Available online at www.sciencedirect.com
ScienceDirect
Fuzzy Sets and Systems 163 (2011) 45–53
FUZZY sets and systems
www.elsevier.com/locate/fss

Parameter estimation of K-distributed sea clutter based on fuzzy inference and Gustafson–Kessel clustering

Atefeh Davari[a,*], Mohammad Hamiruce Marhaban[b], Samsul Bahari Mohd Noor[a], Mohammad Karimadini[c], Ali Karimoddini[c]

[a] *Electrical and Electronic Engineering Department, Faculty of Engineering, Universiti Putra Malaysia, Malaysia*
[b] *Institute of Advanced Technology, Faculty of Engineering, Universiti Putra Malaysia, Malaysia*
[c] *Department of Electrical & Computer Engineering, Faculty of Engineering, National University of Singapore, Singapore*

Received 24 June 2008; received in revised form 20 July 2009; accepted 16 September 2010
Available online 25 September 2010

Abstract

The detection performance of maritime radars is restricted by the unwanted sea echo or clutter. Although the number of these target-like data is small, they may cause false alarm and perturb the target detection. K-distribution is known as the best fit probability density function for the radar sea clutter. This paper proposes a novel approach to estimate the parameters of K-distribution, based on fuzzy Gustafson–Kessel clustering and fuzzy Takagi–Sugeno Kang modelling. The main contribution of the proposed method is the ability to estimate the parameters, given a small number of data which will usually be the case in practical applications. This is achieved by a pre-estimation using fuzzy clustering that provides a prior knowledge and forms a rough model to be fine tuned using the least square method. The algorithm also improves the calculations of shape and width of membership functions by means of clustering in order to improve the accuracy. The resultant estimator then acts to overcome the bottleneck of the existing methods in which it achieves a higher performance and accuracy in spite of small number of data.

Keywords: K-distribution; Parameter estimation; Fuzzy GK-clustering; Fuzzy TSK modelling

1. Introduction

K-distribution model has been demonstrated to fit the sea clutter [1] and successfully has been used in many signal processing and radar detection applications [2]. The statistics of a K-distributed random variable x have been described in [3] by the probability density function as

$$p(x) = \frac{2}{a\Gamma(v)}\left(\frac{x}{2a}\right)^{v} \mathrm{K}_{v-1}\left(\frac{x}{a}\right), \quad x > 0 \tag{1.1}$$

* Corresponding author. Tel.: +60 6591759285.
E-mail addresses: atefehdavari@gmail.com (A. Davari), hamiruce@eng.upm.edu.my (M. Hamiruce Marhaban), samsul@ieee.org (S. Bahari Mohd Noor), karimadini@nus.edu.sg (M. Karimadini), akarimoddini@gmail.com (A. Karimoddini).

0165-0114/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.fss.2010.09.008

where a is the scale parameter, v is the shape parameter, $\Gamma(\cdot)$ denotes the gamma function and $K_\lambda(\cdot)$ is modified Bessel function of order λ. For high resolution sea clutters, the shape parameter (v) is typically in the range $0.1 \leq v \leq \infty$, where v close to 0.1 represents a very spiky clutter and $v = \infty$ corresponds to thermal noise [4].

For an accurate parameter estimation of K-distribution, a lot of field data is required. However, only a small number of data is available for the said purpose in real practice [5–7]. Therefore, the estimator needs to be able to accurately estimate the parameters in given a small number of samples. In the critical situations on the other hand, the targeting must be performed as fast as possible. Thus, besides the accuracy, the researchers also endeavour to simplify the algorithm in order to reduce the program cycle time.

There are many methods to estimate K-distribution parameters. Maximum likelihood (ML) method can be used to estimate a and v [8]. This method yields an efficient estimate but it involves a complex computation and it is inappropriate to be implemented in real-time system [9]; however, it is suitable for applications with large amounts of data while only a high accurate estimation is required. Method of moments (MOM) have been proposed in [10], that leads to accurate estimate but it still involves with computational complexity problem. The simplest type of the moment methods is composed of the second and fourth order moments (SFMOM) or Raghvavan's method [10]. Fractional lower order moment (FLOM) is another type of moment method which is proposed by Iskander [9]. These approaches perform well when the number of sample data is large. In terms of accuracy, so far the ML/MOM method [7] and Hadamard-transformation-based method [3], show higher accuracy than the others. The existing methods have looked into increasing the accuracy and simultaneously reducing the computational complexity of the algorithm. However, the bottleneck that the methods seem to suffer is due to small number of data.

This paper proposes a novel approach through the design of an accurate and fast estimator to estimate the parameter of the K-distribution, given a very small set of sample data. A combination of GK-clustering and TSK modelling has been chosen to take the advantage of clustering as a pre-estimator and later fine-tuning by least square regression. The main contribution of the proposed method is the capability of handling the estimation using a small number of data.

The rest of the paper is outlined as follows. The preliminaries including Takagi–Sugeno Kang (TSK) fuzzy modelling and Gustafson–Kessel (GK) clustering are addressed in Section 2. The estimation of K-distribution parameter based on GK-clustering and TSK modelling is proposed in Section 3. The simulation results and comparisons with the previous methods are illustrated in Section 4. Finally, the paper is concluded in Section 5.

2. Preliminaries

2.1. Takagi–Sugeno Kang (TSK) fuzzy modelling

Fuzzy set theory has been introduced by Zadeh [11], addressing the partial membership in set theory. The application of fuzzy set theory in control systems was pioneered by Mamdani [12]. An alternative fuzzy system was introduced by Takagi and Sugeno [13] to facilitate the fuzzy system with mathematical consequents. The TSK modelling has been demonstrated to function as an efficient model for systems that can be fully represented by their input/output relationships. Decision can be made depending on a "rule base", where the rule consequents are linear parametric equations in terms of inputs of the system. Theoretically, the consequents part could be any function, even a non-linear one. However, linear functions have been employed in most cases due to their simplicity. A typical rule of TSK model with M rules, each rule having p antecedents can be expressed as

$$R^l : \text{IF } x_1 \text{ is } \tilde{F}_1^l \text{ and } x_2 \text{ is } \tilde{F}_2^l \text{ and} \cdots \text{and } x_p \text{ is } \tilde{F}_p^l \text{ Then } y^l = f^l(x_1, x_2, \ldots, x_p) \tag{2.1}$$

In particular, f can be a linear combination of inputs as

$$R^l : \text{IF } x_1 \text{ is } \tilde{F}_1^l \text{ and } x_2 \text{ is } \tilde{F}_2^l \text{ and} \cdots \text{and } x_p \text{ is } \tilde{F}_p^l \text{ Then } y^l = c_0^l + c_1^l x_1 + c_2^l x_2 + \cdots + c_p^l x_p \tag{2.2}$$

where c_i^l ($i = 0, 1, \ldots, p, l = 1, 2, \ldots, M$) are the consequent parameters; y^l are the output of the l-th IF–THEN rule which are already crisp numbers; and $\tilde{F}_j^l (j = 1, 2, \ldots, p)$ are fuzzy sets in the antecedent parts. x_i are crisp numbers belonged to fuzzy sets $\tilde{F}_j^l$ and they are to be fuzzified. Given an input $(x_1, x_2, \ldots, x_p)$, the final output of the first-order

TSK model is inferred as

$$y = \sum_{l=1}^{M} \frac{\omega^l y^l}{\sum_{l=1}^{M} \omega^l} \tag{2.3}$$

where ω^l are rule firing strengths of the rules. If product is used as T-norm, in premise of the rules, the applicability degree can be derived as

$$\omega^l = \tilde{F}_1^l \cdot \tilde{F}_2^l \cdots \tilde{F}_p^l \tag{2.4}$$

The construction of a TSK model is usually done in two steps. At first, the fuzzy sets (membership functions) in the rule antecedents are determined. This can be done manually, using knowledge of the process, or by some data-driven techniques. Next, the parameters of the consequent functions are estimated which will be discussed in the following sections.

2.2. Gustafson–Kessel (GK) clustering

Fuzzy clustering uses the fuzzy set theory and clustering techniques based on objective functions on distances between clusters and distances between data and clusters. The early works on fuzzy clustering [14,15] were based upon the Euclidean distance [16] that could not capture the correlation between data. This problem has been solved in GK-clutering [18] by using Mahalanobis distance [17] to consider the distribution of data by incorporating the covariance of data. The Mahalanobis distance takes the correlation of the dataset into account and is scale-invariant, i.e. not dependent on the scale of measurements. An improved version of fuzzy GK-clustering was introduced in [19] to tackle singularity and overfitting.

Assume $Z \in \mathrm{R}^{n\times N}$ is the identification data, where n is the number of dimensions in the feature space and N is the number of data. The aim is to classify N data into c clusters. Therefore, the partition matrix can be denoted as [18]

$$u = [\beta_{ik}]_{c\times n}, \quad \beta_{ik} \in [0, 1] \tag{2.5}$$

where β_{ik} represent the membership degree of k-th data to i-th cluster. The cluster centres are called "prototype". Prototypes are denoted as

$$V = [V_1, V_2, \ldots, V_c], \quad V_i \in \mathrm{R}^n \tag{2.6}$$

These parameters are derived through out the algorithm by minimizing the cost function

$$L = \sum_{i=1}^{c} \sum_{k=1}^{N} (\beta_{ik})^e D_{ik}^2 \tag{2.7}$$

where $e > 1$ is the degree of fuzzification and is typically set to 2, D_{ik} is Mahalanobis distance between k-th data to i-th prototype, and is calculated as

$$D_{ikA_i}^2 = (Z_k - V_i)^T A_i (Z_k - V_i) \tag{2.8}$$

$$A_i = \rho_i \det(F_i)^{1/n} F_i^{-1} \tag{2.9}$$

$$F_i^l = \frac{\sum_{i=1}^{N} (\beta_{ik}^{(l-1)})^e (Z_k - V_i^l)(Z_k - V_i^l)^T}{\sum_{i=1}^{N} (\beta_{ik}^{(l-1)})^e} \tag{2.10}$$

Thus given the dataset Z, the number of clusters $1 < c < N$, the weighting exponent $e > 1$ (usually 2), the termination tolerance $\varepsilon > 0$ (usually 10^{-3}) and the cluster volumes ρ_i (usually 1) are the parameters to be chosen. Covariance matrix is considered to take the distribution of data into account through the process of minimizing of the cost function, as Eqs. (2.7)–(2.10).

The implementation of GK-clustering algorithm is represented in the following steps [18,19]:

Step 1: Determine the number of clusters in terms of priori knowledge about the system. Then initialize partition matrix $u = [\beta_{ik}]$ with random values taken from [0,1] interval. Rows of partition matrix correspond to clusters while its columns associate with data. Hence each row of partition matrix has to include at least one non-zero entry and the summation of each column of this matrix should be 1;

Step 2: Calculate prototypes using

$$v_i^l = \frac{\sum_{k=1}^{N} (\beta_{ik}^{l-1})^e Z_k}{\sum_{k=1}^{N} (\beta_{ik}^{l-1})^e}; \tag{2.11}$$

Step 3: Compute $D_{ikA_i}^2$ values by using Eqs. (2.8)–(2.10).

Step 4: Update partition matrix based on

$$\beta_{ik}^l = \frac{1}{\sum_{j=1}^{C} \left(\frac{D_{ikA_i}}{D_{jkA_j}} \right)^{2/(e-1)}}; \tag{2.12}$$

Step 5: Go to step 2 and repeat the procedures, until convergence inequality $\|u^{(l+1)} - u^{(l)}\| < \varepsilon$ is satisfied, where $\|u^{(l+1)} - u^{(l)}\|$ is the change in the norm of partition matrix, l is the number of iterations and ε is a small real positive value (e.g. 0.001).

The GK-clustering algorithm is a powerful clustering technique; however, numerical problems may occur frequently often in the standard GK-clustering when the number of data in some cluster is small or when the data within a cluster are almost linearly correlated. In such cases, the cluster covariance matrix becomes singular and cannot be inverted to compute the matrix determinant as well as norm-inducing matrix. Babuska in [19] presented a method to overcome the singularity problem by fixing the ratio between the maximal and minimal eigenvalues of the covariance matrix. The other problem is overfitting in which some clusters are formed extremely long in the direction of the largest eigenvalues. In this instance the covariance matrix does not reflect the actual distribution of data and consequently a poor model will be obtained. The following modification for GK algorithm has been proposed in [19] to overcome this problem:

$$F_i^{new} = (1 - \gamma)F_i + \gamma \det(F_0)^{1/n} I \tag{2.13}$$

where $\gamma \in [0, 1]$ is a tuning parameter and F_0 is the covariance matrix of the whole dataset. The scaling parameter γ should be big enough to overcome the overfitting problem. On the other hand when γ is close to 1, all the covariance matrices converge to the same matrix $\det(F_0)^{1/n} I$. In this case GK-clustering is reduced to K-means clustering in which no covariance matrix is taken into account.

A partitional fuzzy clustering method based on adaptive quadratic distances has been introduced in [20]. It uses Babuska's method [19] in order to overcome the singularity. Although this method is able to modify the clustering distances adaptively in each iteration, it suffers from the overfitting problem for the application of K-distribution parameter estimation, especially when experimenting with small number of samples. Moreover, it utilises the Euclidean distance that has a limitation to capture a fair distance due to the fact that the correlation of dataset is not considered.

To benefit the expressivity of Mahalanobis distance as well as overcoming the singularity and overfitting, Karimoddini et al. in [21], proposed an improvement on Eq. (2.13) to overcome the over fitting problem as well as to keep the ability of GK-clustering algorithm. The main idea is that the scaling parameter needs to be increased when the number of data in a cluster is too small to cope with the overfitting problem while it is necessary to be decreased when the cluster is rich enough to keep the effect of covariance matrix in Mahalanobis distance. Therefore, the scaling parameter is required to be updated subject to the value of contribution of data in each cluster as

$$\rho_i = \frac{\sum_{i=1}^{N} u_{ij}}{N} \times 100 \tag{2.14}$$

where ρ_i represents the value of contribution (density of presence or probability) of data in i-th cluster, u_{ij} denotes membership value of j-th data in i-th cluster and N is number of all data.

Now, Eq. (2.13) can be modified as

$$F_i^{new} = (1 - \gamma^{\rho_i})F_i + \gamma^{\rho_i} \det[F_0]^{1/n} I \tag{2.15}$$

where $\gamma \in [0, 1]$ is chosen to be close to 1 (e.g. $\gamma = 0.9$). In this way, GK-clustering algorithm is facilitated to do clustering regardless of the number of data in each cluster.

3. Parameter estimation based on GK-clustering and TSK modelling

This section proposes a novel estimator for parameter estimation of K-distributed random variables (RVs) by using fuzzy TSK modelling and GK-clustering. At first, a set of v's is considered, where for each v different sets of K-distributed samples are randomly generated and their respective means and variances are calculated. The set of all secondary variables (means and variances) together with v's are then stored. A fuzzy TSK modelling is then used to model the relationship between v's and corresponding means and variances. The GK-clustering is used to obtain the membership functions of the antecedent part and least square method is used to estimate the parameters of the consequents part of the fuzzy rules in a super matrix. Each page of this super matrix contains the information of the respective v. The model is trained and the matrix of consequent parameters and the matrix of membership function centres are stored to be used later in the test stage. In the testing stage, a different set of K-distributed random variables are considered and their means and variances are calculated. Each v is then estimated based on the train matrices, and the estimation error is calculated. Finally, the accuracy and the performance for the proposed method are compared with the existing methods.

The K-distributed samples are randomly generated [4,22] by using

$$x = 2a\sqrt{G_1 G_v} \tag{3.1}$$

where $G_v = \Gamma(v, 1, m, n)$ denotes gamma random variables (RV) with shape parameter ρ, scale parameter 1 and size of $G_v \in \mathrm{R}^{m\times n}$. Then, the respective means and variances of each set of K-distributed RVs are calculated as the secondary data to be fed as inputs to the estimator. Fuzzy TSK modelling is used to model the relationship between v's and corresponding means and variances. The mean and variance are considered as inputs of fuzzy model, since these two statistics are related to v, monotonically. Moreover, they are simple to calculate and well-known to be used.

A typical rule of TSK model can be represented as

$$R^i : \text{If } m \in MF^i_{m_j} \text{ and } q \in MF^i_{q_k} \text{ Then } v_i = A_i m + B_i q + C_i \tag{3.2}$$

where m is mean, q is variance, $MF^i_{m_j} and\ MF^i_{q_k}$ are the j-th and k-th membership functions of mean and variance fuzzy variables in the i-th rule, $j = 1, 2, \ldots, cm$, $k = 1, 2, \ldots, cq$ and $i = 1, 2, \ldots, R = cm \times cq$ denote the number of fuzzy sets for mean, the number of fuzzy sets for variance and the number of rules, respectively. A_i, B_i, C_i are the consequent parameters of i-th rule R^i. Furthermore,

$$v_i = X_d \theta_i \tag{3.3}$$

where

$$X_d = [m_d \ \ q_d \ \ 1] \tag{3.4}$$

$$\theta_i = [A_i \ \ B_i \ \ C_i]^T \tag{3.5}$$

For each set of crisp inputs (m,q), total output is the weighted average of individual outputs v_i and is calculated as

$$v^d = \frac{\sum_{i=1}^{R} \beta_i v_i}{\sum_{i=1}^{R} \beta_i} \tag{3.6}$$

where v_i and β_i are the fuzzy output set and applicability degree of the i-th rule, respectively. Using product as T-norm in premise and consequent parts of rules, the applicability degree which is a quantity that determines the degree of firing of the output of each rule can be derived as

$$\beta_i = MF^i_{m_j}(m) \cdot MF^i_{q_k}(q) \tag{3.7}$$

The consequents can be easily calculated by mean of matrix operations.

Now, the main issue is to take care of the accuracy of the model. If the antecedent parts can be determined based on a prior knowledge, then the error will be reduced, significantly. The prior knowledge to build the fuzzy sets in the antecedent parts can be acquired through clustering technique. GK-clustering classifies the samples not only based on the distance, but also with respect to the covariance between the variables. Each cluster centre is determined using GK-clustering while the individual width of membership functions is designed directly using the centroid matrix. The range of the field data is also considered in the design of the membership functions to improve the accuracy and performance. Therefore, improved GK-clustering is used to obtain the membership functions (MF) of the antecedent parts.

The estimator uses triangular membership functions instead of Gaussian ones to ensure the completeness and consistency of universe of discourse. Furthermore, in reality the clusters are not distributed uniformly and symmetrically. The triangular membership functions can easily express asymmetric distributions. Moreover, unlike Gaussian functions, triangular functions can converge to zero in finite distance, representing clear boundaries for adjacent clusters. Therefore, the width of membership functions can be calculated accurately using the adjacent cluster centres. The first and the last clusters would be illustrated using trapezoidal membership functions instead of triangular ones in order to cover the data smaller (greater) than the first (last) cluster centre. The completed model is now ready to be trained.

In training stage, the parameter matrix of the consequent part (θ) can be obtained by using least square method in order to test the algorithm and consequently to estimate the parameter v in the real field.

Eq. (3.6) can be represented as

$$v^d = \frac{\sum_{i=1}^{R} \beta_i v_i}{\sum_{i=1}^{R} \beta_i} = \sum_{i=1}^{R} \frac{\beta_i}{\sum_{i=1}^{R} \beta_i} v_i = \sum_{i=1}^{R} \varphi_i(X_d) v_i \tag{3.8}$$

By substituting v_i from Eq. (3.3) in Eq. (3.8), the d-th train v is obtained as

$$v^d = \sum_{i=1}^{R} \varphi_i(X_d) X_d \theta_i \tag{3.9}$$

Therefore, the results of N train can be expressed as

$$v = X_\varphi \theta \tag{3.10}$$

where

$$v = [v^1 \; v^2 \; \cdots \; v^N]^T \in \mathrm{R}^{N\times 1} \tag{3.11}$$

$$X_\phi = \begin{bmatrix} \phi_1(X_1)X_1 & \phi_2(X_1)X_1 & \dots & \phi_R(X_1)X_1 \\ \phi_1(X_2)X_2 & \phi_2(X_2)X_2 & \dots & \phi_R(X_2)X_2 \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(X_N)X_N & \phi_2(X_N)X_N & \dots & \phi_R(X_N)X_N \end{bmatrix} \in \mathrm{R}^{N\times R} \tag{3.12}$$

$$\theta = [\theta_1 \; \theta_2 \cdots \; \theta_R]^T \in \mathrm{R}^{R\times 1} \tag{3.13}$$

where N is the number of training data (m_d and q_d; the mean and variance of d-th train samples, $d = 1, 2, \dots, N$) and R is number of rules. Therefore, the consequent parameters of TSK model are the entries of matrix θ which are calculated as

$$\theta = X_\varphi^+ v \tag{3.14}$$

where, X_φ is the pseudo-inverse matrix defined as

$$X_\varphi^+ = (X_\varphi^T X_\varphi)^{-1} \cdot X_\varphi^T \tag{3.15}$$

4. Results and discussion

A comparison is made between the proposed algorithm and previous methods. Figs. 1 and 2 show the bias and variance of the estimation errors for different numbers of sample data (N=512 and 64) with scale parameter a=1 and shape parameter v ranging from 0.1 to 1.5. The figures also suggest that in the conventional methods the bias and

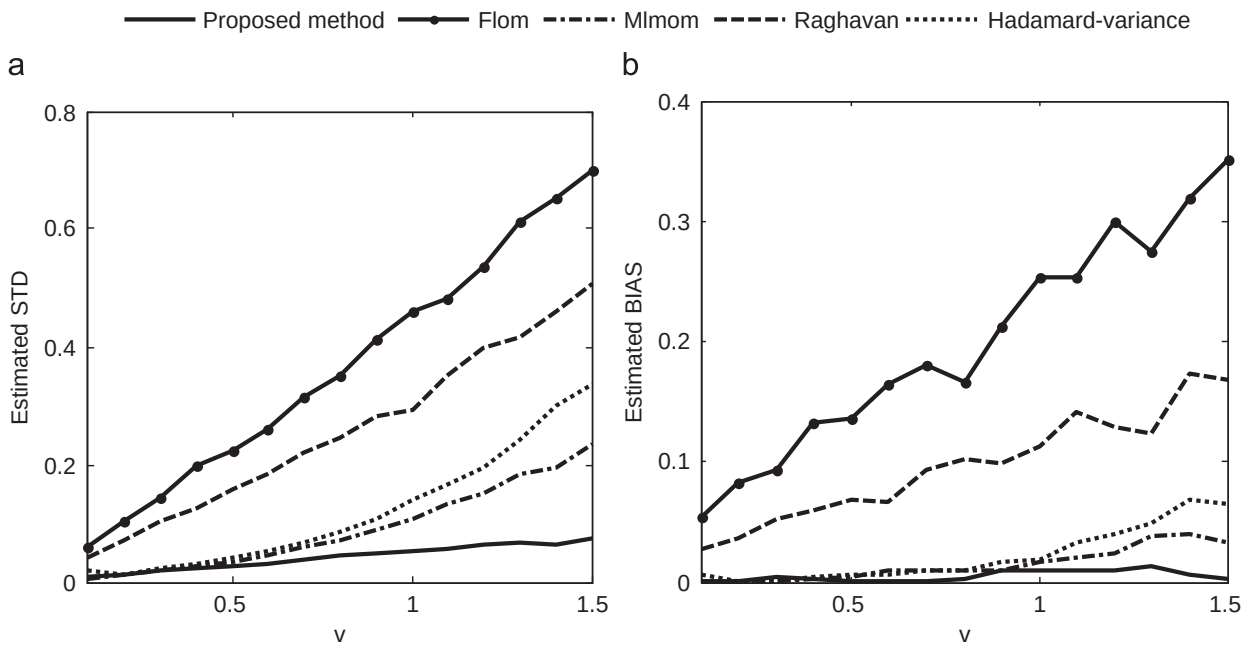

Fig. 1. Bias and STD of the estimators for N =512.

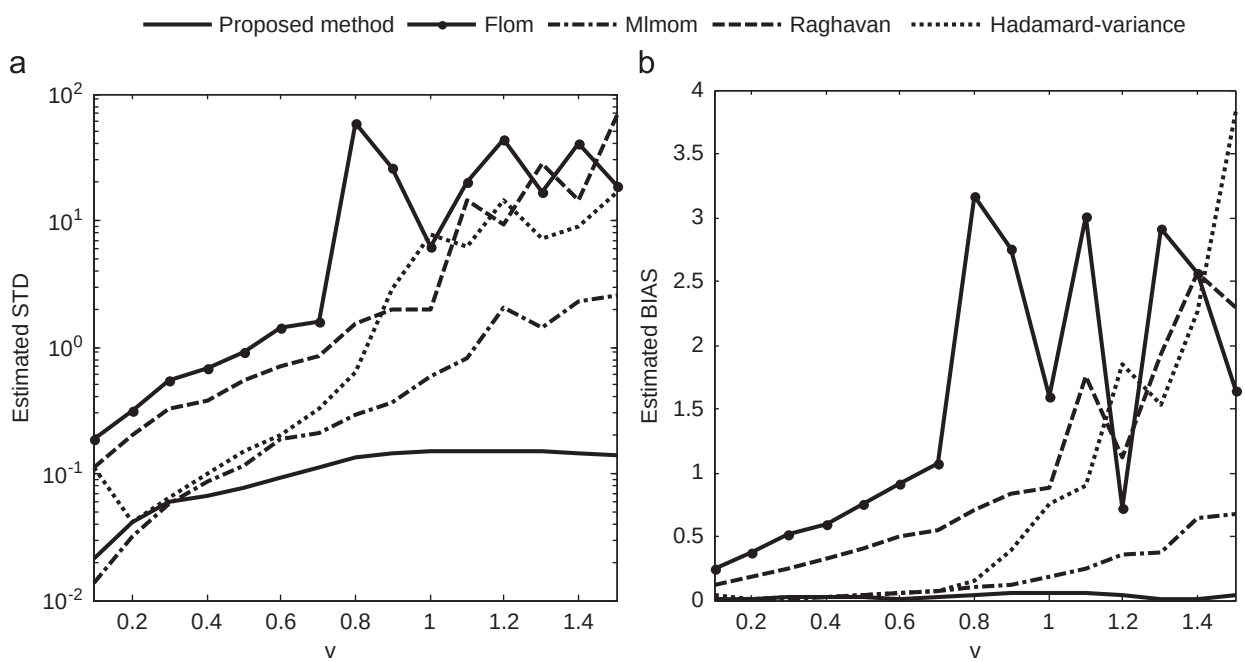

Fig. 2. Bias and STD of the estimators for N = 64.

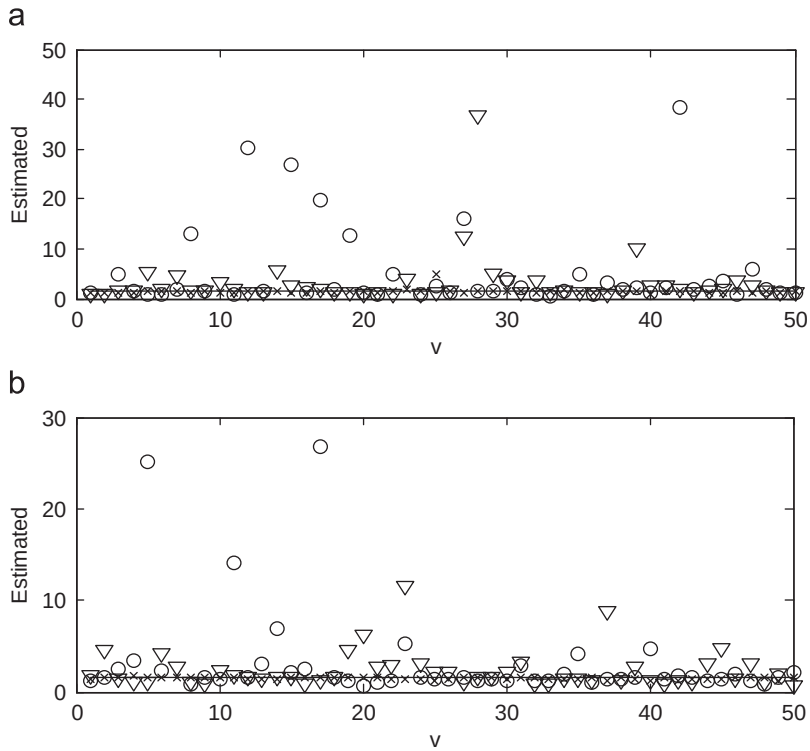

Fig. 3. (a) Independent trials of the estimates of v for N=32. (b) Independent trials of the estimates of v for N=64. ∇—Mlmom, O—Hadamard-variance, ×—proposed method.

Table 1
The comparison of the proposed method with Mlmom, Hadamard-based, Raghavan and Flom methods in terms of accuracy and execution time.

$v = 1.5 \quad N = 100$			
	Execution time (ms)	Bias	Standard deviation
Proposed method	4.8	0.021	0.33
Mlmom	3.3	0.260	1.028
Hadamard-based	2.9	1.307	6.68
Raghavan	13.6	0.947	2.95
Flom	0.8	232.6	7300

variance are dramatically increased as the number of data is decreased. The proposed method yields significantly less bias and standard deviation in comparison with the other methods especially for large v's and small number of data where the other methods fail to estimate accurately and precisely.

Fig. 3(a) and (b) illustrate the independent trials of estimates of parameter v for N=32 and 64, respectively. The estimates for 50 independent trails were successfully obtained, where in both cases v was estimated to be 1. A comparison with Mlmom and Hadamard-variance method shows that the estimates obtained by using the proposed method are always better than those of any other methods. It can also be seen that the Hadamard-variance method shows a large variability at this sample size where, in one case out of 50 trials (N=32), v was estimated as 42.

Table 1 shows the comparison of the proposed method with Mlmom, Hadamard-based, Raghavan and Flom methods in terms of accuracy and execution time for the shape parameter $v = 1.5$ and N=100 sample data. As this table suggests, the execution time of the proposed method is in the same range as two successful previous methods; Mlmom and Hadamard-based methods. Flom has the smallest execution time; however, as highlighted earlier, this method is not suitable to be used at this sample size, due to its large value for bias and variance. The proposed method is therefore

almost as fast as the previous methods with the merit of significant improvements in bias and standard deviation. For example, in comparison with the Hadamard method, the proposed method is 39.6% slower, but it has 98.7% and 95% improvement in bias and standard deviation, respectively.

5. Conclusions

One of the most important issues in maritime remote sensing and surveillance is estimation of sea clutter among radar data. It has been found that K-distribution is the most suitable PDF to model the probability of sea clutters. Many statistical methods have been introduced to estimate the K-distribution parameter and some of them have acceptable bias and variance of error. However, they need a large number of data to achieve a reasonable performance while in practical applications only a small number of samples is available.

A novel fuzzy estimation approach to overcome the problem of small number of data is discussed in this paper. The method uses TSK fuzzy modelling which utilizes GK-clustering as a pre-estimator in order to determine the antecedent membership functions and least square method to estimate the consequent parameters. GK-clustering classifies the samples not only based on the distance, but also with respect to the covariance between the variables, leading to more accurate estimation. In addition, the centres of clusters are determined directly by the method while the widths of membership functions are designed by using the centroid matrix. The range of field data is also taken into account in the design of the membership functions, aimed at improving the accuracy and performance.

The proposed method has shown significantly lower estimation bias and standard deviation especially for the high values of v parameter. It has the ability to learn from the real data and most importantly the method is capable of achieving an accurate estimation in spite of small number of data.

References

[1] E. Jakeman, P.N. Pusey, A model for non-Rayleigh sea echo, IEEE Transactions on Antennas and Propagation AP-24 (1976) 808.
[2] K.D. Ward, R.J.A. Tough, S. Watts, Sea clutter: scattering the K distribution and radar performance, United Kingdom, 2006, pp. 1–43.
[3] M.H. Marhaban, A parameter estimation method for K-distribution, IEICE Transactions on Communications E87-B (10) (2004).
[4] M.H. Marhaban, Estimation of K-distribution parameters with application to target detection, Ph.D. Thesis, University of Surrey, UK, May 2003, pp. 68–69.
[5] E. Jakeman, Speckle statistics with a small number of scatterers, Optical Engineering 23 (4) (1984) 453–461.
[6] E. Jakeman, R.J.A. Tough, Generalized K-distribution: a statistical model for weak scattering, Journal of the Optical Society of America A 4 (9) (1987) 1764–1772.
[7] D.R. Iskander, A.M. Zoubir, B. Boashash, A method for estimating the parameters of the K distribution, IEEE Transactions on Signal Processing 47 (4) (1999) 1147–1151.
[8] T.K. Moon, W.C. Stirling, Mathematical Methods and Algorithm for Signal Processing, Prentice-Hall, 2000.
[9] D.R. Iskander, A.M. Zoubir, Estimation of the parameters of the K-distribution sing higher order and fractional moments, IEEE Transactions on Aerospace and Electronic Systems 35 (4) (1999) 1453–1457.
[10] R.S. Raghavan, A method for estimating parameters of K-distributed clutter, IEEE Transaction on Aerospace and Electronic Systems 27 (2) (1991) 238–246.
[11] L. Zadeh, A. Fuzzy Sets, Information and Control 8 (1965) 338–353.
[12] E.H. Mamdani, Applications of fuzzy algorithms for simple dynamic plant, Proceedings of IEE 121 (12) (1974) 1585–1588.
[13] T. Takagi, M. Sugeno, Fuzzy identification of systems and its applications to modelling and control, IEEE Transactions on Systems, Man, and Cybernetics 15 (1) (1985) 116–132.
[14] J.C. Dunn, A fuzzy relative to the ISODATA process and its use in detecting compact, well-separated clusters, Journal of Cybernetics 3 (1974) 32–57.
[15] J.C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, 1981.
[16] K. Jajuga, L1-norm based fuzzy clustering, Fuzzy Sets and Systems 39 (2001) 63–83.
[17] P.C. Mahalanobis, On the generalised distance in statistics, Proceedings of the National Institute of Sciences of India 2 (1) (1936) 49–55 Retrieved on 2008-11-05.
[18] D.E. Gustafson, W.C. Kessel, Fuzzy clustering with a fuzzy covariance matrix, in: Proceedings of IEEE CDC, San Diego, CA, USA, 1979, pp. 761–766.
[19] R. Babuska, P.J. van der Veen, U. Kaymak, Improved covariance estimation for Gustafson–Kessel clustering, in: Proceedings of the 2002 IEEE International Conference on Fuzzy Systems, FUZZ-IEEE apos, vol. 2, 2002, pp. 1081–1085.
[20] F.de A.T. de Carvalho, C.P. Tenório, N.L. Cavalcanti Jr, Partitional fuzzy clustering methods based on adaptive quadratic distances, Fuzzy Sets and Systems 157 (21) (2006) 2833–2857.
[21] A. Karimoddini, K. Salahshoor and A. Fatehi, A new approach for overcoming to overfitting problem in fuzzy GK clustering algorithm, in: ATTCE 2006, August 2006, Malaysia.
[22] M.P. Wachowiak, R. Smolikova, J.M. Zurada, A.S. Elmaghraby, Estimation of K-distribution parameters using neural networks, IEEE Transactions on Biomedical Engineering 49 (6) (2002) 617–620.