Poster 1-48

Automatic Ground Clutter Identification Technique based on Neural Networks

Tomomi Aoki*, Keiichi Yamaguchi*, Koichiro Gomi*, Takashi Murano†, Akiko Yamada†, Satoshi Kida†, Fumihiko Mizutani†, and Masakazu Wada†,

*TOSHIBA Corporation, Kawasaki, Japan

†TOSHIBA Infrastructure Systems and Solutions Corporation, Kawasaki, Japan

I. INTRODUCTION

Fast and precise detection of signs of severe weather is important for mitigating flood damage caused by torrential rainfall and damage by tornados. As a member of the Strategic Innovation Promotion Program (SIP), TOSHIBA has been developing X-band multi-parameter phased-array weather radar (MP-PAWR) with rapid 3-dimensional volume scanning. The developed radar system can make observations out to a distance of 80 km in less than 60 s by using a fan-transmission beam and multi-pencil digital reception beams. However, the radar returns of the fan-beam-type MP-PAWR are likely to be contaminated by ground clutter. Using classical methods, ground clutter can be filtered by using finite-impulse response (FIR) or infinite-impulse response (IIR) filters. As radar-signal processors have improved, adaptive-frequency-domain filters have become feasible for use in actual systems [1], [2]. These adaptive filters can not only remove clutter but also reconstruct the removed precipitation echo. However, these filters should not be applied to data not contaminated by clutter; therefore, high-accuracy techniques to identify ground clutter are necessary.

Clutter should be identified automatically according to the weather condition, taking into account both the anomalous propagation of clutter and how the strength of the relation between clutter and precipitation changes with weather conditions. Many automatic identification algorithms have been proposed. Most use statistical features such as clutter-to-signal-power ratio (CSR), clutter phase alignment (CPA) [3], spectral power distribution (SPD), and spectral phase fluctuations (SPF) [4]. In one practical algorithm, clutter contamination can be automatically identified by comparing features against threshold values [5]. In the algorithm, the thresholds should be decided in advance, and adjusting the thresholds imposes a large burden. To mitigate this burden, we propose a completely automatic ground-clutter identification algorithm based on neural networks (NN). The networks are trained using MP-PAWR I/Q data, which is composed of clutter I/Q data collected in the clear-air condition and precipitation I/Q data collected from a slightly high elevation angle. After training, the validation I/Q data are fed into the networks and a ground-clutter filter is applied according to the identification result.

To quantitatively assess the performance of the proposed method, an estimation of error as a function of weather echo truth (power, velocity, and spectrum width) will be shown in this paper.

II. PROPOSED CLUTTER-IDENTIFICATION METHOD

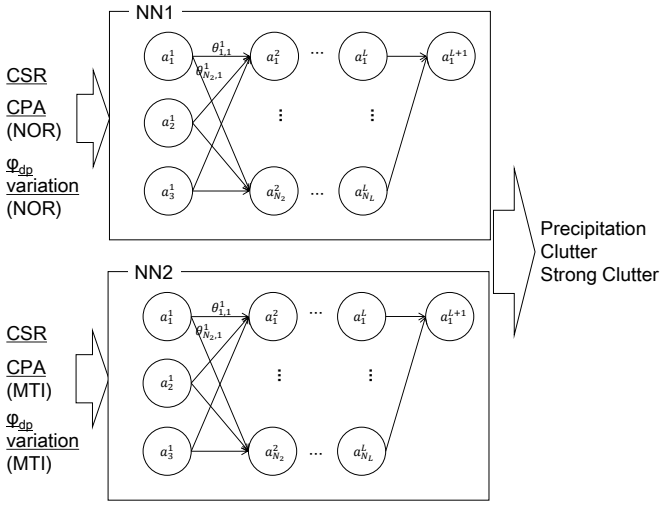


Fig. 1. Concept of proposed clutter-identification method

Figure 1 shows the concept of the proposed clutter-identification method, which is based on NNs. The number of nodes for the lth hidden layer is N_l, and the number of hidden layers is $L-1$. Every input node is connected to every hidden node, and the activation function $f(z)$ is the sigmoid function

$$f(z) = \frac{1}{1+e^{-z}}. \tag{1}$$

The proposed method uses two NNs to classify observed data into 3 classes (precipitation, clutter, or strong clutter). The final data class is decided by integrating the classification by each NN. For data classified as clutter, a clutter-mitigation processing is applied; for data classified as strong clutter, the data are assumed to be invalid because the quality of these data is so deteriorated that the clutter effects cannot be removed. In the method, each NN uses 3 features, CSR, CPA, and ϕ_{dp} variation, which matches the features used in [5]. CSR can be calculated as the ratio of the estimated-clutter signal power to the precipitation-echo power. CPA is a measure of pulse-to-pulse I/Q phase variability for a gated range.

The value of ϕ_{dp} variation is computed as the variability of the difference in ϕ_{dp} between adjacent gates. After these features from data, both before clutter mitigation (NOR) and after clutter mitigation processed data (MTI), are calculated, the features are fed into the NNs and the output is computed by forward propagation.

III. Training Datasets

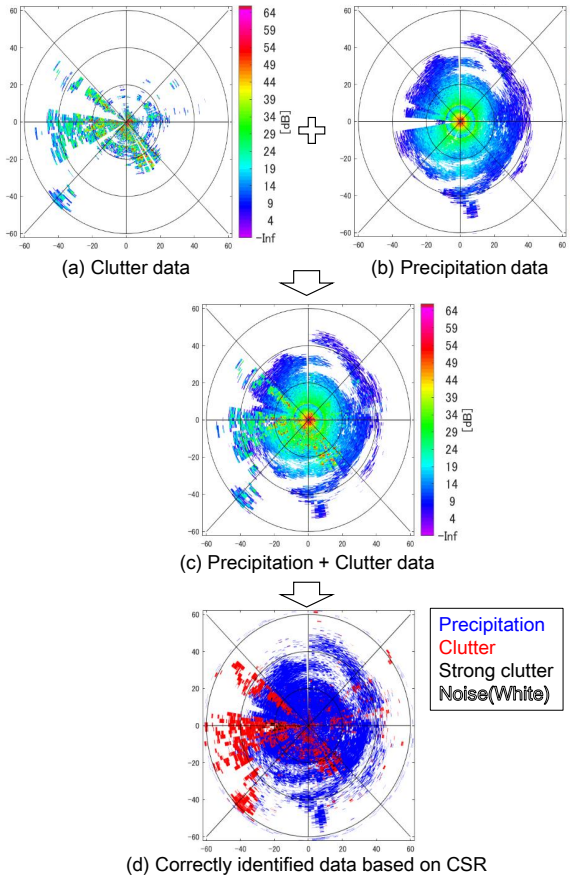


Fig. 2. Method for generating correctly identified data

TABLE I. Observation parameters

Radar frequency	9,325 – 9,445 MHz
Observation date (Clutter)	23 July 2018 09:10 (JST)
Elevation angle (Clutter)	0 deg.
PRF (Clutter)	1310 Hz
Observation date (Precipitation)	31 January 2019 18:35 (JST)
Elevation angle (Precipitation)	6 deg.
PRF (Precipitation)	1392 Hz
Number of hits	20

If real-time operation is needed, training NNs during operation is difficult. Fortunately, because the features characteristics needed to identify clutter do not change daily, the NNs can be trained before operation in our method, and we have done so.

To train NNs, it is necessary to prepare a training dataset for known conditions, including knowing for the data items whether they have been contaminated by clutter, by strong clutter, or not contaminated. In this paper, the training dataset is generated with I/Q data, which comprises clutter I/Q data collected in the clear-air condition and precipitation I/Q data collected from a slightly higher elevation angle. These data are classified into three classes on the basis of true CSR. The datasets can be generated automatically by observing conditions with and without rain. This method thus allows easy collection of big training datasets.

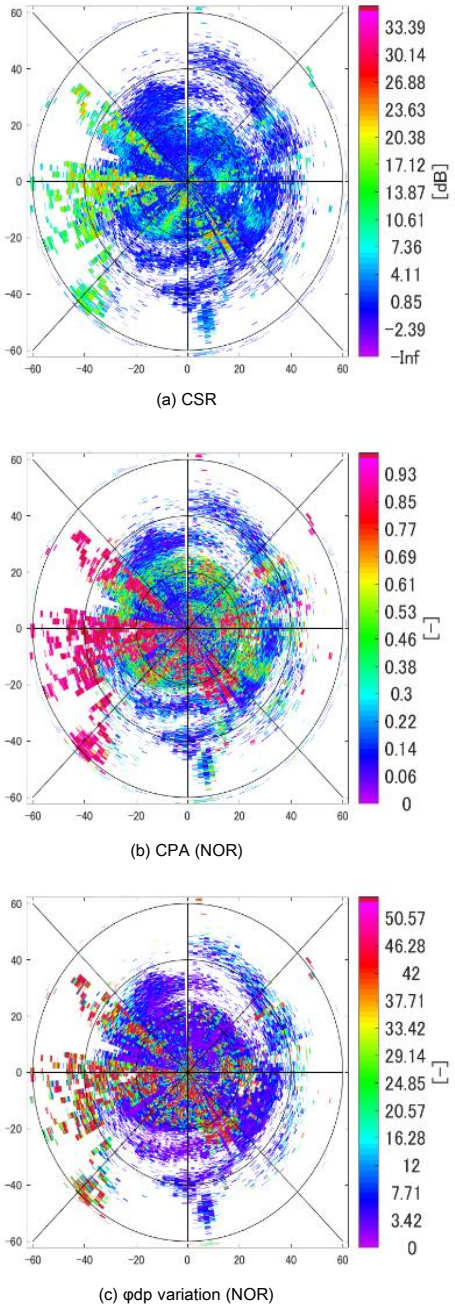
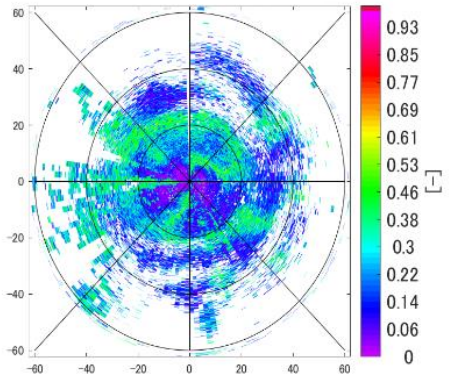
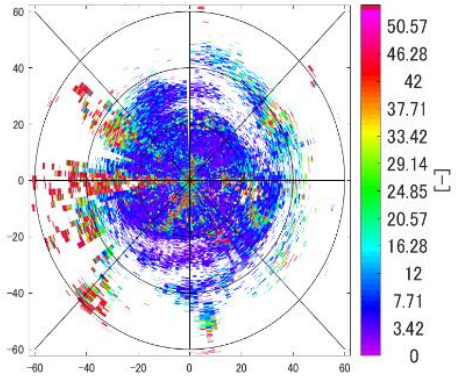


Fig. 3. Features of correctly classified data (NN 1)



(d) CPA (MTI)



(e) φdp variation (MTI)

Fig. 4. Features of correctly classified data (NN 2)

Figure 2 shows an example of generating correctly identified data. Figures 2 (a)–(c) are estimated from the signal-to-noise power ratio of clutter, precipitation, and precipitation I/Q data with clutter added. Figure 2 (d) shows the identification results. Table I lists the observation parameters used to obtain the training I/Q data. These clutter and precipitation I/Q data were collected from the MP-PAWR installed at Saitama University in Japan with the elevation angle fixed at 0° and 6° on different days. The input features were computed to generate the datasets as shown in Figs. 3 and 4. In these figures, CSR input to NN 2 is the same as that in Fig. 3 (a).

In this paper, NNs were trained by back propagation with cross entropy defined as follows.

$$J(\theta^1_{1,1},\ldots,\theta^l_{i,j},\ldots,\theta^L_{1,N_L}) = -\frac{1}{N_s}\sum_{k=1}^{N_s}\Big[y_k \ln\{f(z_1^{L+1,k})\}+(1-y_k)\ln\{1-f(z_1^{L+1,k})\}\Big] \quad (2)$$

$$z_i^{l+1,k} = \sum_{j=1}^{N_l}\theta^l_{i,j}a_j^{l,k} \quad (3)$$

In this, $\theta^l_{i,j}$ is the branch weight between the jth node of the lth layer and the ith node of the $l+1$th layer. The values N_s and y_k are, respectively, the number of samples and the (correctly identified) training data of the kth sample; N_l is the total number of nodes in the lth layer; $a_j^{l,k}$ is the jth node output of the lth layer for the kth sample.

IV. Performance Evaluation

The proposed method is evaluated by comparing its identification accuracy and estimation error against precipitation-echo truth with the accuracy and error levels obtained when using the algorithm described in [5]. The observation parameters are the same between methods (see Table I). To generate validation datasets different from the training datasets, the azimuth of the precipitation I/Q data is rotated, and the clutter and azimuth-rotated precipitation I/Q data are added together.

A. Identification accuracy

Figure 5 (a) shows the correct identifications, and Figs. 5 (b) and (c) show the identifications as made by the proposed method and the method of [5]. The results confirm that the proposed method classifies data well. By contrast, the method of [5] identifies most data belonging in the clutter class as being in the strong clutter class. Consequently, most clutters are invalidated by the method of [5] despite the fact that clutter-mitigation algorithms could remove clutter effects from those data. Note that, in this evaluation, we used the thresholds given in [5] for evaluation of that method; these values were optimized for a different radar system than we used, which may explain the poor performance of the method given in [5].

To evaluate the identification performance of our proposed method, the probability of detection (POD) and false-alarm ratio (FAR) were computed. Table II gives the resulting identification matrix. In the table, TP(P) is the number of precipitation data items that the classifier correctly judged to be precipitation; FC(P) is the number of precipitation data items that the classifier incorrectly classified as either clutter of strong clutter. Occupancy is the ratio, for each class, between the number of items that are correctly in the class and the total size of data. The values of POD and FAR for each class are defined as below.

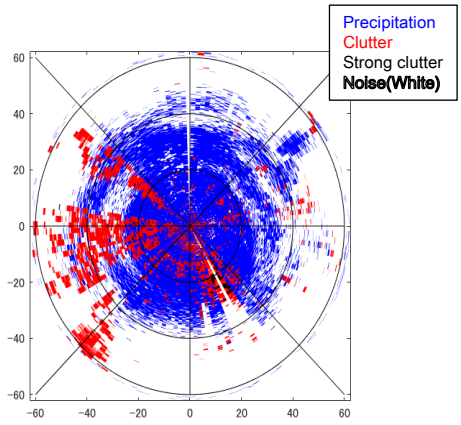
TABLE II. Matrix of identification results

		Estimation results			Occupancy
		Precipitation	Clutter	Strong clutter	
True values	Precipitation	TP(P)	FC(P)	FS(P)	74.7%
	Clutter	FP(C)	TC(C)	FS(C)	23.8%
	Strong clutter	FP(S)	FC(S)	TS(S)	1.5%

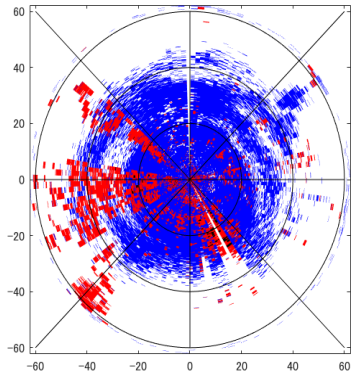
$$POD_P = \frac{TP(P)}{TP(P)+FC(P)+FS(P)} \quad (4)$$

$$POD_C = \frac{TC(C)}{FP(C)+TC(C)+FS(C)} \quad (5)$$

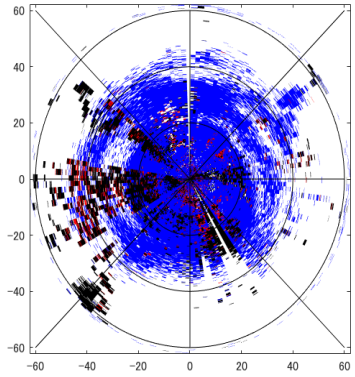
$$POD_S = \frac{TS(S)}{FP(S)+FC(S)+TS(S)} \quad (6)$$



(a) Correctly identified result



(b) Estimated result by proposed method



(c) Estimated result by [5]

Fig. 5. Identification results

$$FAR_P = \frac{[FP(C)+FP(S)]}{TP(P)+FP(C)+FP(S)} \quad (7)$$

$$FAR_C = \frac{[FC(P)+FC(S)]}{FC(P)+TC(C)+FC(S)} \quad (8)$$

$$FAR_S = \frac{[FS(P)+FS(C)]}{FS(P)+FS(C)+TS(S)} \quad (9)$$

The accuracies of identifying our validation data by the proposed method and the method of [5] are shown in Table III. Focusing on the PODs and FARs of the precipitation class, the methods had accuracies of 98.6% (the proposed method) and 89.6% (the method of [5]), with corresponding FARs of only 1.5% and 1.9%, respectively. That is, both methods are highly accurate. In contrast with classification into the precipitation class, the identification accuracies (POD and FAR) for clutter are high for the proposed method and low for the method of [5]. The reason the accuracy of identifying clutter by the method of [5] is low is because most clutter data are identified as being strong clutter data. Finally, for the strong clutter class, the method of [5] has a very high POD, but this is because it wrongly identifies clutter as strong clutter, as reflected in the high value of FAR. These results suggest that precipitation-echo data overlaid with clutter will be removed at an excessively high level by the method of [5]. With the proposed method, for the strong clutter class, the POD is 42.0% and the FAR is 29.8%. This is because the number of strong clutter items is low, and the classifier preferentially classifies items into the largest class. From the above, although the accuracy of the proposed method identified for strong clutter is not high, we believe it should be possible to reduce the number of wrongfully invalidated data.

TABLE III. ACCURACY OF CLASSIFICATION RESULTS

Class	Proposed method/Method of [5]	
	POD [%]	FAR [%]
Precipitation	98.6/89.6	1.5/1.9
Clutter	94.4/17.2	7.7/34.1
Strong clutter	42.0/98.5	29.8/93.7

B. Estimation error relative to precipitation truth

Figure 6 shows a PPI image of the power estimation. Figure 6 (a) is the PPI image as calculated with the I/Q data of precipitation and (unmitigated) clutter added, and, Fig. 6 (b) is the image as calculated from I/Q data from precipitation only. Figures 6 (c) and (d) show the power-estimation results with clutter mitigation and clutter invalidation applied according to the identification results.

Comparing Fig. 6 (a) with Fig. 6 (c), we find that the clutter located at the lower-right quadrant and left of center at a range from 20–50 km can be removed. Moreover, comparing Figs. 6 (b) and (c), the result when using the proposed method is similar to the precipitation truth. Focusing on Fig. 6 (d) shows that most clutter can be removed; however, the precipitation data located right of center at a range of 0–20 km and left of center at a range of 10–40 km have disappeared (i.e., been invalidated). To assess the amount of wrongfully removed precipitation data, the removal ratio of precipitation data P_{remove} is defined as

$$P_{remove} = \frac{N_{remove}}{N_{precip}}. \quad (10)$$

In this, N_{remove} is the number of removed precipitation data items and N_{precip} is the total number of data items that are true precipitation echo. Because P_{remove} was 3.8% with the proposed method and 30.8% with the method of [5], the method of [5] seems to have

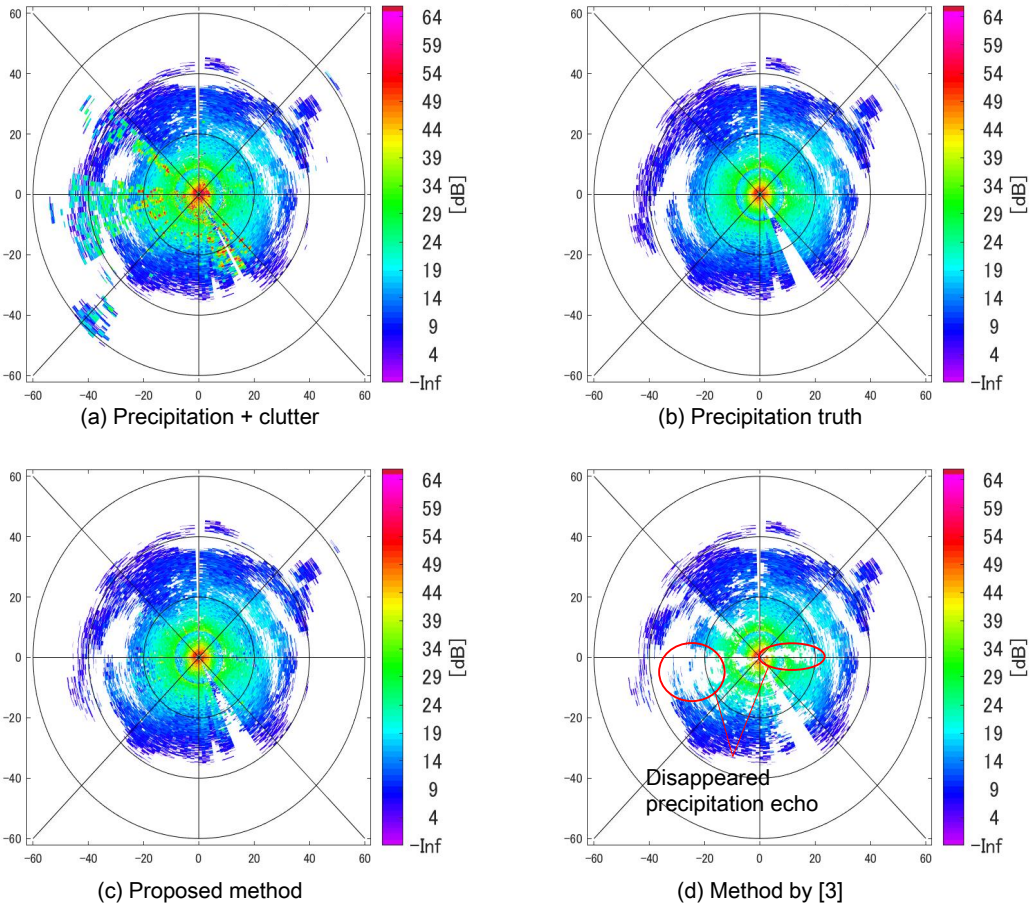


Fig. 6. Power-estimation results

been too aggressive, removing an excessive amount of precipitation-echo data.

In this paper, to quantitatively evaluate the performance of the proposed method, the cumulative distribution function (C.D.F.) of the estimation error relative to the precipitation-echo truth (power, velocity, and spectrum width) were computed. Note that because the denominator of the cumulative probability is N_{precip}, the maximum value of the C.D.F. is $1 - P_{remove}$.

The C.D.F.s of power, velocity, and spectrum width are illustrated in Fig. 7. The blue and red lines show the C.D.F.s of the proposed method and the method of [5], respectively. As illustrated in Fig. 7, all blue lines are bigger than the corresponding red lines. This means the number data with low error data are higher for the proposed method than for the method of [5]. As a specific example, by applying the proposed method, the C.D.F. of the power error at 1 dB was improved by 13 percentage points from the C.D.F. obtained for the method of [5]. For velocity and spectrum width, the C.D.F. differences at 1 m/s were 19 percentage points, respectively.

V. Conclusion and Future Work

In conclusion, we developed a new clutter-identification method based on NNs. Our proposed method uses two NNs to classify input data class into three classes (precipitation, clutter, and strong clutter).

Evaluation of the meteorological data observed by MP-PAWR confirmed that the proposed method effectively improved the accuracy of identification and estimation relative to the accuracy of the conventional identification method when using the same features. In particular, the proposed method reduced misclassification of weak clutter as strong clutter, allowing it to avoid over-zealous removal of precipitation echo overlaid with weak clutter. By evaluating the estimation error against the precipitation ground truth data, we found that the amount of data for which the estimation error was low could be increased by the proposed method.

In future work, we will test whether the proposed method can work correctly in various locations and with different weather conditions by using data collected at other radar sites and under other weather conditions. In addition, we will develop this method to identify additional classes (e.g., biological clutter, interference from other transceivers).

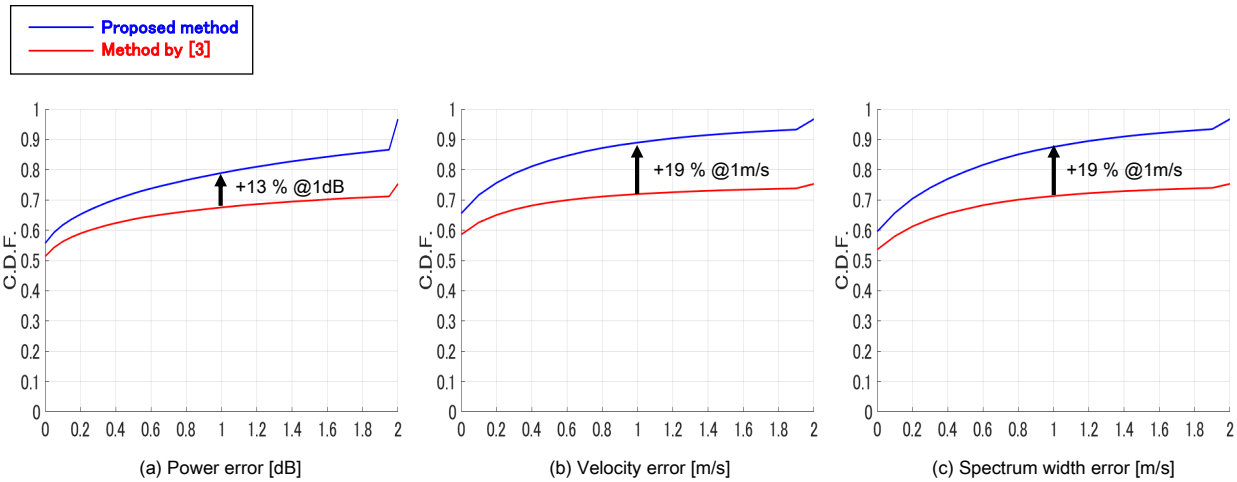


Fig. 7. Cumulative density function of estimation error

ACKNOWLEDGMENT

This work was supported by the Council for Science, Technology and Innovation (CSTI)'s Cross-ministerial Strategic Innovation Promotion Program (SIP), "Enhancement of societal resiliency against natural disasters" (Funding agency : Japan Science and Technology Agency).

REFERENCES

[1] A. Siggia and R. Passarelli, "Gaussian Model Adaptive Processing (GMAP) for Improved Ground Clutter Cancellation and Moment Calculation," *Proc. of ERAD*, p. 67–73, 2004.

[2] S. Torres and D. Warde, "Ground Clutter Mitigation for Weather Radars Using the Autocorrelation Spectral Density," *J. Atmos. Oceanic Technol.*, vol. 31, no. 10, p. 2049–2066, Oct. 2014.

[3] J. Hubbert, M. Dixon, S. Ellis, and G. Meymaris, "Weather Radar Ground Clutter. Part I: Identification , Modeling, and Simulation," *J. Atmos. Oceanic Technol.*, vol. 26, no. 7, p. 1165–1180, July 2009.

[4] Y. Li, G. Zhang, R. Doviak, L. Lei, and Q. Cao, " A New Approach to Detect the Ground Clutter Mixed with Weather Signals, '' *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 4, p. 2373–2387, April 2013.

[5] H. Yamauchi, "Quality control of weather radar data by using dual-polarization," WMO/ASEAN Training Workshop on Weather Radar Data Quality and Standardization, 2.4 Weather radar data Quality Control PDF3, Feb. 2018.