STATISTICAL ANALYSIS OF MSTAR GROUND DATA

M. Greco Dept. of Information Engineering, University of Pisa, Italy
F. Gini Dept. of Information Engineering, University of Pisa, Italy

1 STATISTICAL ANALYSYS

In this paper, we deal with the problem of modeling SAR clutter data from different vegetated areas. We analyzed MSTAR dataset by means of histogram, moment analysis and covariance estimation*. The real data used throughout this document are high-resolution SAR data in X band, collected in strip map mode during September of 1995 at the Redstone Arsenal, Huntsville, Alabama, by the Sandia National Laboratory (SNL) SAR sensor platform. DARPA and Air Force Research Laboratory jointly sponsored the collection as part of the Moving and Stationary Target Acquisition and Recognition (MSTAR) program. The SAR features are summarized in Table 1.

Data Collectors	Sandia National Lab
Acquisition date	05/09/1995
Site	Huntsville, Alabama (USA)
Sensor name	Twin Otter
Range resolution	0.3047 m
Cross-range resolution	0.3047 m
Range pixel spacing	0.2021048 m
Cross range pixel spacing	0.203125 m
Additive Noise	-32 to -34 dB
Central frequency	9.60 GHz
Bandwidth	0.591 GHz
Dynamic Range	64 dB
Azimuth Beamwidth	8.8°
Elevation Beamwidth	6.8°
Polarization	HH
Bits per pixel	16

Table 1 - MSTAR characteristics

The whole dataset comprises more than 100 files, recorded on urban and rural zones, with different kind of vegetation and percentage of manmade and natural objects. Based on these characteristics, we divided the whole dataset in 7 subsets. In this paper, we report the results of files relating to grass fields and dense vegetated areas (wood and trees). The images of the two analyzed files are shown in Figure 1 and Figure 2. Each image is 1784×1472 pixels large then each file contains 2,626,048 real amplitude data.

* This work has been funded by EOARD contract N6255804P6114

1.1 Amplitude Analysis

The first step of our statistical analysis focused on the validation of theoretical models for the first order distribution of the data. Many distributions have been proposed in the literature to model the amplitude probability density function (PDF) of ground SAR data[1]

In this work, we compare the histogram or empirical pdf of the data with Log-normal (LN), Weibull (W), K, and Generalized K (GK) PDFs. The expressions of these PDFs and their moments are reported in the reference[1].

The characteristic parameters of the theoretical PDFs were estimated by the classical method of moments (MoM)[1,2].

The results of the histogram analysis are reported in Figure 3a for the grass field of the file HB06176 and Figure 3b for the trees. The results in Figure 3a show that, for this file, the PDFs are very similar each other, but for the LN model, and to the shape of the histogram. The analysis of moments shows that GK and Weibull models provide the best fitting. For the trees none of the tested models exhibits a very good fitting as show in Figure 3b, particularly on the tails. In this scenario the clutter is much spikier than for the grass field

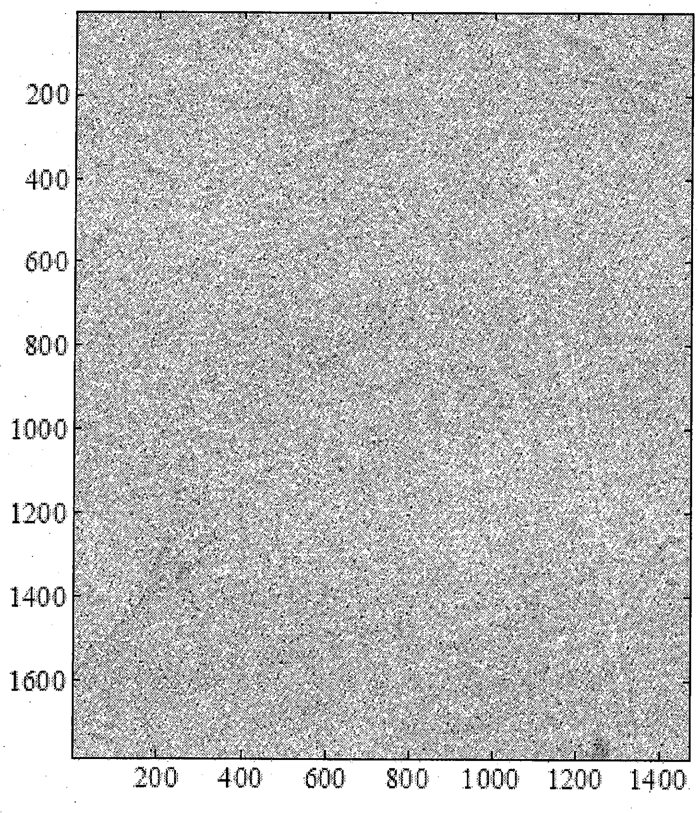


Figure 1 – Grass field, file HB06176

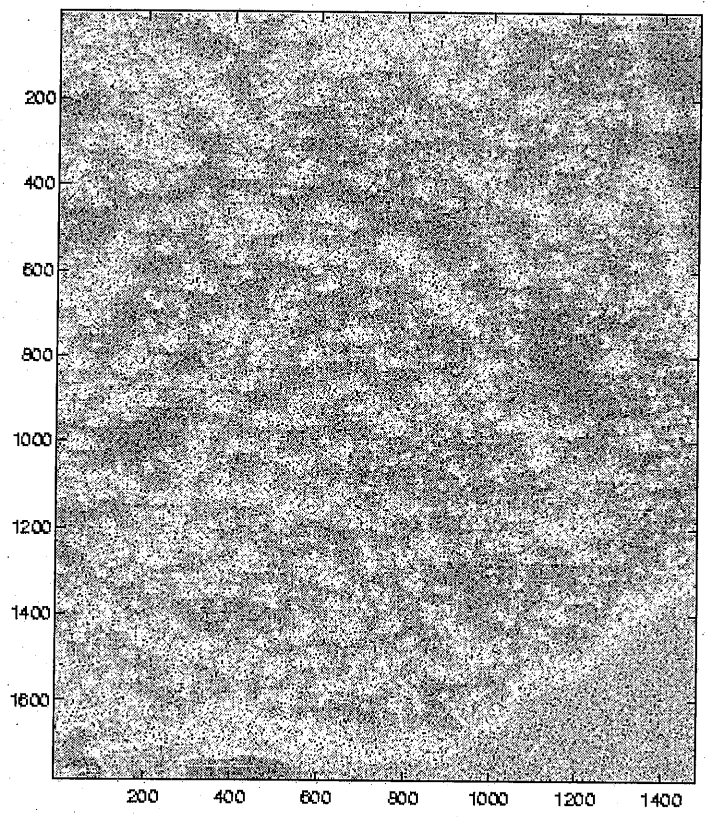
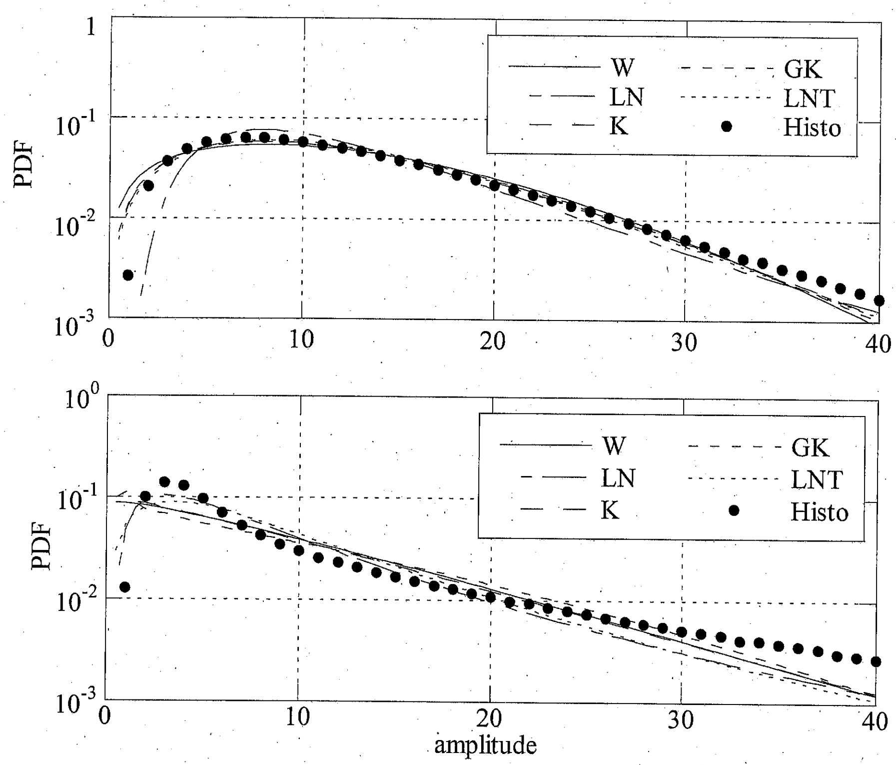


Figure 2 – Dense vegetated area (trees), file HB06192



Figures 3a-b – PDFs and histogram of clutter amplitude

1.2 Amplitude Analysis

After testing some statistical model on the overall amplitude of the clutter from dense vegetated areas, we performed a statistical analysis on speckle and texture separately[1]. We considered the data of file HB06192 where each tested model failed in modeling the amplitude histogram and the file HB06176 where almost all the models perform well.

In our analysis we supposed that the clutter process is a compound process $r(n)=\sqrt{\tau(n)}x(n)$, where $\tau(n)$ is the texture and $x(n)$ is the speckle[2]. As well know in the literature, texture and speckle have very different correlation times. Due to the physical nature of the texture that takes into account the space variations of the local power of the image, it can be considered as a long time (or space) correlated process. Then the texture can be supposed constant (or very slowly varying) on small patches of the image and it can be estimated from the amplitude data[2].

$$\hat{\tau}(i,j)=\frac{1}{N_r N_{cr}}\sum_{k=iN_r}^{(i+1)N_r-1}\sum_{m=jN_{cr}}^{(j+1)N_{cr}-1} r(k,m)^2 \tag{1}$$

where N_r and N_{cr} are the number of samples in range and cross-range respectively used for the estimation, then the dimensions (in pixels) of each patch. In performing this estimation we fixed N_r=N_{cr}=10, without overlap. It is apparent from eq. (1) that every 100 samples of clutter amplitude we obtained only 1 texture sample. The dimension of the square window has been chosen as a compromise between accuracy of the texture estimate and variation time of the texture itself. A larger window could improve the texture estimate only if the texture would be constant in a larger space interval, otherwise the estimation calculates only an average texture.

The speckle can be estimated by normalizing the data amplitude with respect to the estimated texture, thus

$$\hat{x}(k,m)=\frac{r(k,m)}{\sqrt{\hat{\tau}(i,j)}} \tag{2}$$

where $k\in[iN_c,(i+1)N_c]$ and $m\in[jN_{cr},(j+1)N_{cr}]$.

On speckle and texture we carried out a statistical analysis very similar to that performed on the amplitude. We used same models, Weibull, Rayleigh, K, LN and LNT for the speckle, and Gamma, Generalized Gamma (GΓ) and LN for the texture. The results are shown in Figures 4-7. For the grass field we obtained a good fitting of the texture histogram with the Gamma and LN PDFs. For the speckle the best fitting is to the Weibull PDF, that, in this case, is very similar to the Rayleigh PDF. Therefore, we can conclude that for the grass, even with very high resolution, the compound-Gaussian model is still good. We cannot state the same for the tree vegetated areas. The speckle is not more Rayleigh distributed. In our analysis we obtained the best fitting with the LNT PDF and the texture PDF does not follow anyone of the tested PDFs.

2 COVARIANCE ANALYSIS

The second step of our analysis was to estimate the autocovariance function of the data, both in range and cross-range. We performed the covariance analysis on the data amplitude, on the texture and on the speckle for both files. We show here only the results about the texture. The texture covariance function has been estimated as:

$$\hat{C}(m)=\frac{1}{N-|m|}\sum_{n=0}^{N-|m|-1}[\tau(n)-\hat{\mu}_\tau][\tau(n+m)-\hat{\mu}_\tau] \tag{3}$$

where $0 \le m \le N-1$ and $\hat{\mu}_\tau = \sum_{n=0}^{N_\tau - 1} \tau(n) \Big/ N_\tau$. This algorithm has been applied in both directions, then row-by-row and column-by-column of the texture estimated matrix. N_τ is the dimension of the texture matrix in range or cross-range. In the following figures we show the average covariance coefficient, that is:

$$\hat{c}(m) = \frac{1}{N_\tau} \sum_{k=1}^{N_\tau} \frac{\hat{C}_k(m)}{\hat{C}_k(0)} \tag{4}$$

where $\hat{C}_k(m)$ is the covariance for each kth row or column and $\hat{C}_k(0)$ is the relating variance.

For the grass clutter we can state that in both directions the differences in the covariances are small. The covariance coefficient plotted in Figure 8 shows that the covariance exhibits a fast drop in a distance of 5 m in both cross-range and range. After this interval it goes to zero slowly with some oscillations, particularly in cross-range. Figure 9 shows the covariance coefficient for dense vegetated area. The drop of the covariance in cross-range direction is faster than in range direction and it presents negative values. These phenomena are due to the presence of shadows, stronger and larger in cross-range direction than in range[3]. The speckle for both images and both directions is almost white.

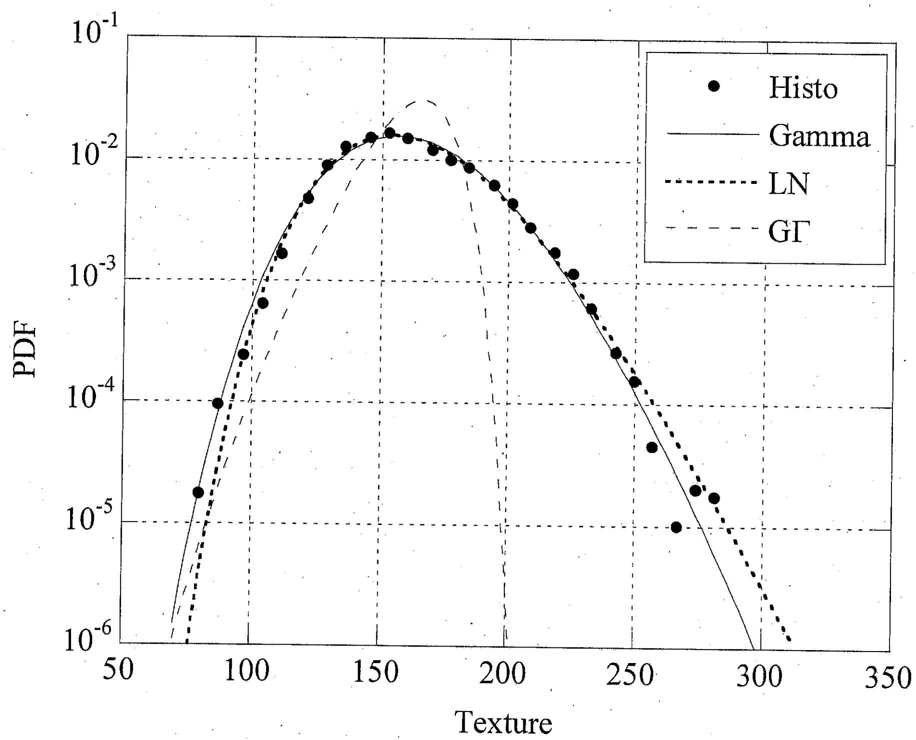


Fig.4 – PDFs and histogram of grass field texture

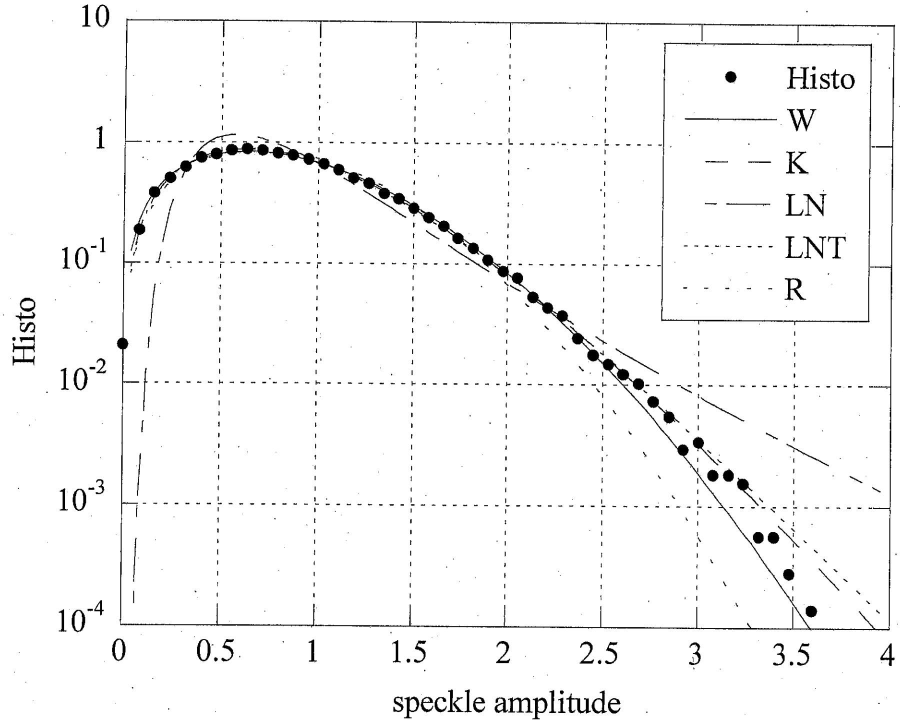


Fig. 5- PDFs and histogram of grass field speckle

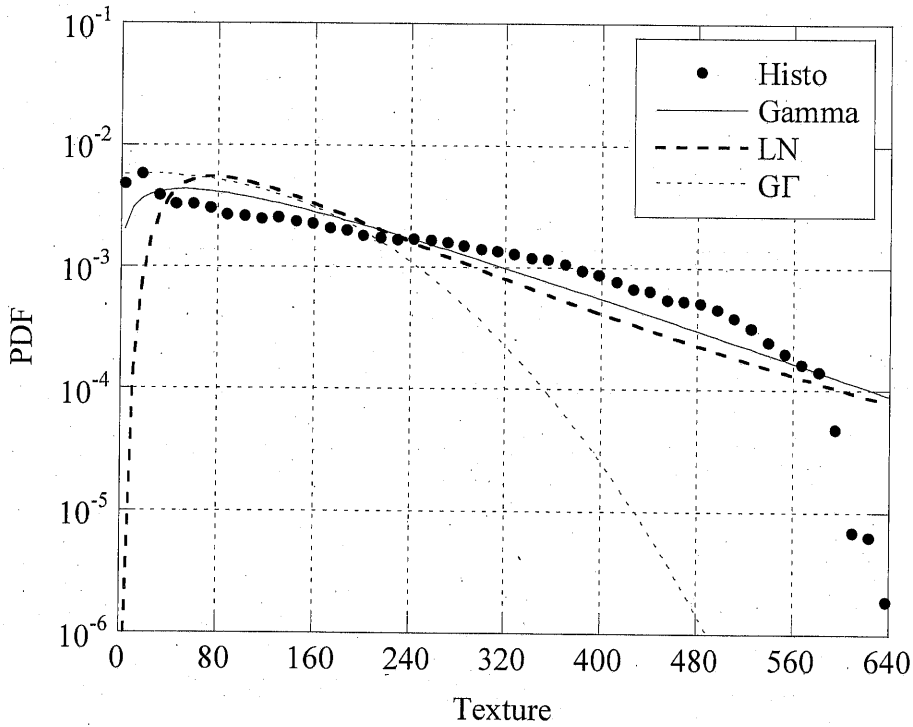


Fig. 6 - PDFs and histogram of the texture of tree vegetated area

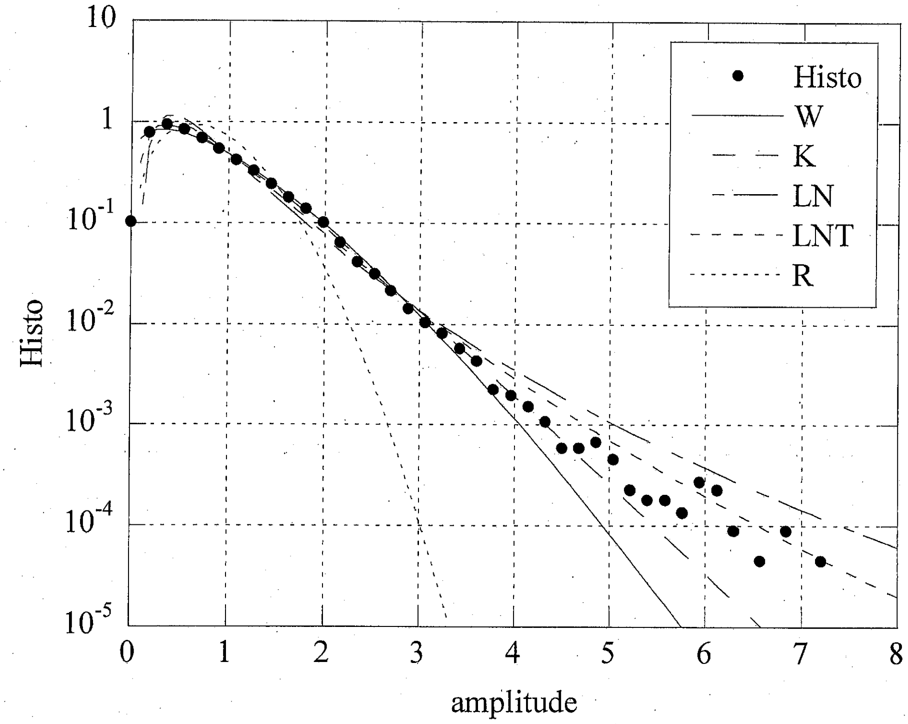


Fig. 7 - PDFs and histogram of the speckle of tree vegetated area

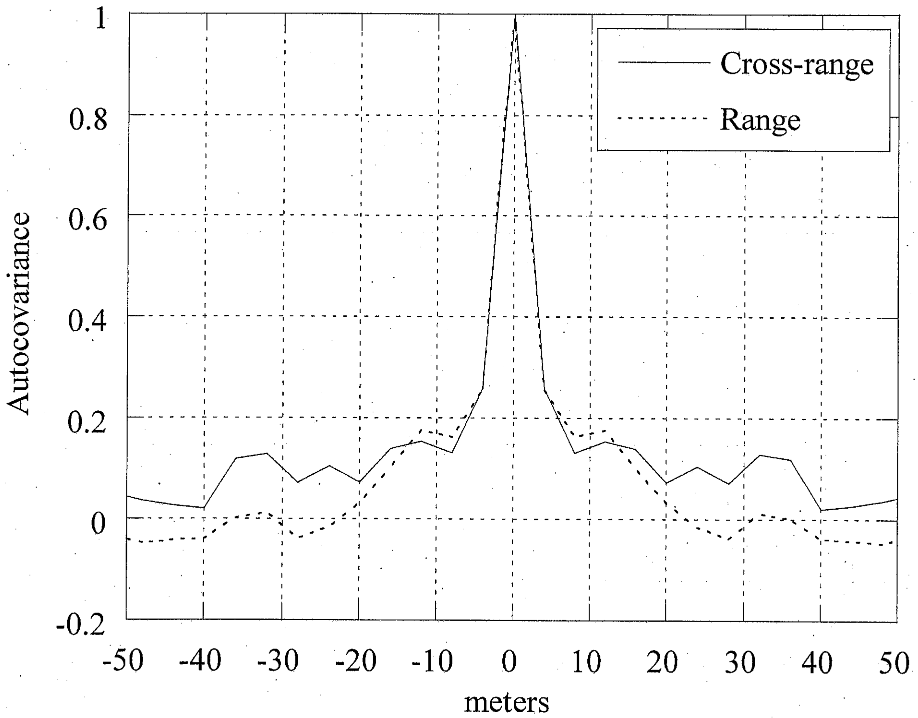


Fig. 8 – Covariance function of data of grass field

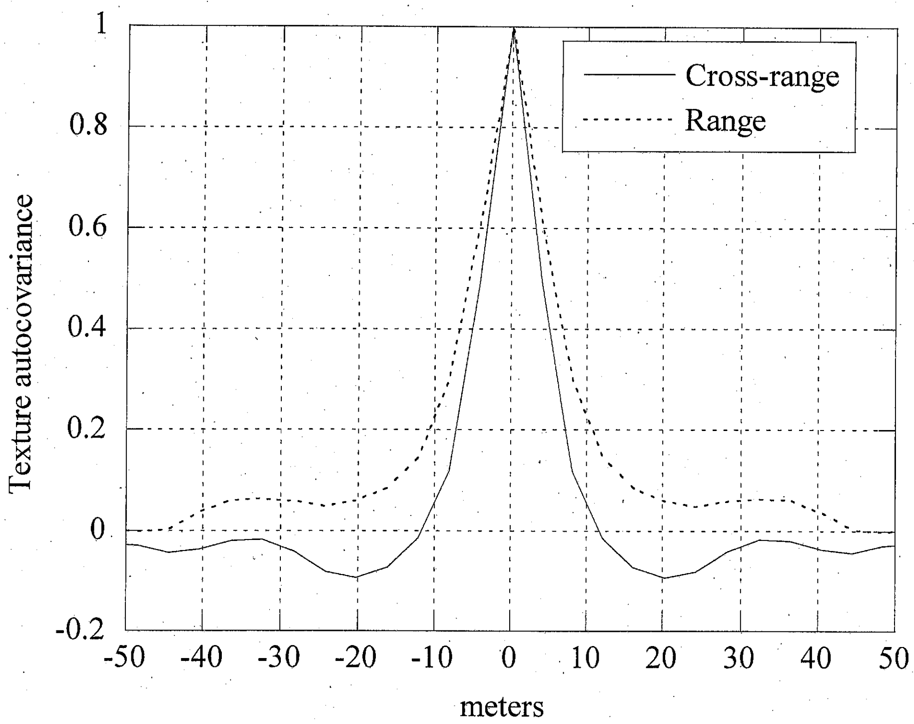


Fig. 9 – Covariance function of data of tree vegetated area

3 CONCLUSIONS

A detailed statistical analysis has been performed on MSTAR clutter data scattered by grass field and dense vegetated areas, manly covered by trees, with the aim of highlighting the differences in the scattering due to the different vegetation. Based on our results we can conclude that, while the compound-Gaussian model can be successfully applied to describe the grass clutter echoes, it cannot for dense vegetation, where speckle is not more Rayleigh distributed and the texture does not follow any tested PDF. Moreover, the correlation characteristic of the speckle does not depend on the vegetation but only on the sensor system, then the speckle is a white process. On the contrary, the texture generally presents a faster drop off in the grass clutter and in the cross-range direction.

Acknowledgments: The authors wish to thank the EOARD that kindly provided the MSTAR datasets.

4 REFERENCES

1. V. Anastassopoulos, G.A. Lampropoulos, A. Drosopoulos and M. Rey, 'High resolution radar clutter statistics', *IEEE Trans on AES* 35 (1) 43-60. (January 1999).
2. F. Gini, M. Greco, M. Diani, and L. Verrazzani, 'Performance Analysis of Two Adaptive Radar Detectors against Non-Gaussian Real Sea Clutter Data', *IEEE Trans. on AES* 36 (4) 1429-1439 (October 2000).
3. F.L. Posner, 'Texture and Speckle in High Resolution Synthetic Aperture Radar Clutter', *IEEE Trans. on GRS* 31(1) 192-203 (January 1993).