	2.37E-3
3rd VV	4.55E-2	3.71E-4	0.658	3.89	0.41	2.78E-3
4th HH	8.48E-2	7.49E-5	0.749	4.63	0.50	2.54E-3
4th VV	7.07E-2	1.44E-4	0.720	4.32	0.47	2.91E-3

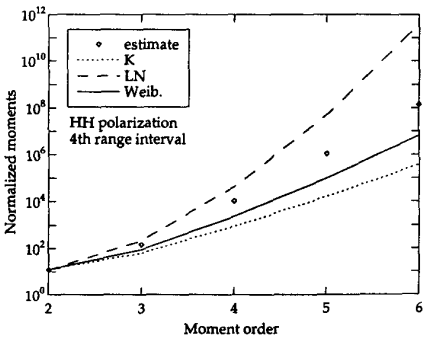


Fig. 11. Normalized moments, 4th range interval, HH polarization.

For comparison, in the same figure, the theoretical normalized moments are reported.

D. Weibull Paper

To verify the good fit of the data to the Weibull distribution, we also used the *Weibull paper* (or *Boothe diagram*). The scale of this paper is obtained by the logarithmic transformations [6]:

$$X = 10\log(Z),$$
$$Y = 10\log\left[-\ln\left(1 - \int_0^Z p_Z(z)dz\right)\right]. \quad (33)$$

If we apply the transformations (33) to the Weibull pdf of (29) we obtain

$$Y = cX - 10c\log(b). \quad (34)$$

Equation (34) represents a line whose slope gives the shape parameter c of the Weibull pdf.

In Fig. 12 we compare the plot relative to the histogram and the straight line, relative to the Weibull pdf with the same mean and mean square value. The straight line obtained by linear LS fitting, [17, p. 520], is also reported; it furnishes the estimates of b and c reported in Table IV.

The MoM and the LS techniques provided very similar values for the parameters b and c; compare

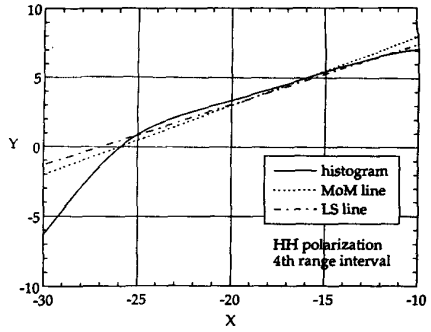


Fig. 12. Weibull paper, 4th range interval, HH polarization.

TABLE IV
b and c Estimates with LS Method

Range Interval	Weibull Param. c	b	RMS Error
1st HH	0.28	5.740E-4	6.28E-2
1st VV	0.32	9.798E-4	3.00E-2
2nd HH	0.33	1.516E-3	2.48E-2
2nd VV	0.41	2.661E-3	8.88E-3
3rd HH	0.44	2.602E-3	3.59E-2
3rd VV	0.47	3.890E-3	5.32E-2
4th HH	0.44	2.015E-3	1.36E-1
4th VV	0.50	3.220E-3	1.33E-1

the sixth and seventh columns of Table III with the second and third columns of Table IV. In Table IV, the rms error of the LS method is also reported. The fitting is good for the first and the second range intervals, for HH and VV data. For the third and fourth range intervals, a nonnegligible deviation is present for small values of X (see Fig. 12). This deviation is due to the presence of radar noise corruption at low signal levels. If the radar did not have this noise limitation, the Weibull fit would be certainly good over the whole X axis.

E. Modified Kolmogorov–Smirnoff Statistical Test

The analysis was concluded by applying a statistical hypothesis test. The KS goodness-of-fit test has been largely used to determine which distribution (Rayleigh, log-normal, Weibull, or K in our case) best fits the data. Unfortunately, in some cases it is not useful, because it places an equal importance on all regions in the probability space. In practical radar applications a good fit is important in the tail regions of the pdfs. The tails, in fact, contain the strong values (i.e., the spikes) that, considered as target returns by a detector, can increase the FAR. Since good fitting in the tails is mandatory for correct design of CFAR processors, especially when low P_{FA} values are required, the KS test is of limited use for clutter data (as recognized also in [3 and 7]). To overcome this problem we propose here a *modified* Kolmogorov–Smirnoff (MKS) goodness-of-fit test. The idea is simple: apply the standard KS test by

taking into account only the tail regions, i.e., by considering only the data above a given threshold λ_{MKS} and the *modified* theoretical pdfs $p_{Z,MKS}(z) = p_Z(z)u(z - \lambda_{MKS})$, where $u(\cdot)$ is the unit step function.

The standard two-sample KS test verifies whether the recorded data are distributed in accordance with a hypothetical pdf. The test is characterized by the parameter α that represents the probability of Type I error, namely the probability of having an error if we reject the null hypothesis H_0 (empirical distribution equal to hypothetical distribution). If this probability is very low, for instance $< 1\%$, then the hypothesis H_0 should be rejected [18].

We first applied the standard two-sample test, which considers the entire definition range of the random variable Z under investigation, and we obtained a probability of Type I error always $< 1\%$, for all the distributions (Rayleigh, log-normal, Weibull, and K); thus, in the classical formulation, the KS test is not useful with these data. This is due to the differences exhibited by the distributions in the region of low values of Z (see Fig. 10). On the contrary, if we apply the MKS test setting the threshold λ_{MKS} in the region (0.01,0.03), where the distributions are very similar (except for the Rayleigh), the result for the parameter α is always 100%. This means that, in this central region, the KS test cannot distinguish between the different proposed models. But, as written above, we are mainly interested in the tails of the distribution, so we applied the MKS test with $\lambda_{MKS} = 0.03$. In this way we left out of the analysis the central region of the pdf, corresponding to high values of P_{FA} (say, $P_{FA} \geq 10^{-2}$). The results are reported in Table V. Generally, the value of α for Weibull pdf is higher than for the other pdfs, so the good fit of our ground clutter data to the Weibull model is confirmed.

TABLE V
MKS Test for Amplitude

Range Interval	α (Type I Error) LN	Weibull	K
1st HH	99%	99%	19%
1st VV	6.4%	96%	4.2%
2nd HH	1.4%	96%	31%
2nd VV	14.6%	99%	3.3%
3rd HH	11%	87%	1.3%
3rd VV	62%	55%	$< 1\%$
4th HH	84%	97%	1.5%
4th VV	94%	79%	$< 1\%$

VI. CONCLUSIONS

The primary objective of our analysis was to quantify the statistics applicable to the cell-to-cell spatial variation of ground clutter returns from open farmland for surface-sited radar. For contrast, some results are also provided in an Appendix involving the temporal variation of clutter returns from fixed cells containing windblown trees.

The results of statistical analysis performed on a ground clutter data set recorded at Wolseley site by the Phase One radar have been reported. The non-Gaussianity of I and Q components of these data has been revealed by means of histogram analysis and a cumulant-based statistical test. The test based on third-order cumulants shows that the I and Q components have a pdf symmetric around zero, which does not mean, however, that they are Gaussian distributed; in fact, the results based on fourth-order cumulants give strong evidence of the non-Gaussianity of the data.

The spectral analysis was performed in both the azimuth and range directions. The study highlighted that, in both range and azimuth, the signal largely decorrelates (~ 0.3) in one spatial cell (either one 15 m range cell, or one azimuthal beamwidth), and completely decorrelates (to zero) in few spatial cells. That the decorrelation time in seconds in the two directions is very different (tens of nanoseconds in range, seconds in azimuth) is simply because the sampling rate in range is fast (10 MHz) and in azimuth is slow (15.625 Hz). The azimuthal PSD was shown to fit an AR(3) model well, and this result was used for the Gaussianity test based on cumulants of order higher than the second. The range PSD is almost flat, and this supports the usual assumption of independence of the data from different range cells.

The analysis was completed by comparing the empirical with some theoretical pdfs (Rayleigh, log-normal, K, and Weibull); the moments estimated up to the sixth order were compared with the theoretical ones. The best fit has been obtained with the Weibull model.

These results were confirmed by means of an MKS statistical test, which we proposed here for fitting the clutter data only on the tails of the distributions. Because the standard KS test shows a probability of Type I error always lower than 1%, for all the polarizations and all the distributions, the KS test in the classical formulation was inefficacious. The differences noted between the two like-polarizations HH and VV in land clutter are negligible, both in the spectrum and in the distribution; this is in contrast to what is observed in the sea clutter [11], where the HH data are spikier than the VV data and a *peak-separation* in the spectra is often evident (up-wind condition).

The interest in radar clutter analysis is due to its importance for successful radar design. The development of statistical models that properly characterize radar clutter processes are required both for optimum detection algorithm design and for performance prediction. The ultimate goal of our analysis is to check through recorded live data the analytical and simulated performance of optimum and suboptimum coherent receivers operating in

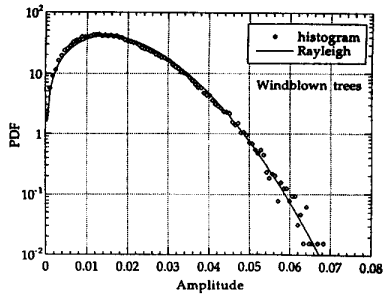


Fig. 13. Amplitude histogram, file H067032.3.

non-Gaussian environment (see, e.g., [28]) and to develop a complete clutter model, applicable to polarimetric radar.

APPENDIX. WINDBLOWN TREES DATA

We performed the same analysis of Section IV on a different Phase One X-band file, namely H067032.3, for which the data were measured from range cells containing windblown trees in contrast to open farmland. A previous analysis [9] demonstrated the data of file H067032.3 to be Gaussian distributed. This data set was recorded at Katahdin Hill site at the Lincoln Laboratory by the same Phase One radar. It contains 30,720 pulses, with a PRF equal to 500 Hz. Data were recorded from 76 contiguous range gates using the Phase One X-band stationary antenna in a fixed azimuth position (235°). The 76 range cells were located from 2.0 km to 3.1 km, covering windblown trees at depression angle $\sim 0.65°$. In Fig. 13 the histogram of the data relative only to the range cells 34–36 is compared with the Rayleigh pdf with the same variance. We can assume the I and Q components to be Gaussian distributed. In Fig. 14 the results relative to the statistic $d_{G,3}$ are reported. It was shown in [4] that the spectral shape from windblown vegetation (the most pervasive source of clutter internal motion) is exponential (see also [5]); in order to implement the cumulant-based Gaussianity test we fitted these data with an AR process. We found that for our purposes a third-order AR model is a good approximation, consequently, the number of third-order cumulant lags was fixed at 45. The statistic values were always below the threshold. This characteristic holds also for $d_{G,4}$, as evident in Fig. 15.

The results obtained for the open farmland data (two data files N007001.34 and N007001.35) are very different from the windblown trees data reported in this Appendix and in [9], where the clutter was found to be approximately Gaussian distributed. This is partly due to the different land covers of the illuminated areas, but also importantly, partly due to the forest data embodying temporal variations, not spatial. The area relative to the H067032.3 data file was homogeneously tree covered, primarily with

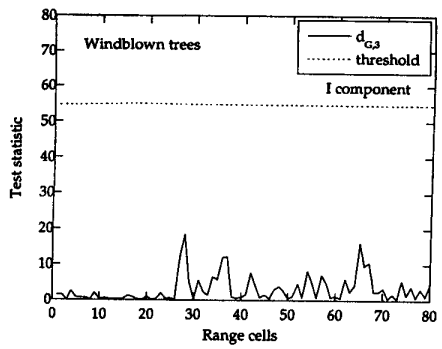


Fig. 14. $d_{G,3}$ statistic for I component (z_I), file H067032.3 (windblown trees).

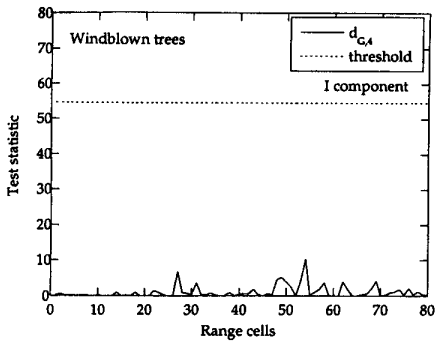


Fig. 15. $d_{G,4}$ statistic for I component (z_I), file H067032.3 (windblown trees).

mixed deciduous trees and with occasional pine and cedar. In the two analyzed data files of this paper, on the other hand, the returns came from a large spatial population of fixed discrete sources on open farmland. This heterogeneity introduces a considerable spread in the distributions as already noted in [2]. The differences are also due to the way of recording the data: the windblown tree data are temporal statistics (variations in time on a given range cell, or on few cells) recorded with fixed antenna; the Wolseley farmland data are spatial statistics, recorded with a scanning antenna on many range cells.

ACKNOWLEDGMENT

We acknowledge the accurate reviewing work of the four referees who provided insightful comments and constructive criticisms that greatly improved the manuscript.

REFERENCES

[1] Anastassopoulos, V., and Lampropoulos, G. A. (1995) High resolution radar clutter classification. In *Proceedings of IEEE International RADAR Conference*, Washington, DC, May 1995, 662–667.

[2] Billingsley, J. B., and Larrabee, J. F. (1991) Multifrequency measurements of radar ground clutter at 42 sites. Technical report 916, MIT, Lincoln Laboratory, Nov. 1991; DTIC AD-A246710.

[3] Billingsley, J. B. (1991)
Radar ground clutter measurements and models Part 1: Spatial amplitude statistics.
In *AGARD Conference Proceedings 501*, Ottawa, Canada, May 1991, 1–15.

[4] Billingsley, J. B. (1996)
Exponential decay in windblown radar clutter Doppler spectra.
Technical report 997, MIT, Lincoln Laboratory, July 1996; DTIC AD-A312399.

[5] Billingsley, J. B., Farina, A., Gini, F., Greco, M. V., and Lombardo, P. (1997)
Impact of experimentally measured Doppler spectrum of ground clutter cancellation on MTI and STAP.
In *Proceedings of 1997 International Radar Conference*, Edinburgh, UK, Oct. 14–16, 1997, 290–294.

[6] Boothe, R. R. (1969)
The Weibull distribution applied to the ground clutter backscattering coefficient.
U.S. Army Missile Command, RE-TR-69-15, ADA691109, 1969.

[7] Chan, H. C. (1990)
Radar sea-clutter at low grazing angles.
IEE Proceedings, Pt. F, **137**, 2 (Apr. 1990), 102–112.

[8] Chen, K. S., and Fung, A. K. (1995)
Frequency dependence of backscattered signals from forest components.
IEE Proceedings, Pt. F, **142**, 6 (Dec. 1995), 310–315.

[9] Farina, A., Gini, F., Greco, M. V., and Verrazzani, L. (1995)
Preliminary analysis of ground clutter data.
Technical report, ETS, Pisa, July 1995.

[10] Farina, A., Gini, F., Greco, M. V., and Lee, P. (1996)
Improvement factor for real sea-clutter Doppler frequency spectra.
IEE Proceedings, Pt. F, **143**, 5 (Oct. 1996), 341–344.

[11] Farina, A., Gini, F., Greco, M. V., and Verrazzani, L. (1997)
High resolution sea clutter data: A statistical analysis of recorded live data.
IEE Proceedings, Pt. F, **144**, 3 (June 1997), 121–130.

[12] Jao, J. K. (1984)
Amplitude distribution of composite terrain radar clutter and the K-distribution.
IEEE Transactions on Antennas Propagation, **AP-32** (Oct. 1984), 1049–1062.

[13] Giannakis, G. B., and Tsatsanis, M. K. (1992)
A unifying maximum-likelihood view of cumulant and polyspectral measures for non-Gaussian signal classification and estimation.
IEEE Transactions on Information Theory, **38**, 2 (Mar. 1992), 386–406.

[14] Giannakis, G. B., and Tsatsanis, M. K. (1994)
Time-domain tests for Gaussianity and time-reversibility.
IEEE Transactions on Signal Processing, **42**, 12 (Dec. 1994), 3460–3472.

[15] Haykin, S., and Steinhardt, A. (1992)
Adaptive Radar Detection and Estimation.
New York: Wiley, 1992.

[16] Helstrom, C. W. (1995)
Elements of Signal Detection and Estimation.
Englewood Cliffs, NJ: Prentice-Hall, 1995.

[17] Kay, S. M. (1993)
Fundamentals of Statistical Signal Processing: Estimation Theory.
Englewood Cliffs, NJ: Prentice-Hall, 1993.

[18] Kim, P. J., and Jennrich, R. I. (1973)
Tables of the exact sampling distribution of the two sample Kolmogorov–Smirnoff criterion $D_{mn}(m < n)$.
In H. L. Harter and D. B. Owen (Eds.), *Selected Tables in Mathematical Statistics*, Vol. 1.
Providence, RI: American Mathematical Society, 1973.

[19] Nohara, T. J., and Haykin, S. (1991)
Canadian East Coast radar trials and the K-distribution.
IEE Proceedings, Pt. F, **138**, 2 (Apr. 1991), 80–88.

[20] Rajalakshmi Menon, K., Balakrishnan, N., Janakiraman, M., and Ramchand, K. (1995)
Characterization of fluctuating statistics of radar clutter for Indian terrain.
IEEE Transactions on Geoscience and Remote Sensing, **33**, 2 (Mar. 1995), 317–323.

[21] Sadler, B. M., Giannakis, G. B., and Lii, K. S. (1994)
Estimation and detection in non-Gaussian noise using higher order statistics.
IEEE Transactions on Signal Processing, **42**, 10 (Oct. 1994), 2729–2741.

[22] Sekine, M., and Mao, Y. (1990)
Weibull Radar Clutter.
London: Peter Peregrinus Ltd., 1990.

[23] Skolnik, M. I. (1990)
Radar Handbook (2nd ed.).
New York: McGraw-Hill, 1990.

[24] Stoica, P., and Moses, R. (1997)
Introduction to Spectral Analysis.
Englewood Cliffs, NJ: Prentice-Hall, 1997.

[25] Stuart, A., and Ord, J. K. (1987)
Kendall's Advanced Theory of Statistics.
London: Griffin, 1987.

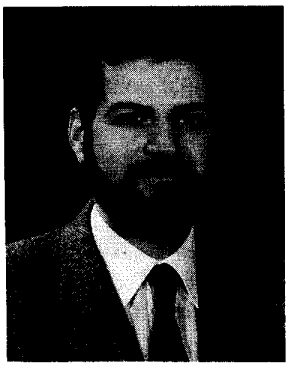
[26] Therrien, C. W. (1992)
Discrete Random Signals and Statistical Signal Processing.
Englewood Cliffs, NJ: Prentice-Hall, 1992.

[27] Watts, S., Baker, C. J., and Ward, K. D. (1990)
Maritime surveillance radar Part 2: Detection performance prediction in sea clutter.
IEE Proceedings, Pt. F, **137**, 2 (Apr. 1990), 63–72.

[28] Sangston, K. J., Gini, F., Greco, M. V., and Farina, A.
Optimum and sub-optimum coherent radar detection in compound-Gaussian clutter.
IEEE Transactions on Aerospace and Electronic Systems, **35**, 2 (July 1999), 445–457.

J. Barrie Billingsley (S'63—M'67) received B.A.Sc. and M.A.Sc. degrees in electrical engineering from the University of Toronto. From 1967 to 1978 he was employed by Calspan Corporation, Buffalo, NY, where much of his work involved analysis of scattering from radar targets. Since 1978 he has been a member of the Technical Staff at M.I.T. Lincoln Laboratory where he is the principal investigator for a large-scale program of radar ground clutter measurements.

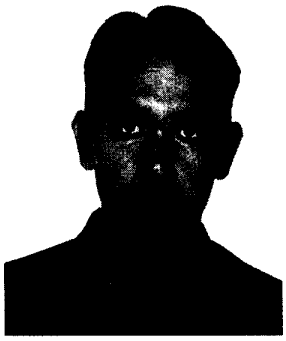
Mr. Billingsley received the 1987 Best Paper Award from the *Journal of Defense Research* for his paper entitled "Ground Clutter Measurements for Air Defense Radars." From 1992 to 1997 he served as U.S. National Leader of TTCP's (The Technical Cooperation Program's) Technical Panel on Radar Clutter Research, and was awarded the 1993 TTCP Achievement Award for work in characterizing the low-angle clutter phenomenon.



Alfonso Farina (M'85—SM'97) received his doctorate in electronic engineering from the University of Rome, Italy, in 1973.

In 1974 he joined Selenia S.p.A. (now part of Alenia Marconi Systems) where he is now a Manager in the Radar and C^2 Division and Head Deputy of the System Analysis Group. He was "Professore Incaricato" of Radar Techniques at the University of Naples since 1979, and in 1985 he was appointed Associate Professor at the same university. His main work interests are in signal, radar data processing techniques for ground-based and AEW radars, ECCM and synthetic aperture radars. He is the author of more than 170 published papers, the co-author with Dr. F. A. Studer of a two-volume book, *Radar Data Processing*, the editor of Optimised Radar Processors, and the author of the recent book by Artech House, *Antenna Based Signal Processing Techniques for Radar Systems*.

Dr. Farina is a Fellow of IEE (UK). He has been Session Chairman at several International Radar Conferences. He received the 1987 Radar System Panel Award of IEEE for development of radar data processing techniques. Also, he is a member of the International Radar System Panel of AES, IEEE.



Fulvio Gini (M'92) was born in Fececchio, Italy, in 1965. He received the Doctor Engineer (cum laude) and the Research Doctor degrees in electronic engineering from the University of Pisa, Italy, in 1990 and 1995, respectively. During his military service, from 1991 to 1992, he joined the "Istituto per le Telecomunicazioni e l'Elettronica" of the Italian Navy, assigned to the Radar division. In 1993 he joined the Department of "Ingegneria dell'Informazione" of the University of Pisa, where he is now a Research Scientist. From July 1996 through January 1997, he was a visiting researcher at the Department of Electrical Engineering, University of Virginia, Charlottesville.

His general interests are in the areas of statistical signal processing, estimation and detection theory. In particular, his research interests include non-Gaussian signal detection and estimation using higher order statistics, cyclostationary signal analysis, and estimation of nonstationary signals, with applications to radar processing and communication. He is the author of over 70 publications.

Maria Valerina Greco was born in Leece, Italy, in 1968. She graduated in electronic engineering in 1993 and received the Ph.D. in 1998, from University of Pisa, Italy. From December 1997 to May 1998 she joined the Georgia Tech Research Institute, Atlanta, USA as a visiting research scholar where she carried on research activity in the field of radar detection in non-Gaussian background.

Her main interests include clutter models, coherent and incoherent detection in non-Gaussian clutter and CFAR techniques.

Lucio Verrazzani received the Dr. Ing. degree (cum laude) in electrical engineering from the University of Pisa, Italy, in 1962 and the "Libera Docenza" in applied electronics from the same University in 1972.

Since 1962 he has worked at the Institute of Electronics and Telecommunications. He is currently a Full Professor of Signal Theory at the Engineering Faculty of Pisa, Italy, and Invited Professor at Accademia Navale of the Italian Navy in Livorno. He is engaged in research on communication systems design, signal processing with application to radar and data transmission, radar target classification and identification. His research interest include non-Gaussian clutter models, deconvolution methods relying on polispectral analysis, and data fusion in multisensor systems.