Parameter Estimation of Weibull Distribution Based on Second-Kind Statistics

Zengguo Sun and Chongzhao Han

Original scientific paper

Abstract: **The log-cumulant estimator is proposed to estimate the parameters of Weibull distribution based on second-kind statistics. With the explicit closed form expressions, the log-cumulant estimator is computationally efficient. Parameter estimation results from Monte Carlo simulation and real synthetic aperture radar (SAR) image demonstrate that the log-cumulant estimator leads to better performance when compared to the moment estimator.**

Index Terms: **Weibull distribution, parameter estimation, second-kind statistics, log-cumulant estimator**

I. Introduction

The Weibull distribution has been widely applied to synthetic aperture radar (SAR) images of sea, land, weather, and sea-ice clutter, and it contains the classical Rayleigh and exponential distributions as special cases [1-5]. With two parameters (shape parameter and scale parameter), the Weibull distribution can fit the experimental data better than the one-parameter distributions such as Rayleigh [1, 6]. For example, the Rayleigh distribution describes the early low-resolution SAR images well enough, but for the higher resolution SAR images, the two-parameter Weibull distribution can characterize the image contrast precisely [1].

In order to use the Weibull model in practical applications, its parameters should be estimated accurately. The estimation methods of Weibull distribution are summarized in [7], including linear estimator, maximum likelihood estimator, moment estimator, and Bayesian estimator. The linear estimator is the linear combinations of order statistics with suitably chosen coefficients. However, the determination of the coefficients is very difficult owing to a huge amount of computation, so it usually requires table look-ups. The maximum likelihood estimator is the parameter value that maximizes the likelihood function, given the data available. However, the maximum likelihood estimator has to solve the nonlinear equation, which usually requires the iterative method. The moment estimator estimates the parameters by directly using the statistical moments of Weibull distribution, but it does not have the explicit closed form and needs some numerical optimization techniques. The Bayesian estimator is the value of parameter that maximizes the posterior density in terms of the Bayesian theorem. However, the Bayesian estimator requires the prior distribution, which is not easy to determine.

In this letter, the log-cumulant estimator is proposed for the Weibull distribution based on second-kind statistics, which relies on the Mellin transform [8, 9]. We compare the log-cumulant estimator with the moment estimator, and we have observed that the performance of the moment estimator is degraded seriously for the small values of the shape parameter, but the log-cumulant estimator leads to high estimation accuracy no matter what values are chosen for the shape parameter, which is validated by parameter estimation results from Monte Carlo simulation and real SAR image experiment. Consequently, we recommend the log-cumulant estimator instead of the moment estimator.

This letter is organized as follows. The Weibull distribution is introduced in Section II. The moment estimator is briefly introduced in Section III, and the log-cumulant estimator based on second-kind statistics is proposed in Section IV, including the derivation process, Monte Carlo simulations, and real SAR image experiment. Lastly, this letter is concluded in Section V.

II. Weibull Distribution

The Weibull distribution has the following probability density function (pdf) [1]

$$f_{c,b}\left(x\right)=\frac{cx^{c-1}}{b^{c}}\exp\left[-\left(\frac{x}{b}\right)^{c}\right],\quad x\ge 0\,, \tag{1}$$

where c ($c>0$) is the shape parameter and b ($b>0$) is the scale parameter. The appearance of Weibull pdf is determined by the shape parameter c. When $c<1$, the pdf curve is J-shaped. When $c>1$, the pdf curve becomes skewed unimodal [7]. Denoting X as the Weibull-distributed random variable with parameters c and b, it can be demonstrated that the new random variable $\frac{X}{b}$ is still Weibull-distributed with the shape parameter c and the unit scale parameter ($b=1$).

Manuscript received June 12, 2010, revised September 4, 2010.

This work has been supported by the State Key Development Program for Basic Research of China (Grant No. 2007CB311006), the National High Technology Research and Development Program of China (Grant No. 2006AA01Z126), the National Natural Science Foundation of China (Grant No. 60602026), and the Specialized Research Fund for the Doctoral Program of Higher Education of China (Grant No. 20070698002).

Z. Sun and C. Han are with the Xi'an Jiaotong University, Xi'an Shaanxi, China. (email: duffer2000@163.com, czhan@mail.xjtu.edu.cn).

1845-6421/10/8213 © 2010 CCIS

This means that the Weibull distribution can be readily normalized. For various values of c, the pdf of Weibull distribution is plotted in Fig. 1. Obviously, the value of c controls the shape of the pdf. It should be noted that the Weibull distribution reduces to the exponential distribution when $c=1$ and to the Rayleigh distribution when $c=2$.

The Weibull distribution can be simulated by [7]

$$X = b\left[-\log(Y)\right]^{1/c}, \tag{2}$$

where X is the Weibull random variable with shape parameter c and scale parameter b, and Y is the random variable uniformly distributed in the interval $(0,1)$. With the help of (2), the Weibull-distributed samples can be simulated, which are shown in Fig. 2 for various values of c. It is apparent that the Weibull samples with $c=0.15$ show much severer impulsiveness than the ones with $c=2$. In general, the smaller the value of c is, the more impulsive the Weibull samples are. Since the Weibull-distributed samples can be simulated readily, we can use the Monte Carlo simulation to compare the performance of various parameter estimators.

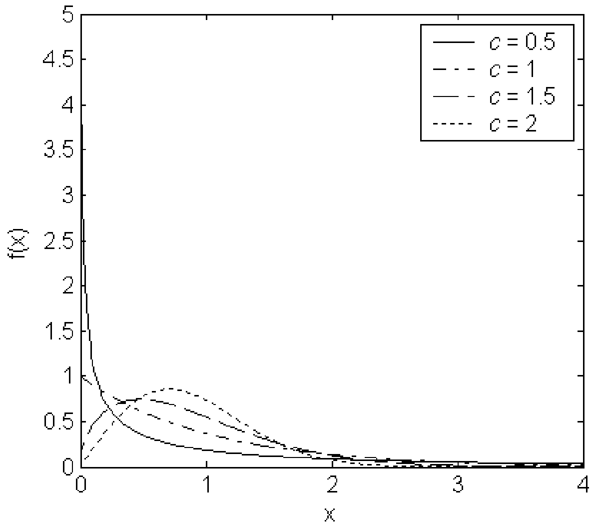
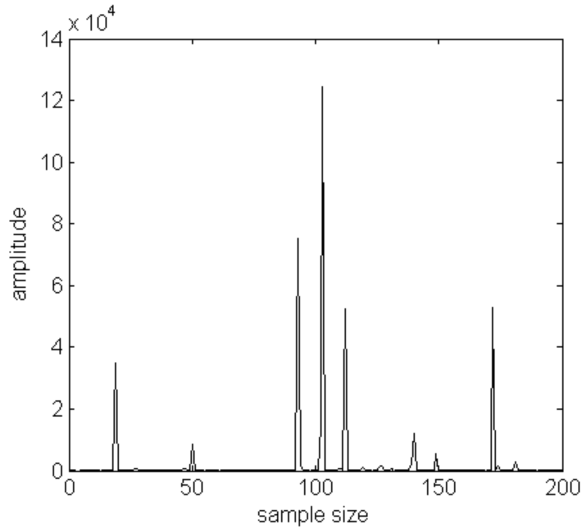
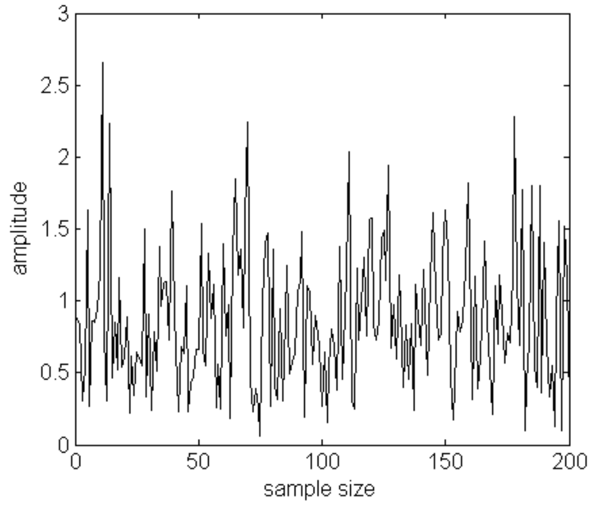


Fig. 1. Pdfs of Weibull distribution ($b=1$)



(a) $c=0.15$



(b) $c=2$

Fig. 2. Weibull-distributed samples ($b=1$ and the number of samples is 200)

III. Moment Estimator

The n th order moment of Weibull distribution can be written as

$$E\left(X^n\right) = b^n \Gamma\left(1+\frac{n}{c}\right), \quad n = 1, 2, \ldots, \tag{3}$$

where X is the Weibull random variable with parameters c and b, and $\Gamma(\cdot)$ is the Gamma function. Hence, the moment estimator for the Weibull distribution is straightforward as follows [1, 7]:

$$\frac{E\left(X^2\right)}{E^2(X)} = \frac{\Gamma(1+2/c)}{\Gamma^2(1+1/c)}, \tag{4}$$

$$E(X) = b\Gamma\left(1+\frac{1}{c}\right). \tag{5}$$

By replacing the actual moments with the sample moments, parameters c and b can be subsequently estimated from (4) and (5), using some numerical optimization techniques such as bisection [10].

The moment estimator was tested for various true values of parameter c according to Monte Carlo simulation. The Weibull-distributed samples were simulated independently by using (2), and the number of samples is 10000. For each true parameter c, the Monte Carlo simulation experiment was repeated 100 times independently, and then the average and standard deviation of the estimates were computed. The results are shown in Table I with standard deviations in parentheses. Obviously, the performance of the moment estimator relies on the true values of c. For the larger values of c, the moment estimator can lead to high estimation accuracy (e.g., $c=2$). However, if the smaller values are chosen for the c (e.g., $c=0.15$), the moment estimator results in poor performance. In other words, the moment estimator is sensitive to samples. If the samples show severe impulsiveness, which corresponds to the small values of the shape parameter (e.g., $c=0.15$), the moment estimator cannot achieve high estimation accuracy.

TABLE I
MONTE CARLO SIMULATION OF MOMENT ESTIMATOR (TRUE $b=1$)

True Value	$c=0.15$	$c=0.2$	$c=0.25$	$c=0.5$	$c=1$	$c=2$
$\hat{c}$	0.1819 (0.0178)	0.2203 (0.0170)	0.2616 (0.0168)	0.5027 (0.0111)	1.0024 (0.0090)	2.0017 (0.0176)
$\hat{b}$	9.7405 (5.9173)	2.2442 (1.0153)	1.3134 (0.3930)	1.0061 (0.0411)	0.9998 (0.0095)	0.9991 (0.0051)

IV. LOG-CUMULANT ESTIMATOR

In [8], the log-cumulant estimator was used for parameter estimation of the α - stable positive distributions due to its explicit expressions. The log-cumulant estimator is based on second-kind statistics, which relies on the Mellin transform. Denoting g as a function defined over $[0,+\infty)$, its Mellin transform is defined as

$$\mathrm{M}\left[g(x)\right](s)=\int_0^{+\infty} x^{s-1} g(x)\,dx\,, \tag{6}$$

where s is the complex variable of the transform. Specifically, for a pdf f defined in $[0,+\infty)$, analogous to the case of common statistics based on Fourier transform, the second-kind statistic functions are defined as follows [9]:

·Second-kind first characteristic function

$$\Phi(s)=\int_0^{+\infty} x^{s-1} f(x)\,dx \tag{7}$$

·Second-kind second characteristic function

$$\Psi(s)=\log\left(\Phi(s)\right) \tag{8}$$

· r th order second-kind cumulant (log-cumulant)

$$\tilde{k}_r=\left.\frac{d^r\Psi(s)}{ds^r}\right|_{s=1} \tag{9}$$

The first two log-cumulants $\tilde{k}_1$ and $\tilde{k}_2$ can be estimated empirically from N samples y_i as follows [9]:

$$\begin{aligned}\hat{\tilde{k}}_1&=\frac{1}{N}\sum_{i=1}^{N}\left[\log(y_i)\right]\\ \hat{\tilde{k}}_2&=\frac{1}{N}\sum_{i=1}^{N}\left[\left(\log(y_i)-\hat{\tilde{k}}_1\right)^2\right]\end{aligned}. \tag{10}$$

By substituting (1) into (7) and after some manipulation, the second-kind first characteristic function of Weibull distribution can be written as

$$\Phi(s)=\frac{c}{b^c}\int_0^{+\infty} x^{s+c-2}\exp\left(-\frac{x^c}{b^c}\right)dx\,. \tag{11}$$

Using the following identity [11]

$$\begin{gathered}\int_0^{+\infty} x^{\nu-1}\exp\left(-\mu x^p\right)dx=\frac{1}{p}\mu^{-\frac{\nu}{p}}\Gamma\left(\frac{\nu}{p}\right),\\ \mathrm{Re}(\mu)>0,\ \mathrm{Re}(\nu)>0,\ \ p>0\end{gathered} \tag{12}$$

one arrives at

$$\Phi(s)=b^{s-1}\Gamma\left(1+\frac{s-1}{c}\right). \tag{13}$$

Then, substituting (13) into (8) and subsequently into (9), the first and second orders log-cumulants for the Weibull distribution are obtained as follows:

$$\tilde{k}_1=\log(b)-\frac{\mathrm{C}_e}{c}\,, \tag{14}$$

$$\tilde{k}_2=\frac{\pi^2}{6c^2}\,. \tag{15}$$

Here, C_e is the Euler's constant. Replacing the actual log-cumulants $\tilde{k}_1$ and $\tilde{k}_2$ with the sample log-cumulants $\hat{\tilde{k}}_1$ and $\hat{\tilde{k}}_2$ in (10) respectively, the log-cumulant estimator for the Weibull distribution is finally obtained as follows:

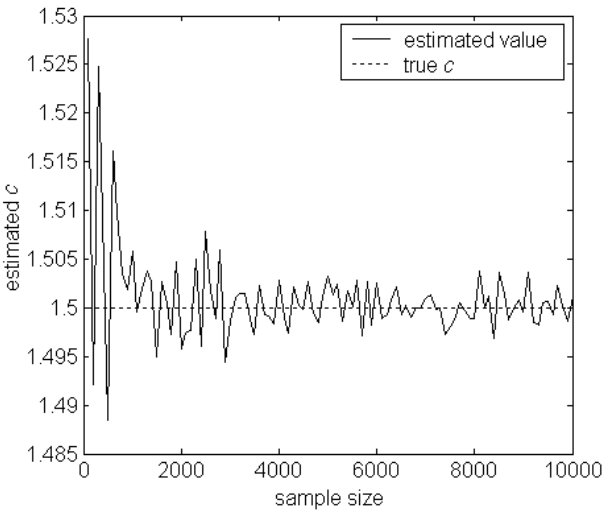
$$\hat{c}=\frac{\pi}{\sqrt{6\hat{\tilde{k}}_2}}\,, \tag{16}$$

$$\hat{b}=\exp\left(\hat{\tilde{k}}_1+\frac{\mathrm{C}_e}{\hat{c}}\right). \tag{17}$$

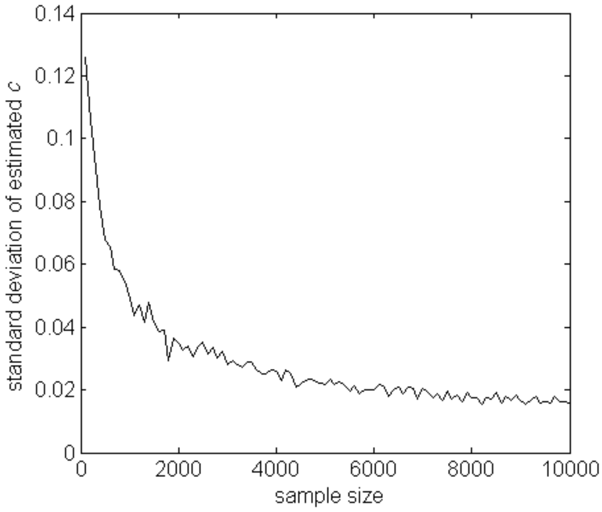
Obviously, after calculating the sample log-cumulants, the shape parameter c can be estimated firstly, and then the scale parameter b can be estimated. Compared to the moment estimator (equations (4) and (5)), the log-cumulant estimator does not need to solve the nonlinear equation due to its explicit closed form expressions, so it is computationally simple.

Firstly, the log-cumulant estimator was tested on Monte Carlo simulations. From equation (10), it is obvious that the sample size N is an important factor that determines the performance of the log-cumulant estimator, which is illustrated in Fig. 3. For each sample size chosen, the Monte Carlo simulation was repeated 100 times independently, and the average and standard deviation of all estimated parameter values were selected as the final estimation results. In general, when the sample size is getting larger, the estimated parameter values approach to the true parameter, and the standard deviations are becoming lower. This means that the larger sample size leads to the higher performance of the log-cumulant estimator. For Monte Carlo simulations, the Weibull-distributed samples are generated randomly according to (2), so the Monte Carlo simulation should be independently repeated many times, and the average of all estimated parameters should be selected as the final estimated parameter value. This indicates that the running times is another important factor that determines the estimation accuracy of the log-cumulant estimator, which is illustrated in Fig. 4. Obviously, more running times leads to higher estimation accuracy. For various true values of the shape parameter c, Table II illustrates the average and standard deviation values (in parentheses) of Monte Carlo simulation results based on the log-cumulant estimator, and Fig. 5 shows the performance comparison of the log-cumulant estimator and the moment estimator as a function of true c. For each true parameter c, the Monte Carlo simulation experiment was repeated 100 times independently, and the number of samples was 10000 for each time. Obviously, the log-cumulant estimator leads to high estimation accuracy no matter what values are chosen for the true c. Therefore, the log-cumulant estimator is robust and not sensitive to samples. Even for the samples with severe impulsiveness (e.g., $c = 0.15$), the log-cumulant estimator can achieve high estimation accuracy. For the moment estimator, on the other hand, the estimated parameters may be close to that of the log-cumulant estimator for the larger values of c. However, as shown in Table I, the performance of the moment estimator is deteriorated for the smaller values of c (e.g., $c = 0.15$). In a word, the log-cumulant estimator is superior to the moment estimator, which is validated by the Monte Carlo simulations.

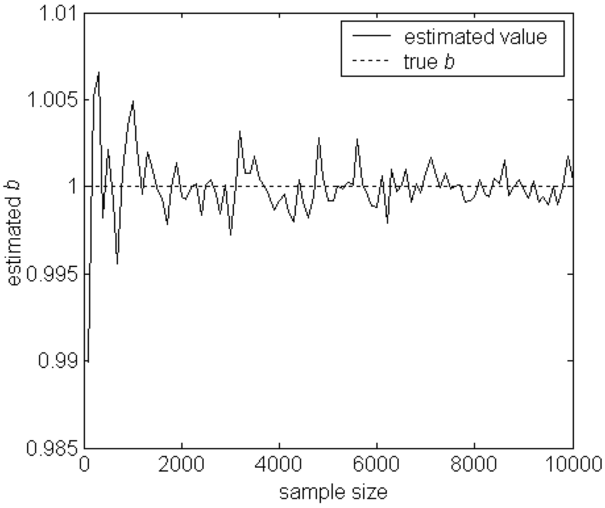
Secondly, the log-cumulant estimator was tested on the real SAR image in Fig. 6 which was obtained from the Sandia National Laboratories. Modeling the Fig. 6 with the Rayleigh distribution and the Weibull distribution whose parameters were estimated from the moment estimator and the log-cumulant estimator respectively, the results are shown in Fig. 7, and the estimated parameters and the corresponding K-S fit probability are provided in Table III [1, 10] (It should be noted that the Rayleigh distribution is just the Weibull distribution in (1) when $c = 2$). The K-S fit probability describes the fitness of statistical distribution to real SAR image. The larger value of the K-S fit probability, the better fit of distribution to the image. In this letter, the K-S fit probability is selected as a quantitative measure to test the estimation accuracy of various estimators. The larger the K-S fit probability is, the higher estimation accuracy the estimator has. Compared to the moment estimator, the log-cumulant estimator leads to the Weibull distribution that fits the SAR image well enough especially the high peak, so the log-cumulant estimator corresponds to the larger value of K-S fit probability. This indicates that the log-cumulant estimator has higher estimation accuracy compared to the moment estimator. In addition, it is obvious that the Rayleigh distribution cannot describe the statistical characteristics of the high-resolution SAR image, which results in the lowest K-S fit probability.



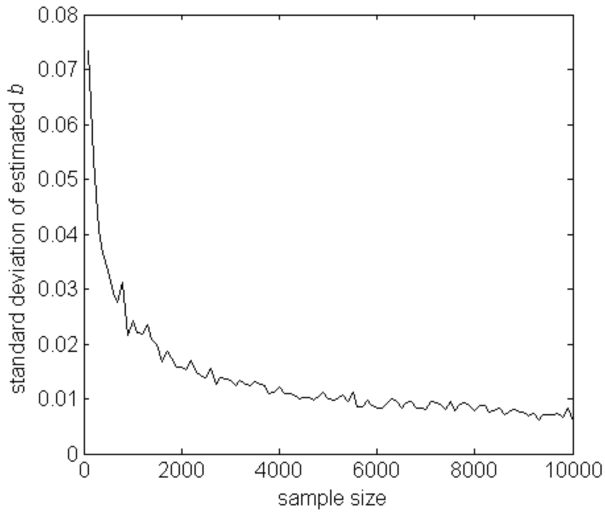
(a) $\hat{c}$ varying with sample size



(b) Standard deviation of $\hat{c}$

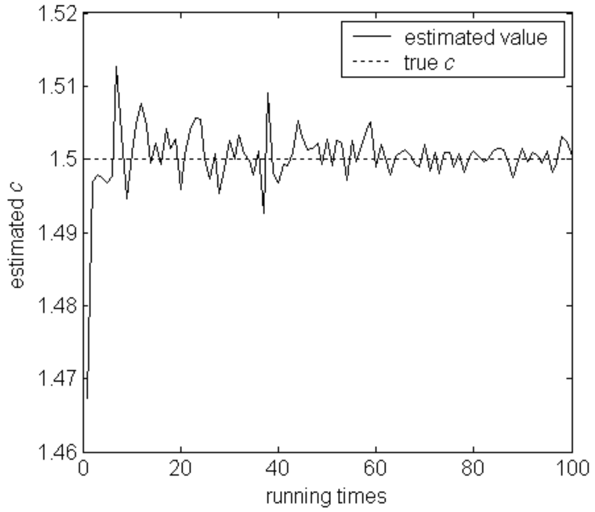


(c) $\hat{b}$ varying with sample size

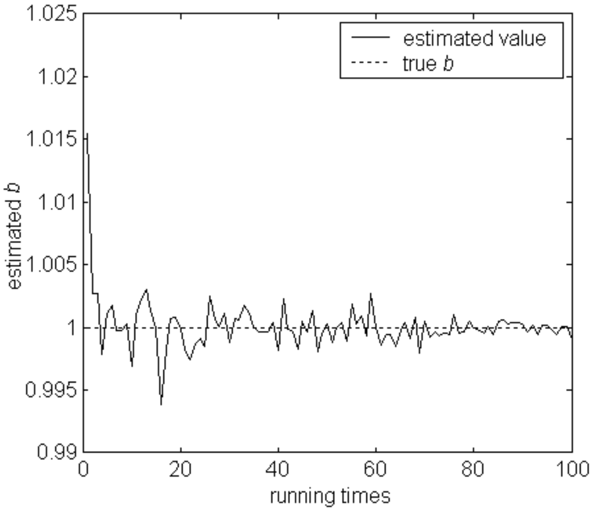


(d) Standard deviation of $\hat{b}$

Fig. 3. Performance of log-cumulant estimator as a function of sample size (true $c = 1.5$, $b = 1$)



(a) $\hat{c}$ varying with running times

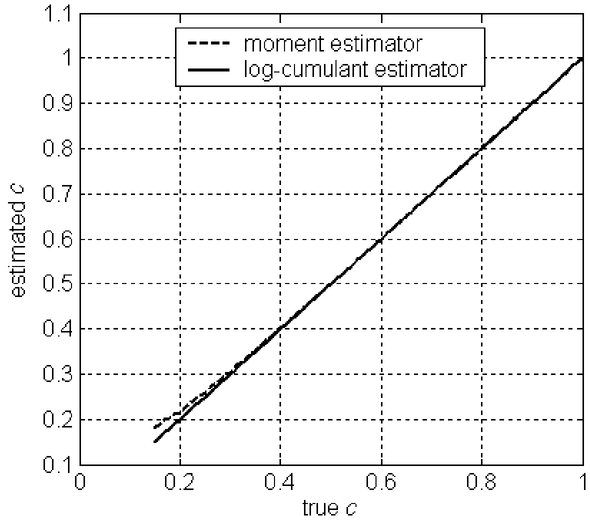


(b) $\hat{b}$ varying with running times

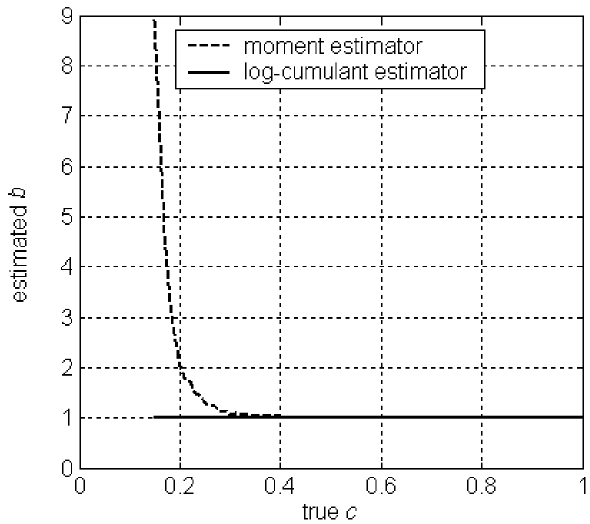
Fig. 4. Estimation accuracy of log-cumulant estimator as a function of running times (true $c = 1.5$, $b = 1$)

TABLE II
MONTE CARLO SIMULATION OF LOG-CUMULANT ESTIMATOR (TRUE $b = 1$)

True Value	$c=0.15$	$c=0.2$	$c=0.25$	$c=0.5$	$c=1$	$c=2$
$\hat{c}$	0.1503 (0.0015)	0.2004 (0.0023)	0.2504 (0.0025)	0.4995 (0.0051)	1.0012 (0.0105)	2.0000 (0.0228)
$\hat{b}$	1.0010 (0.0761)	0.9998 (0.0487)	0.9994 (0.0404)	1.0012 (0.0193)	1.0001 (0.0120)	0.9996 (0.0059)



(a) $\hat{c}$ as a function of true c



(b) $\hat{b}$ as a function of true c

Fig. 5. Monte Carlo simulation comparison of log-cumulant estimator and moment estimator (true $b = 1$)

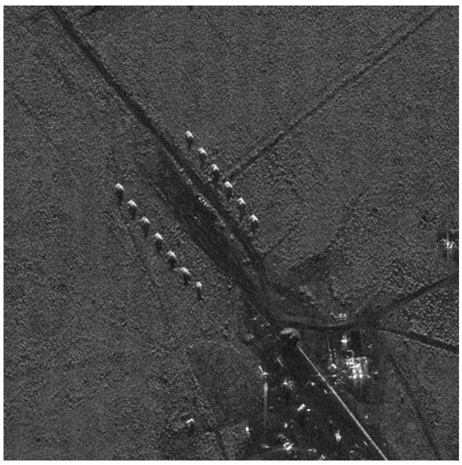


Fig. 6. High-resolution SAR image (Kirtland Air Force Base, Albuquerque, NM, 1-ft resolution)

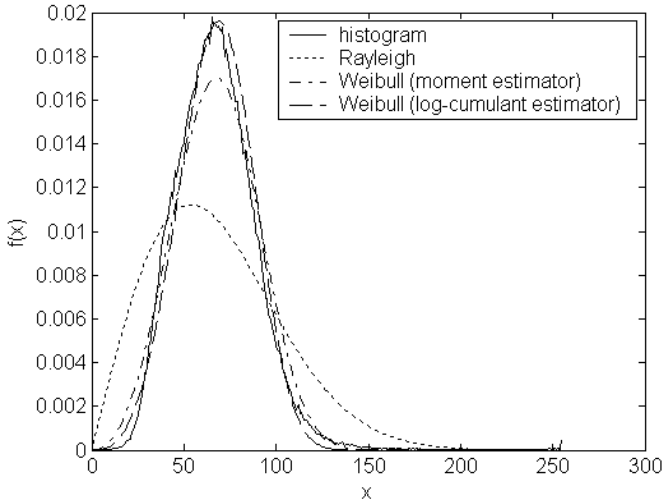


Fig. 7. Modeling Fig. 6 with Rayleigh and Weibull distributions

TABLE III
ESTIMATED PARAMETERS AND K-S FIT PROBABILITY OF VARIOUS DISTRIBUTIONS FOR FIG. 7

	Estimated Parameters	Probability
Rayleigh	$\hat{b} = 76.4200$	0.8516
Weibull (moment estimator)	$\hat{c} = 3.3126$ $\hat{b} = 75.4865$	0.9616
Weibull (log-cumulant estimator)	$\hat{c} = 3.8275$ $\hat{b} = 74.5996$	0.9748

V. CONCLUSION

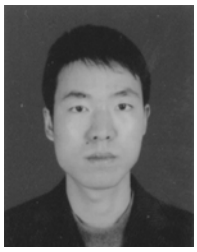
The log-cumulant estimator based on the second-kind statistics is proposed to estimate the parameters of Weibull distribution in this letter. Compared to the moment estimator, the log-cumulant estimator has explicit closed form expressions, and it can achieve good performance even for the severely impulsive samples. Parameter estimation results from Monte Carlo simulations and real SAR image demonstrate that the log-cumulant estimator leads to higher estimation accuracy compared to the moment estimator. Therefore, we suggest adopting the log-cumulant estimator instead of the moment estimator. In this work, we test the log-cumulant estimator experimentally. In the future research, we should evaluate the log-cumulant estimator theoretically.

ACKNOWLEDGMENT

The authors would like to thank the Sandia National Laboratories for providing the precious SAR image.

REFERENCES

[1] C. Oliver, S. Quegan: "Understanding Synthetic Aperture Radar Images", Boston, MA, Artech House, 1998.
[2] M. Sekine, T. Musha, Y. Tomita, T. Hagisawa, T. Irabu, E. Kiuchi: *Weibull-Distributed Sea Clutter*, IEE Proceedings F, Vol. 130, No. 5, p. 476, Aug. 1983.
[3] M. Sekine, S. Ohtani, T. Musha, T. Irabu, E. Kiuchi, T. Hagisawa, Y. Tomita: *Weibull-Distributed Ground Clutter*, IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-17, No. 4, pp. 596-598, Jul. 1981.
[4] M. Sekine, T. Musha, Y. Tomita, T. Hagisawa, T. Irabu, E. Kiuchi: *On Weibull-Distributed Weather Clutter*, IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-15, No. 6, pp. 824-830, Nov. 1979.
[5] H. Ogawa, M. Sekine, T. Musha, M. Aota, M. Ohi, H. Fukushi: *Weibull-Distributed Radar Clutter Reflected from Sea Ice*, IEICE Transactions, Vol. E70-E, No. 2, pp. 116-120, Feb. 1987.
[6] M. I. Skolnik: "Introduction to Radar Systems", New York, McGraw-Hill, Third Edition, 2001.
[7] H. Rinne: "The Weibull Distribution: A Handbook", Boca Raton, CRC Press, 2009.
[8] J. M. Nicolas: *Alpha-Stable Positive Distributions: A New Approach Based on Second Kind Statistics*, in Proc. EUSIPCO, Toulouse, France, Sept. 2002, pp. 197-200.
[9] C. Tison, J. M. Nicolas, F. Tupin, H. Maitre: *A New Statistical Model for Markovian Classification of Urban Areas in High-Resolution SAR Images*, IEEE Transactions on Geoscience and Remote Sensing, Vol. 42, No. 10, pp. 2046-2057, Oct. 2004.
[10] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery: "Numerical Recipes in C", Cambridge, UK, Cambridge University Press, 1994.
[11] I. S. Gradshteyn, I. M. Ryzhik: "Table of Integrals, Series, and Products", New York, Academic Press, Seventh Edition, 2007.



Zengguo Sun is a Ph. D. candidate at the Institute of Integrated Automation, School of Electronic and Information Engineering, Xi'an Jiaotong University. He received his BS (2003) from Xi'an Jiaotong University, China, in computer science and technology. Sponsored by China Scholarship Council, he stayed in Pennsylvania State University, USA, as a collaborative researcher from Sept. 2007 to Sept. 2008. His research interest is the synthetic aperture radar image processing, including noise suppression, edge detection, image modeling, and parameter estimation. He has published several journal and conference papers on these fields.

Chongzhao Han was graduated from Department of Electric Machinery Engineering, Xi'an Jiaotong University, Xi'an, China, in 1968. He received his ME (1981) from Graduate School of Chinese Academy of Sciences in automatic control. Now he is a Professor at the Department of Automation, School of Electronic and Information Engineering, Xi'an Jiaotong University. His research interest covers multi-source information fusion, stochastic control and adaptive control, and nonlinear spectral analysis.