

# Flood Damage Assessment for Cultivated Lands: BVDA based Adaptive Thresholding and Local Statistics

N. Gökhan Kasapoğlu  
Dept. of Physics and Technology  
University of Tromsø  
Tromsø, Norway  
gokhan.kasapoglu@uit.no

Verda Kocabas  
Research and Development  
BlackBridge Geomatics  
Lethbridge, Canada  
verda.kocabas@blackbridge.com

**Abstract**—Determining the land lost due to excess moisture for a harvest season is very important for the farmers. Quick and easy way of doing that is to utilize satellite imagery which provides accurate moisture damage calculations. In this study, an automated analysis technique is developed to process SPOT 5 Panchromatic data for delineating excess moisture on cultivated lands. The proposed method was validated with the maps that were manually digitized by imagery technicians. A total of six different SPOT 5 Panchromatic scenes acquired from May to July in 2011 over six different farming regions in Alberta and Saskatchewan, Canada were used in excess moisture damage assessment experiments. The preliminary results show that using BVDA as a dynamic threshold method resulted in more consistent and accurate results than CFAR. Incorporating local statistics with the global ones decreases the false alarm rate and increases the robustness of the results. Excess moisture layer over cultivated lands that is produced with the proposed methodology can be useful to farmers applying for crop insurance programs which might result in potential payments for unseeded acres.

**Keywords**— *Border vector detection and adaptation (BVDA); classification; SPOT HRG; excess moisture*

## I. INTRODUCTION

After a heavy rain and storm, small and shallow water bodies appear as well as slurry soil on the cultivated lands. Therefore, two extreme target spectral characteristics need to be detected for proper damage assessment. Water body generally absorbs most of the irradiation in the near-infrared (NIR) part of the electromagnetic spectrum and appears dark, therefore it can be detected by simply applying a method of maximum thresholding [1]. On the other hand, slurry soil shows relatively higher spectral responses than water bodies. Moreover, electromagnetic radiation from background is modulated by cultivated lands with different soil types and some other land cover types which may have high surface roughness values (e.g., rocks, manmade objects). Cultivated lands which may contain different soil types can be responded within large dynamic range. This may cause large background variance and may increase false alarm rate for a conventional method of maximum thresholding. These ambiguities may be more pronounced for panchromatic images which cover larger electromagnetic spectrum (e.g., SPOT 5 Panchromatic 0.48 – 0.71  $\mu\text{m}$ ) than multispectral channels (e.g., SPOT 5 NIR 0.78 –

0.89  $\mu\text{m}$ ). Signals in a panchromatic channel represent reflected average electromagnetic radiation from imaging areas within large spectral range. For instance SPOT 5 Panchromatic channel covers both green (0.50 – 0.59  $\mu\text{m}$ ) and red (0.61 – 0.68  $\mu\text{m}$ ) bands which are in visible region of the electromagnetic spectrum [2]. However, high spatial resolution of panchromatic images (e.g. 2.5 m for SPOT 5) and its lower cost make panchromatic images an important data source for excess moisture damaged cultivated land assessment applications.

Hence, this paper aims to evaluate the use of SPOT 5 panchromatic satellite imagery and propose a unique methodology for detecting and classifying flood damage areas from panchromatic images. The proposed methodology employs unique combination of border vector detection and adaptation (BVDA) based adaptive thresholding and local statistics to map excess moisture damage zones on cultivated lands accurately and efficiently. BVDA consist of two parts. In the first part, initial border vectors are selected from the training set via special selection process called border vector detection. In the second part of the algorithm, the border vectors are adapted by moving them toward the decision boundaries. At the end of the adaptation process, the border vectors are finalized. The method next uses the minimum distance to border vector rule for target detection. Final position of the border vectors are adaptively defined thresholds for separating targets from the background. For operational requirements, the process is finalized by applying morphological filters to the BVDA outputs for smoothing noisy pixels.

Conventional target detection problems consist of separating deserved signals from background with a reasonable false alarm rate. If the global statistics of the background is not consistent for parts of a scene, then false alarm rate increases. When the false alarm rate increases, then unaffected cultivated lands may be detected as excess moisture damaged areas. To overcome this problem, local statistics as well as global statistics of the background during the damage assessment process were employed. Surface roughness and different soil types may cause this kind of large changes in the statistics of local background. Therefore, a moving window is used to

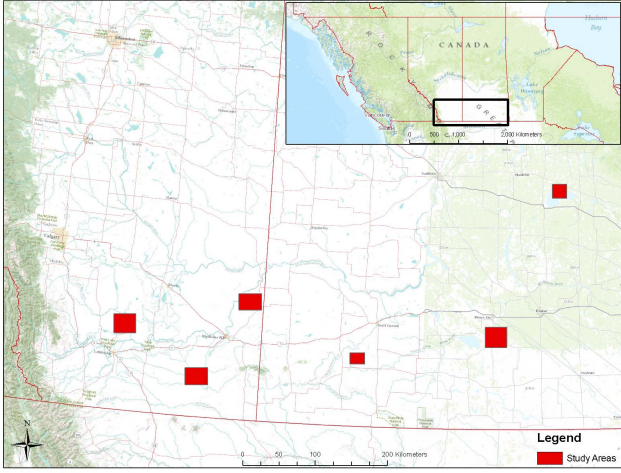


Fig. 1. Study Areas

extract local statistics of the background. Window size should not be chosen too small to extract meaningful statistical parameters or too big to miss the local changes. As a result, sizes of 7, 13, 27 and 201 pixel window are selected after trial-and-error experiments. Decisions are made for the center pixel of the moving window by using both local and global statistics of the target and background.

## II. STUDY AREA AND IMAGE DATASET

In 2011, flooding due to heavy rains has caused damage to many acres of crop, hay and posture land in Alberta, Saskatchewan and Manitoba provinces in Canada. This resulted in many acres of unseeded agricultural land. Therefore, a total of six farming regions in Alberta and Saskatchewan that were affected by the excess rainy season in 2011 were selected as study areas (Fig. 1).

SPOT 5 2.5m panchromatic image data were acquired right after the flood season. These datasets were selected for each study area from the SPOT 5 image archive. Table I outlines image data that have been used in the experiments.

## III. IMAGE PREPROCESSING

Geometric correction and atmospheric calibration are two important aspects in the image preprocessing. The datasets for each selected study area were derived from SPOT 5 2.5m panchromatic Level 1A (raw imagery) products. These image data have been processed using the most accurate control available: this includes National Road Network from GeoBase and Alberta Government Access Vectors. The control points were well distributed within and around the edges of the image.

The PCI High Resolution Satellite Ortho Package has been used to orthorectify image data. The orthorectification process computes a Rational Functions Math Model where each individual scene contains coefficients, called Rational Polynomial Coefficients (RPC). RPC's define the math model together with the collected Ground Control Points (GCPs). The procedure also uses a Digital Elevation Model (DEM) for correction and the Canadian Digital Elevation Data (CDED) at 1:50 000 scale was utilized in this step.

TABLE I. SUMMARY OF IMAGE DATA

Image Date	Image Gain	Image Sun Elevation	Instrument
05 June 2011	2.36	59.63	HRG2
05 July 2011	2.38	59.53	HRG1
16 July 2011	2.38	58.03	HRG1
02 May 2011	2.39	54.43	HRG1
15 July 2011	2.38	60.40	HRG1
24 June 2011	2.39	62.01	HRG1

The atmospheric calibration of SPOT images were conducted by converting the Digital Number (DN) values of each image to top of atmosphere reflectance (as in (1)), which normalizes the solar position and irradiance variation between images, following procedures outlined by Chander et al. [3] with sensor calibration coefficients for both SPOT 5 HRG sensors. Thus, the same classification algorithm can be applied to images acquired at different times of the year.

$$R_{\lambda} = \frac{L_{\lambda} * D^2 * \pi}{E_{sun\lambda} * \cos(\theta)} \quad (1)$$

$$L_{\lambda} = \frac{DN_{\lambda}}{A_{\lambda}} \quad (2)$$

where  $L_{\lambda}$  is the apparent at-satellite radiance for spectral band  $\lambda$ ,  $DN_{\lambda}$  is the digital number of selected band  $\lambda$ ,  $A_{\lambda}$  is the calibration factor for spectral band  $\lambda$  of SPOT image,  $R_{\lambda}$  is the calibrated reflectance,  $E_{sun\lambda}$  is exoatmospheric solar irradiance,  $D$  is the distance between the Earth and Sun, and  $\theta$  is the sun zenith angle.

## IV. METHODOLOGY

Selection of training samples is very important for target detection and classification process; thus, training samples should be representative for the problem and there should be enough number of training samples selected. However, limited number of training samples is usually available for wide range of classification problems. Therefore classifiers should produce robust decisions even with less representative samples. In this study, Gaussian maximum likelihood classifier (GMLC) is employed as a constant false alarm rate detector (CFAR) [4] for flood damaged area assessment. For a nonparametric solution, BVDA has been used [5].

### A. Gaussian Maximum Likelihood Classifier as a CFAR detector

A conventional method for target detection problems or binary classification problems is using CFAR detectors. Class specific statistical distributions should be defined properly to achieve high accuracies. Distributions can be assumed as Gaussian for optical data. Conditional probability density function (pdf) is used as discrimination rule both in the maximum likelihood classifier (MLC) and Constant false alarm rate (CFAR) detectors [4]. If the number of classes is  $m$ , then there are  $m$  discrimination functions that can be defined by using conditional probability density function as follows:

$$g_{C_i}(\bar{x}) = p(\bar{x} | C_i), i = 1..m \quad (3)$$

The label of the class which makes the discrimination rule maximum is assigned as the class of  $\bar{x}$  :

$$w = \arg \max \{g_{C_i}(\bar{x})\}, i = 1..m \Rightarrow \bar{x} \in C_w \quad (4)$$

In this approach, the classification problem is reduced to estimate some parameters which are related to probability density function (pdf). If the random variable is one-dimensional, then the Gaussian density function is given by

$$p(x | C_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp \left[ -\frac{(x-\mu_i)^2}{2\sigma_i^2} \right] \quad (5)$$

where  $\mu_i$  and  $\sigma_i^2$  are mean and variance of class  $i$ .

In remote sensing data classification problems, random variable is a vector. Assuming  $N$  dimensions, the pdf can be written in the vector form as

$$p(x | C_i) = (2\pi)^{-n/2} |\Sigma_i|^{-1/2} \exp \left\{ -\frac{1}{2} (\bar{x} - \bar{\mu}_i)^T \Sigma_i^{-1} (\bar{x} - \bar{\mu}_i) \right\} \quad (6)$$

where  $\bar{x}$  is random variable,  $\bar{\mu}_i$  the mean vector of the  $i^{th}$  class, and  $\Sigma_i$  is the covariance matrix of the  $i^{th}$  class, respectively. The Gaussian pdf is also called normal distribution and is depicted by  $N(\bar{\mu}_i, \Sigma_i)$ . The unbiased estimates of the multidimensional Gaussian pdf parameters are calculated as follows: assuming a labeled training dataset  $\{(\bar{x}_1, y_1), (\bar{x}_2, y_2), \dots, (\bar{x}_n, y_n)\}$  where the training vectors are  $\bar{x}_i \in \mathcal{H}^N, i = 1, \dots, n$  the class labels are  $y_i \in \{1, 2, \dots, m\}$ ,  $n$  is the total number of training samples, and  $m$  is the number of classes; class means are estimated as

$$\hat{\mu}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \bar{x}_j, \{\bar{x}_j | y_j = i, i = 1, \dots, m\} \quad (7)$$

where  $n_i$  is the total number of training samples for class  $i$ . The covariance matrix estimate of the  $i^{th}$  class is given by

$$\hat{\Sigma}_i = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (\bar{x}_j - \hat{\mu}_i) (\bar{x}_j - \hat{\mu}_i)^T, \{\bar{x}_j | y_j = i, i = 1, \dots, m\} \quad (8)$$

### B. Border Vector Detection and Adaptation for Adaptive Thresholding

In BVDA, feature space is partitioned by using some of the selected and adapted reference vectors from a training set [5]. Every class should be represented by enough number of training samples in the feature space. As a result, special selection procedure called *border vector detection* is performed. In this procedure class means are calculated for every class as follows:

$$\bar{m}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \bar{x}_j, \{\bar{x}_j | y_j = i, i = 1, \dots, m\} \quad (9)$$

The class centers which are the closest training samples to the class means in Euclidean distance sense are defined as follows:

$$\begin{aligned} \bar{c}_i &= \bar{x}_k, \\ &\left\{ \begin{array}{l} k = \arg \min \{D_j\}, 1 \leq i \leq m, 1 \leq j \leq n \\ D_j(\bar{m}_i, \bar{x}_j) = \|\bar{m}_i - \bar{x}_j\| = \sum_{d=1}^N \sqrt{(m_i(d) - x_j(d))^2}, \{\bar{x}_j | y_j = i\} \end{array} \right\} \end{aligned} \quad (10)$$

For every class, *reference vector selection* procedure is performed to define initial values of the border vectors. For a specific class, all class centers are used in the process; however reference vectors are selected and used only for this class. During the reference vector selection process for class  $i=q$ , every member of the training samples belonging to class  $q$  is randomly selected only once as an input. Assume that  $(\bar{x}_k, y_k = q)$  is selected. Then, the Euclidean distances calculated between this sample and the current reference set members of class  $q$  and class centers. If the closest vector to the randomly selected training sample  $(\bar{x}_k, y_k = q)$  causes wrong decision (having different label with the randomly selected training sample), then the training sample is added to the reference vector set of class  $q$ . The reference vector selection process is performed for every class by checking its all training samples. Initial values of the border vectors consist of class centers and reference vector sets. Border vector detection process produces enough number of vectors for adaptation. This procedure has been proposed by Kasapoğlu et al. in [5].

In the adaptation process, initial values of border vectors adaptively changed based on competitive learning principles as follows: Let  $\bar{x}_j$  be one of the training samples with label  $y_j$ . Assume that  $\bar{b}_w(t)$  is the nearest border vector to  $\bar{x}_j$  with label  $y_{b_w}$ . If  $y_j \neq y_{b_w}$ , then the adaptation is applied as follows:

$$\bar{b}_w(t+1) = \bar{b}_w(t) - \eta(t) \cdot (\bar{x}_j - \bar{b}_w(t)) \quad (11)$$

On the other hand, if  $\bar{b}_l(t)$  is the nearest border vector to  $\bar{x}_j$  with label  $y_{b_l}$  and  $y_j = y_{b_l}$ , then

$$\bar{b}_l(t+1) = \bar{b}_l(t) + \eta(t) \cdot (\bar{x}_j - \bar{b}_l(t)) \quad (12)$$

where  $\eta(t)$  is the learning rate and given by

$$\eta(t) = \eta_0 e^{-t/\tau} \quad (13)$$

where  $\eta_0$  is the initial learning rate,  $t$  is the iteration and  $\tau$  is the time constant.

In BVDA threshold values between classes can be defined by using adapted border vectors. These thresholds can be represented as partial linear decision boundaries which are identified by border vector positions in the feature space. For any unlabeled testing sample, labeling is performed based on the closest border vector's label in Euclidean sense.

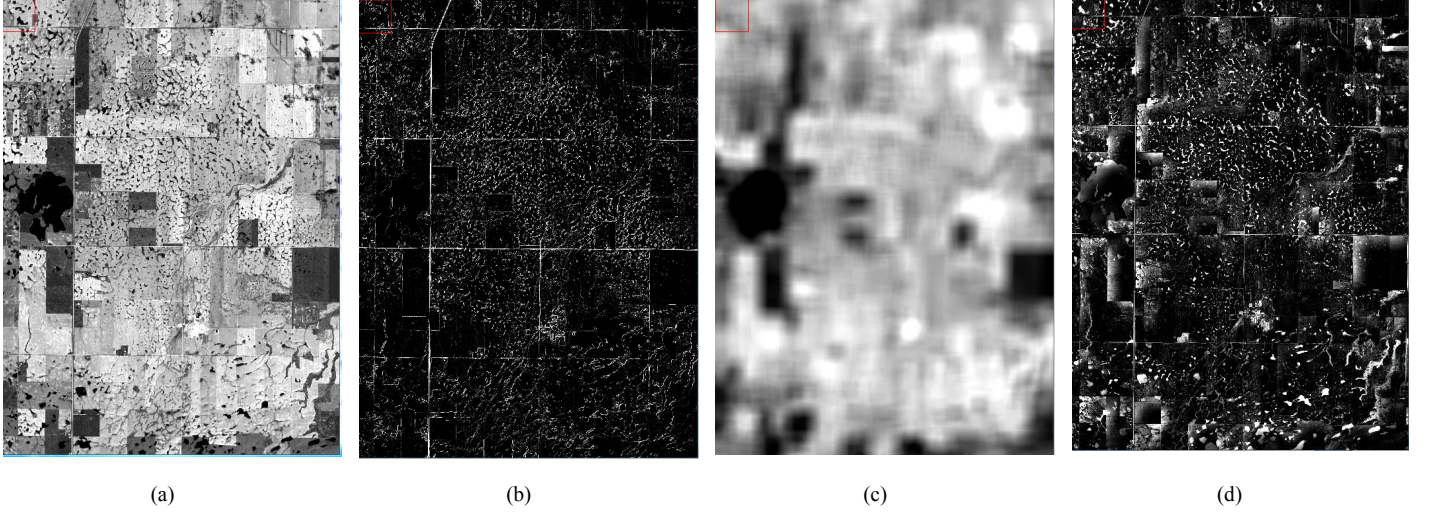


Fig. 2. Extracted features for 05 June 2011 dated data; (a) Top of the atmosphere reflectance. (b) Pixel size of 7 shifted window standard deviation.. (c) Pixel size of 201 shifted window mean. (d) Pixel size of 201 shifted window standard deviation.

### C. Use of Local Statistics

Existence of strong local changes in the background can cause detection problems of targets. Especially for flood damaged assessments for cultivated lands, these effects can be more pronounced than some other applications because of the different type of the soil existence. Separating small water bodies from the dark soil in panchromatic images can be problematic because of lack of the contrast difference. Global statistics of background cannot be sufficient to separate water bodies from dark soil. Using local statistics as well as global ones can improve detection performance and decrease false alarm rates. In this study, statistics such as local mean and local standard deviation for shifting window size of 7, 13, 27 and 201 were calculated. Some of these statistics are shown in Fig. 2 with Spot 5 panchromatic top of the reflectance data. All of the extracted features are shown in Table II.

## V. RESULTS AND CONCLUSION

The BVDA based adaptive thresholding has been performed on the image datasets summarized in Table I. Results of the Gaussian maximum likelihood based CFAR detector has also been reported for comparison. This paper only shows the results from 05 June 2011 dated data. Visual analysis of Fig. 2 (a) shows that reflected radiation from soil can vary from low values to high values related to soil types. Because of this high dynamic range of soil reflectance, some soil types can be seen very bright but some other types can be seen relatively dark. Reflectance from soil can be thought as background signal. Small water bodies and water effected damaged soil are targets and should be properly detected by the algorithms. Normally calm water bodies absorb the most part of incoming sun radiation. Thus, darkest pixels can represent water bodies. Existence of dark soil types cause some detection problem and increase false alarm rates. To overcome this, local statistics are extracted from different sizes of shifted window and employed in the proposed flood damaged assessment scheme.

TABLE II. FEATURES

<i>F.ID</i>	<i>Feature Description</i>
0	Top of the atmosphere reflectance
1	Pixel size of 7 shifted window mean
2	Pixel size of 7 shifted window standard deviation
3	Pixel size of 13 shifted window mean
4	Pixel size of 13 shifted window standard deviation
5	Pixel size of 27 shifted window mean
6	Pixel size of 27 shifted window standard deviation
7	Pixel size of 201 shifted window mean
8	Pixel size of 201 shifted window standard deviation

Analysis results of the BVDA based adaptive thresholding and the Gaussian maximum likelihood based CFAR detector have been produced by using only top of the atmosphere data and has been presented to make basic comparison between the two methods. To present the impact of local statistics, the experiment was redone with the top of the atmosphere data (F.ID 0) and pixel size of 201 shifted window mean (F.ID 7). Results have been depicted in Fig. 3. The BVDA based adaptive thresholding performs better than the Gaussian maximum likelihood based CFAR detector for only use of top of the atmosphere reflectance (F.ID. 0) (see Fig. 3 (a) and (b) for thematic maps). False alarm rate of the Gaussian maximum likelihood based CFAR detector is higher than the BVDA based adaptive thresholding for this experiment (see Fig. 4 (a) and (b) for confusion matrices). Additionally the impact of the local statistics on analysis results have been shown for the Gaussian maximum likelihood based CFAR detector and the BVDA based adaptive thresholding in Fig. 3 (c) and (d) respectively. The results reveal that using local statistics were improved classification results and decrease false alarm rates which is similar for both analysis methods. Statistically meaningful number of training samples has been used in these experiments to reach the best performance of the Gaussian maximum likelihood based CFAR detector. The results indicate that proposed the BVDA based adaptive thresholding performs well for flood damaged assessments for all cases.



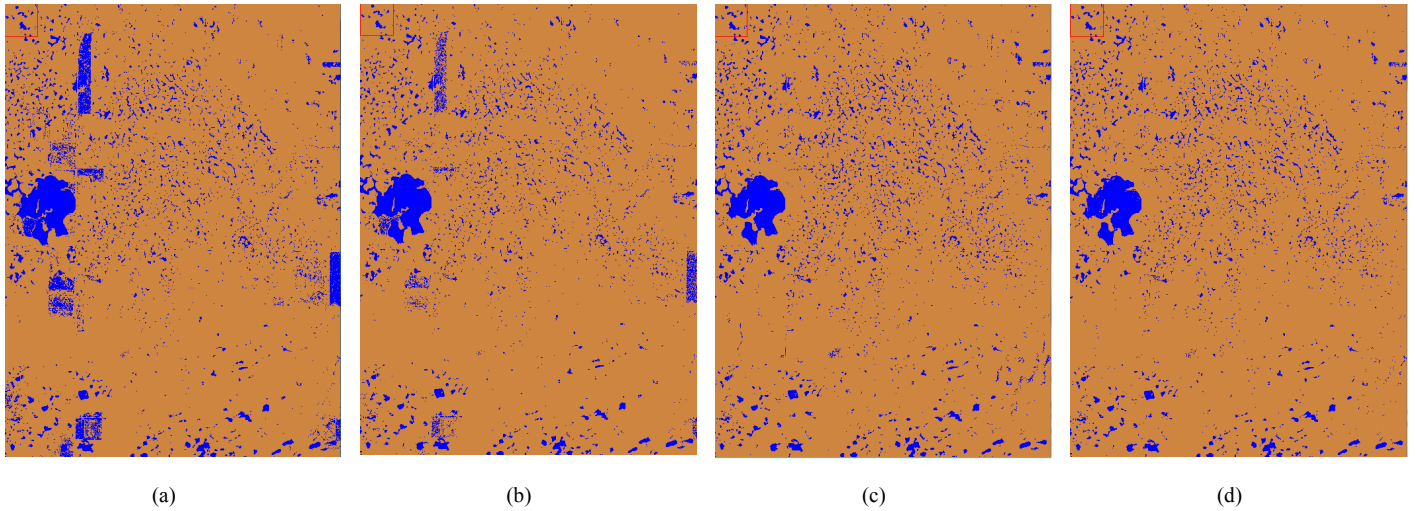


Fig. 3. Analysis results for 05 June 2011 dated data (a) The Gaussian maximum likelihood based CFAR detector for F.ID. 0 (b) The BVDA based adaptive thresholding for F.ID. 0. (c) The Gaussian maximum likelihood based CFAR detector for F.IDs. 0 and 7. (d) The BVDA based adaptive thresholding for F.ID. 0 and 7.

		1	2						
		Soil	Water		$\Sigma$	%			
1	Soil	9539	1231		10770	88.57			
2	Water	102	11175		11277	99.10			
				20714					
	$\Sigma$	9641	12406		22047		%		
	%	98.94	90.08				Overall Accuracy	93.95	
							$\kappa$	0.88	

		1	2						
		Soil	Water		$\Sigma$	%			
1	Soil	10059	711		10770	93.40			
2	Water	254	11023		11277	97.75			
				21082					
	$\Sigma$	10313	11734		22047		%		
	%	97.54	93.94				Overall Accuracy	95.62	
							$\kappa$	0.91	

(a) The confusion matrix of the Gaussian maximum likelihood based CFAR detector for testing and F.ID. 0

(b) The confusion matrix of the BVDA based adaptive thresholding for testing and F.ID. 0

		1	2						
		Soil	Water		$\Sigma$	%			
1	Soil	10770	0		10770	100.00			
2	Water	3	11274		11277	99.97			
				22044					
	$\Sigma$	10773	11274		22047		%		
	%	99.97	100.00				Overall Accuracy	99.99	
							$\kappa$	1.00	

		1	2						
		Soil	Water		$\Sigma$	%			
1	Soil	10770	0		10770	100.00			
2	Water	19	11258		11277	99.83			
				22028					
	$\Sigma$	10789	11258		22047		%		
	%	99.82	100.00				Overall Accuracy	99.91	
							$\kappa$	1.00	

(c) The confusion matrix of the Gaussian maximum likelihood based CFAR detector for testing and F.ID. 0 and 7.

(d) The confusion matrix of the BVDA based adaptive thresholding for testing and F.ID. 0 and 7.

Fig. 4. Confusion matrices for testing samples for the Gaussian maximum likelihood based CFAR and the BVDA based adaptive thresholding.

## REFERENCES

- [1] M. Bochow et al., "On the Use of Airborne Imaging Spectroscopy Data for the Automatic Detection and Delineation of Surface Water Bodies," Remote Sensing of Planet Earth, (2012), Yann Chemin (Ed.), InTech, Online at <http://www.intechopen.com/books/remote-sensing-of-planet-earth/on-the-use-of-airborne-imaging-spectroscopy-data-for-the-automatic-detection-and-delineation-of-surf> ( 12 Dec. 2012).
- [2] Astrium, Spot satellite technical data. Online at [http://www.astriumgeo.com/files/pmedia/public/r329\\_9spotsatellitetechnicaldata\\_en\\_sept2010.pdf](http://www.astriumgeo.com/files/pmedia/public/r329_9spotsatellitetechnicaldata_en_sept2010.pdf) (12 Dec. 2012).
- [3] Chander, G., Markham, B. L., & Helder, D. L. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. Remote Sensing of Environment, 113(5), 2009, 893–903.
- [4] H. Rohling, "Radar CFAR Thresholding in Clutter and Multiple Target Situations," IEEE Trans. on Aero. El. Sys., vol. AES-19, no.4, pp. 608–621, July 1983.
- [5] N. G. Kasapoğlu and O. K. Ersoy, "Border vector detection and adaptation for classification of multispectral and hyperspectral remote sensing images," IEEE Trans. Geosci. Remote Sens., vol. 45, no. 12, pp. 3880–3893, Dec. 2007.