ARMY RESEARCH LABORATORY

Statistical Methods for Analysis of Hyperspectral Anomaly Detectors

by Dalton Rosario

ARL-TR-4266 **September 2007**

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

Army Research Laboratory
Adelphi, MD 20783-1197

ARL-TR-4266 **September 2007**

Statistical Methods for Analysis of Hyperspectral Anomaly Detectors

Dalton Rosario
Sensors and Electron Devices Directorate, ARL

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED (From - To)
September 2007	Final	Fiscal Year 07

4. TITLE AND SUBTITLE
Statistical Methods for Analysis of Hyperspectral Anomaly Detectors

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Dalton Rosario

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Research Laboratory
ATTN: AMSRD-ARL- SE-SE
2800 Powder Mill Road
Adelphi, MD 20783-1128

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-TR-4266

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Laboratory
2800 Powder Mill Road
Adelphi, MD 20783-1128

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Most hyperspectral (HS) anomaly detectors in the literature have been evaluated using a few HS imagery sets to estimate the well-known ROC curve. Although this evaluation approach can be helpful in assessing detectors' rates of correct detection and false alarm on a limited dataset, it does not shed lights on reasons for these detectors' strengths and weaknesses using a significantly larger sample size. This paper discusses a more rigorous approach to testing and comparing HS anomaly detectors, and it is intended to serve as a guide for such a task. Using randomly generated samples, the approach introduces hypothesis tests for two different kinds of data: (i) idealized homogeneous samples and (ii) idealized heterogeneous samples, where model parameters can vary the difficulty level of these tests. In (i), a simulation experiment is devised to address a more generalized concern — the expected degradation of correct detection as a function of increasing noise on a given alternative hypothesis. In (ii), fundamental features of a spectral sample (magnitude and shape) are modeled separately so that strengths and weaknesses of competing detectors can be independently assessed for each feature. Additionally, detectors' ability to suppress transition of regions in the imagery is assessed in (ii).

15. SUBJECT TERMS
Hypothesis tests, hyperspectral, anomaly detection

16. Security Classification of:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Dalton Rosario
U	U	U	U	96	19b. TELEPHONE NUMBER (*Include area code*) (301) 394-4235

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Contents

List of Figures

List of Tables

INTENTIONALLY LEFT BLANK

1. Introduction

Object detection within hyperspectral (HS) images is a highly desired goal for many Army applications. HS sensors are passive sensors that simultaneously record images for many contiguous and narrowly spaced regions of the electromagnetic spectrum (*1*).* Such a dataset lends itself to the ability to search large spatial areas, ideally in an automated and timely fashion. In order for the detector to have value, it should have a high rate of correct detection and a low rate of false alarms. The detector would be based on material detection to take advantage of the high spectral resolution relative to the spatial resolution available. If the material is known, then the spectrum is also known, and the detector would compare each multivariate pixel in the HS dataset to the known spectrum. However, often the exact material of interest is not known a priori, or the number of spectra in a material of interest library is too large to search for all possible materials. Under these situations, an anomaly detector is required.

The goal of an anomaly detector is to identify statistical outliers, that is, data points that are atypical compared to the rest of the data. Multivariate models are used to define the spectral variability of the data, and the majority of the data pixels are assumed to be spectrally homogeneous and are modeled using a multivariate probability density function (PDF) with a single set of parameters. An anomalous pixel is one that has low probability of belonging to the particular PDF. Unlike classification algorithms which try to segment the pixels into homogeneous clusters, anomaly detection algorithms label each pixel as either containing an anomalous material relative to the clutter background or not.

To date, most HS anomaly detectors discussed in the literature (see, for instance, (*2*, *3*, *4*, *5*, *6*, and *7*) have been evaluated by testing a few HS image sets and estimating the well-known ROC (receiver's operational characteristics) curve in order to compare performances among these detectors. Although this evaluation approach can be helpful in assessing detectors' rates of correct detections and false alarms on a limited dataset, it does not shed light on reasons for these detectors' strengths and weaknesses.

This report discusses a more rigorous approach to testing and comparing HS anomaly detectors. The approach offers some useful properties, such as statistical hypothesis tests representing different kinds of data—potentially found in local spatial areas in HS imagery, and data specification parameters that can gradually increase the difficulty level of these tests. Having such properties in a detection-evaluation approach is desired because real HS imagery offers challenging hypotheses to anomaly detectors, and simulating these hypotheses in a meaningful way is not a trivial matter.

* In this report, literature references are referred to using italic numbers, e.g., (*1*). Equations numbers are nonitalic, e.g., equation 1.

A key problem devising such an evaluation approach lies in the fact that, in applications using real HS imagery, sample pairs representing the null hypothesis (H_0) and the alternative hypothesis (H_1) belong to an open set of infinite possibilities. For instance, let's denote the sample pair $\mathbf{X}_0^{(w_j)} = \left(\mathbf{X}_{01}^{(w_j)}, \mathbf{X}_{02}^{(w_j)}, \cdots, \mathbf{X}_{0n_0}^{(w_j)}\right)$ and $\mathbf{X}_1^{(w_j)} = \left(\mathbf{X}_{11}^{(w_j)}, \mathbf{X}_{12}^{(w_j)}, \cdots, \mathbf{X}_{1n_1}^{(w_j)}\right)$ as multivariate random samples drawn from a real (*N* by *N* by *B*) HS imagery, where $\mathbf{X}_{1h}^{(w_j)}$ $\left(h = 1, \cdots, n_1\right)$ can be interpreted as *B*-dimensional vectors forming a block of data W_1, w_j indexes the spatial location of the centroid of W_1 in the imagery's *N* by *N* area, $\mathbf{X}_{0u}^{(w_j)}$ $\left(u = 1, \cdots, n_0\right)$ can be interpreted as *B*-dimensional vectors indexed by the same w_j but forming an outer ring of data W_0 surrounding W_1 (W_0 and W_1 are mutually exclusive), (N x N) >> n_0 and (N x N) >> n_1 for >> denoting *much greater than* (e.g., for a particular application, *N = 640*, the imagery's area is *409600* square pixels, and each n_0 and n_1 is between *100* and *400*). In addition, assume that these multivariate random samples belong $(\in)$ to homogenous classes C_k $\left(k = 1, \cdots K\right)$ or to a mixture of classes $M\left(C_1, \cdots, C_K\right)$. Using these denotations and assumptions, some of the sample pair possibilities representing H_0 and H_1 are as follows:

$$
\left.\begin{array}{l}
\left.\begin{array}{rcl}
\mathbf{X}_0^{(w_1)} = \left(\mathbf{X}_{01}^{(w_1)}, \mathbf{X}_{02}^{(w_1)}, \cdots, \mathbf{X}_{0n_0}^{(w_1)}\right) & \in & \mathrm{C}_1 \\
\mathbf{X}_1^{(w_1)} = \left(\mathbf{X}_{11}^{(w_1)}, \mathbf{X}_{12}^{(w_1)}, \cdots, \mathbf{X}_{1n_1}^{(w_1)}\right) & \in & \mathrm{C}_1
\end{array}\right\} H_{01} \\
\qquad\qquad \text{and/or} \\
\left.\begin{array}{rcl}
\mathbf{X}_0^{(w_2)} = \left(\mathbf{X}_{01}^{(w_2)}, \mathbf{X}_{02}^{(w_2)}, \cdots, \mathbf{X}_{0n_0}^{(w_2)}\right) & \in & \mathrm{C}_1 \\
\mathbf{X}_1^{(w_2)} = \left(\mathbf{X}_{11}^{(w_2)}, \mathbf{X}_{12}^{(w_2)}, \cdots, \mathbf{X}_{1n_1}^{(w_2)}\right) & \in & \mathrm{C}_2
\end{array}\right\} H_{02} \\
\qquad\qquad \text{and/or} \\
\left.\begin{array}{rcl}
\mathbf{X}_0^{(w_3)} = \left(\mathbf{X}_{01}^{(w_3)}, \mathbf{X}_{02}^{(w_3)}, \cdots, \mathbf{X}_{0n_0}^{(w_3)}\right) & \in & \mathrm{M}\left(\mathrm{C}_1, \mathrm{C}_2\right) \\
\mathbf{X}_1^{(w_3)} = \left(\mathbf{X}_{11}^{(w_3)}, \mathbf{X}_{12}^{(w_3)}, \cdots, \mathbf{X}_{1n_1}^{(w_3)}\right) & \in & \mathrm{C}_1
\end{array}\right\} H_{03} \\
\qquad\qquad \text{and/or} \\
\left.\begin{array}{rcl}
\mathbf{X}_0^{(w_4)} = \left(\mathbf{X}_{01}^{(w_4)}, \mathbf{X}_{02}^{(w_4)}, \cdots, \mathbf{X}_{0n_0}^{(w_4)}\right) & \in & \mathrm{M}\left(\mathrm{C}_1, \mathrm{C}_2, \mathrm{C}_3\right) \\
\mathbf{X}_1^{(w_4)} = \left(\mathbf{X}_{11}^{(w_4)}, \mathbf{X}_{12}^{(w_4)}, \cdots, \mathbf{X}_{1n_1}^{(w_4)}\right) & \in & \mathrm{C}_1
\end{array}\right\} H_{04} \\
\qquad\qquad \vdots \\
\left.\begin{array}{rcl}
\mathbf{X}_0^{(w_L)} = \left(\mathbf{X}_{01}^{(w_L)}, \mathbf{X}_{02}^{(w_L)}, \cdots, \mathbf{X}_{0n_0}^{(w_L)}\right) & \in & \mathrm{M}\left(\mathrm{C}_1, \mathrm{C}_2, \mathrm{C}_3, \cdots, \mathrm{C}_\mathrm{K}\right) \\
\mathbf{X}_1^{(w_L)} = \left(\mathbf{X}_{11}^{(w_L)}, \mathbf{X}_{12}^{(w_L)}, \cdots, \mathbf{X}_{1n_1}^{(w_L)}\right) & \in & \mathrm{M}\left(\mathrm{C}_1, \mathrm{C}_2, \mathrm{C}_3, \cdots, \mathrm{C}_\mathrm{K}\right)
\end{array}\right\} H_{0L}
\end{array}\right\} H_0
$$

$$
\left.\begin{array}{l}
\left.\begin{array}{l}
\mathbf{X}_0^{(w_{L+1})} = \left(\mathbf{X}_{01}^{(w_{L+1})}, \mathbf{X}_{02}^{(w_{L+1})}, \cdots, \mathbf{X}_{0n_0}^{(w_{L+1})}\right) \in \mathrm{C}_1 \\
\mathbf{X}_1^{(w_{L+1})} = \left(\mathbf{X}_{11}^{(w_{L+1})}, \mathbf{X}_{12}^{(w_{L+1})}, \cdots, \mathbf{X}_{1n_1}^{(w_{L+1})}\right) \in \mathrm{C}_{\mathrm{K}+1}
\end{array}\right\} H_{11} \\
\qquad \text{and/or} \\
\left.\begin{array}{l}
\mathbf{X}_0^{(w_{L+2})} = \left(\mathbf{X}_{01}^{(w_{L+2})}, \mathbf{X}_{02}^{(w_{L+2})}, \cdots, \mathbf{X}_{0n_0}^{(w_{L+2})}\right) \in \mathrm{C}_1 \\
\mathbf{X}_1^{(w_{L+2})} = \left(\mathbf{X}_{11}^{(w_{L+2})}, \mathbf{X}_{12}^{(w_{L+2})}, \cdots, \mathbf{X}_{1n_1}^{(w_{L+2})}\right) \in \mathrm{C}_{\mathrm{K}+2}
\end{array}\right\} H_{12} \\
\qquad \text{and/or} \\
\left.\begin{array}{l}
\mathbf{X}_0^{(w_{L+3})} = \left(\mathbf{X}_{01}^{(w_{L+3})}, \mathbf{X}_{02}^{(w_{L+3})}, \cdots, \mathbf{X}_{0n_0}^{(w_{L+3})}\right) \in \mathrm{M}\left(\mathrm{C}_1, \mathrm{C}_2, \mathrm{C}_3\right) \\
\mathbf{X}_1^{(w_{L+3})} = \left(\mathbf{X}_{11}^{(w_{L+3})}, \mathbf{X}_{12}^{(w_{L+3})}, \cdots, \mathbf{X}_{1n_1}^{(w_{L+3})}\right) \in \mathrm{C}_{\mathrm{K}+1}
\end{array}\right\} H_{13} \\
\qquad \text{and/or} \\
\left.\begin{array}{l}
\mathbf{X}_0^{(w_{L+4})} = \left(\mathbf{X}_{01}^{(w_{L+4})}, \mathbf{X}_{02}^{(w_{L+4})}, \cdots, \mathbf{X}_{0n_0}^{(w_{L+4})}\right) \in \mathrm{M}\left(\mathrm{C}_1, \mathrm{C}_2, \mathrm{C}_3\right) \\
\mathbf{X}_1^{(w_{L+4})} = \left(\mathbf{X}_{11}^{(w_{L+4})}, \mathbf{X}_{12}^{(w_{L+4})}, \cdots, \mathbf{X}_{1n_1}^{(w_{L+4})}\right) \in \mathrm{M}\left(\mathrm{C}_1, \mathrm{C}_2, \mathrm{C}_3, \mathrm{C}_{\mathrm{K}+1}\right)
\end{array}\right\} H_{14} \\
\qquad \text{and/or} \\
\qquad \vdots \\
\left.\begin{array}{l}
\mathbf{X}_0^{(w_V)} = \left(\mathbf{X}_{01}^{(w_V)}, \mathbf{X}_{02}^{(w_V)}, \cdots, \mathbf{X}_{0n_0}^{(w_V)}\right) \in \mathrm{M}\left(\mathrm{C}_1, \mathrm{C}_2, \mathrm{C}_3, \cdots, \mathrm{C}_{\mathrm{K}}\right) \\
\mathbf{X}_1^{(w_V)} = \left(\mathbf{X}_{11}^{(w_V)}, \mathbf{X}_{12}^{(w_V)}, \cdots, \mathbf{X}_{1n_1}^{(w_V)}\right) \in \mathrm{M}\left(\mathrm{C}_1, \mathrm{C}_2, \mathrm{C}_3, \cdots, \mathrm{C}_{\mathrm{K}}, \mathrm{C}_{\mathrm{K}+1}, \mathrm{C}_{\mathrm{K}+2}\right)
\end{array}\right\} H_{1V}
\end{array}\right\} H_1
$$

where $w_1 \neq w_2 \neq \cdots \neq w_L \neq w_{L+1} \neq \cdots \neq w_V$ as they index the centroid of W_1 at different spatial locations in the imagery where the particular $H_{0i}\ (i = 1, \cdots, L)$ or $H_{1v}\ (v = 1, \cdots, V)$ occur under H_0 or H_1, respectively.

The total number of classes *K* under H_0 depends on the scene's heterogeneity and on the sizes of W_1 and W_0 (especially on the size of W_0), where the sizes of W_1 and W_0 are n_1 and n_0, respectively. Since real HS data are typically heterogeneous, larger sizes of W_1 and W_0 imply a larger number of different classes *K*. In the target application, n_0 and n_1 are typically between *100* and *400* pixels yielding a *K* in W_0 between *3* and *6*.

Notice that under H_1, classes $\mathrm{C}_{\mathrm{K}+1}$ and $\mathrm{C}_{\mathrm{K}+2}$ constitute the only anomalous classes relative to classes under H_0 and, therefore, they are the desired *targets* in this discussion.

In the target application, many of the sample pair possibilities shown above can occur and, therefore, could be used to represent entire hypotheses, H_0 and/or H_1. If we were to conduct simulation experiments using multiple of those possibilities, however, it would be difficult to separate subtle performance differences between detectors. So, in order to focus this study, a few of those possibilities will be chosen to model some of the important challenges an anomaly detector may face in the course of training and testing large HS imagery, and idealized samples will be generated to allow various challenges to appear spontaneously as a detector tests for these

samples. In this report, an idealized multivariate sample is defined as being independent, identically distributed (iid), multivariate normally distributed (Gaussian) sample having both mean and covariance fully specified.

The remainder of this report is organized as follows: section 2 addresses the power test of two competing anomaly detectors using idealized homogeneous pseudo samples, and added noise; section 3 addresses the power of the test of the same competing detectors using idealized pseudo samples of mixtures; section 4 addresses the power of the test of multiple competing anomaly detectors using idealized pseudo cubes, which mimics multispectral cubes of increasing background clutter complexity; and section 5 concludes this report.

2. Statistical Method Using Idealized Homogeneous Multivariate Samples

This section illustrates a statistical method using idealized homogeneous multivariate samples, where simple sets of null and alternative hypotheses are specified with added noise. For two simulation experiments discussed in this section, we estimate parameters (mean and covariance) using real spectral data and use these estimates to generate ideal simulated samples for both experiments. To observe relative strengths and weaknesses of different anomaly detection techniques using this method, two competing anomaly detectors were selected for this study: an univariate based anomaly detector (AsemiP) (*7*) and the industry standard multivariate based anomaly detector (RX) (*2*). (See appendix for their mathematical representations.)

Using the denotations introduced in section 1, the first experiment uses H_{01} to represent the entire null hypothesis H_0. Detectors' outputs using samples under H_0 will be used to obtain cutoff thresholds, and subsequently estimate type I errors of both detectors. (The type I error is the probability of rejecting H_0 given that H_0 is true, see for instance, (*8*)). An alternative hypothesis similar to H_{11} is then used to represent the entire H_1 for the detectors' *power* estimates. (The power of a detector is defined as 1.0 minus type II error, where the type II error is the probability of not rejecting H_0 given that H_1 is true (*8*).) A key factor in the first experiment is that classes C_{K+1} and C_1 are similar classes (see section 1), so, we will add uncorrelated normally distributed noise to samples generated to represent C_{K+1} and will vary the amount of noise to measure the power behavior of both detectors. The goal here is to observe the power of both detectors as H_1 varies from an easier alternative (C_{K+1} with high noise) to a harder alternative (C_{K+1} without noise) relative to a given H_0. Note that *noise* is only applied to the generation of C_{K+1} samples under H_1 in the first experiment.

In the second simulation experiment, using also parameter estimates from real spectra to generate ideal samples, we will use an alternative hypothesis similar to H_{12}, where in this case classes C_{K+2} and C_1 are distinct from each other. We then will check the power behavior of both

detectors by varying, instead, a noise term under H_0, which will be added to one of the random samples in a sample pair similar to H_{01}. The goal here is to observe the power of both detectors as a function of a noise-varying null hypothesis.

2.1 Data Structures and Representations of the Null/Alternative Hypotheses

In this subsection, we describe data models for the generation of random samples $\mathbf{X}_0$ and $\mathbf{X}_1$ representing the null and alternative hypotheses. Models for two simulation experiments will be described: Experiment 1 and Experiment 2. (We will use the denotation introduced in section 1.)

Experiment 1

The goal of this simulation experiment is, in essence, to answer a question from a generalized concern: how much noise would be required under H_1 to sensitize each testing detector, given that—by design—neither detector is able to confidently separate samples under H_0 from those under H_1 without added noise? To answer this question, we will use homogeneous classes and determine which detector can yield higher power as a function of *decreasing asymmetric noise* under H_1. (In this case, asymmetric noise under H_1 relates to noise being added to only one of the samples in the detector's input sample pair, as it will be shown later.)

Our conjecture is that the detection approach that can yield significantly higher power, in the context of Experiment 1, is a better detector, since this detector would potentially demonstrate higher sensitivity to harder alternatives. In the context of Experiment 1 this is a generalized problem. Experiment 1 excludes the presence of sample mixtures in either hypothesis or in both null and alternative hypotheses.

Two sets of real spectra representing two homogenous classes C_1 and C_{K+1} were selected and used to estimate mean and covariance sets, where the 120-dimensional (dim) mean estimate $\hat{\boldsymbol{\mu}}_0$ and the 120 by 120 covariance estimate $\hat{\boldsymbol{\Sigma}}_0$ relates to C_1, and the 120-dim mean estimate $\hat{\boldsymbol{\mu}}_1$ and the 120 by 120 covariance estimate $\hat{\boldsymbol{\Sigma}}_1$ relates to C_{K+1}. Although our selections were arbitrarily made, we used a selection criterion, such that, C_1 and C_{K+1} are not very distinguishable from each other, i.e., both detectors should have problems separating these classes based only on testing samples from both classes. Each spectral set consisted of 2500 spectral samples representing 120 frequency band material responses in the visible to near infrared (VNIR) region of the spectrum. Parameter estimates $\left(\hat{\boldsymbol{\mu}}_0, \hat{\boldsymbol{\Sigma}}_0\right)$ and $\left(\hat{\boldsymbol{\mu}}_1, \hat{\boldsymbol{\Sigma}}_1\right)$ are specified in subsection 2.4.

Null Hypothesis—H_0: We use an idealized case for H_0 where n_0 multivariate samples $\mathbf{X}_{01}, \mathbf{X}_{02}, \cdots, \mathbf{X}_{0n_0}$ are drawn independently from a normal PDF $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and another n_1 samples $\mathbf{X}_{11}, \mathbf{X}_{12}, \cdots, \mathbf{X}_{1n_1}$ are drawn independently from the same PDF, such that, in both cases $(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \left(\hat{\boldsymbol{\mu}}_0, \hat{\boldsymbol{\Sigma}}_0\right)$. Both $\mathbf{X}_{0u}$ $(u = 1, \cdots, n_0)$ and $\mathbf{X}_{1h}$ $(h = 1, \cdots, n_1)$ are 120-dim column vectors. The

null hypothesis H_0 corresponds to an independent, identically distributed (IID) sample pair distributed (~) as follows

$$\begin{aligned}\mathbf{X}_0^{(w)} &= \left(\mathbf{X}_{01}^{(w)}, \mathbf{X}_{02}^{(w)}, \cdots, \mathbf{X}_{0n_0}^{(w)}\right) \sim \text{IID } N\left(\hat{\boldsymbol{\mu}}_0, \hat{\boldsymbol{\Sigma}}_0\right) \\ \mathbf{X}_1^{(w)} &= \left(\mathbf{X}_{11}^{(w)}, \mathbf{X}_{12}^{(w)}, \cdots, \mathbf{X}_{1n_1}^{(w)}\right) \sim \text{IID } N\left(\hat{\boldsymbol{\mu}}_0, \hat{\boldsymbol{\Sigma}}_0\right)\end{aligned} \tag{1}$$

where $\mathbf{X}_0^{(w)}$ and $\mathbf{X}_1^{(w)}$ are independent, and $w = 1, \cdots, m$ indexes independent trials. To relate (6.1), for instance, to the target application, consider, instead, $w = 1, \cdots, m$ indexing independent spatial locations in the HS imagery, where model (1) would be reasonably satisfied, such that, $\mathbf{X}_1^{(w)}$ would represent a reformatted block of data W_1 (centered at w) and $\mathbf{X}_0^{(w)}$ would represent reformatted data from W_1's outer ring W_0, see figure A-1 in appendix A.

Alternative Hypothesis—H_1: In this hypothesis, since by design C_1 and C_{K+1} are not separable classes using either detector, uncorrelated normally distributed random noise $\left(\eta_1, \eta_2, \cdots, \eta_{n_1}\right)$ is added to the samples belonging to C_{K+1} under H_1 in order to yield some distinction between these two classes. The alternative H_1 then corresponds to the following sample pair

$$\begin{aligned}\mathbf{X}_0^{(w)} &= \left(\mathbf{X}_{01}^{(w)}, \mathbf{X}_{02}^{(w)}, \cdots, \mathbf{X}_{0n_0}^{(w)}\right) \sim \text{IID } N\left(\hat{\boldsymbol{\mu}}_0, \hat{\boldsymbol{\Sigma}}_0\right) \\ \mathbf{X}_1^{(w)} &= \left(\mathbf{X}_{21}^{(w)}, \mathbf{X}_{22}^{(w)}, \cdots, \mathbf{X}_{2n_1}^{(w)}\right) + \left(\eta_1^{(w)}, \eta_2^{(w)}, \cdots, \eta_{n_1}^{(w)}\right),\end{aligned} \tag{2}$$

where

$$\left(\mathbf{X}_{21}^{(w)}, \mathbf{X}_{22}^{(w)}, \cdots, \mathbf{X}_{2n_1}^{(w)}\right) \sim \text{IID } N\left(\hat{\boldsymbol{\mu}}_1, \hat{\boldsymbol{\Sigma}}_1\right),$$

$$\left(\eta_1^{(w)}, \eta_2^{(w)}, \cdots, \eta_{n_1}^{(w)}\right) \sim \text{IID } N\left(0, \hat{\boldsymbol{\Sigma}}_\eta\right),$$

$\mathbf{X}_0^{(w)}$ and $\mathbf{X}_1^{(w)}$ are independent, $\mathbf{X}_{0u}^{(w)}$ $(u = 1, \cdots, n_0)$, $\mathbf{X}_{2h}^{(w)}$ $(h = 1, \cdots, n_1)$, and $\eta_h^{(w)}$ $(h = 1, \cdots, n_1)$ are 120-dim column vectors, $w = 1, \cdots, m$ indexes independent trials, $N\left(0, \hat{\boldsymbol{\Sigma}}_\eta\right)$ is a normal PDF centered at zero in an 120-dim space where $\hat{\boldsymbol{\Sigma}}_\eta$ is a 120 by 120 diagonal matrix having non-zero components ζ_b^2 $(b = 1, \cdots, 120)$ related to the diagonal components σ_b^2 $(b = 1, \cdots, 120)$ of $\hat{\boldsymbol{\Sigma}}_1$ by a noise level factor $p > 0$ (i.e., $\zeta_b^2 = p\sigma_b^2$). In this noise model, we consider p = 1.0 as adding 100 percent (%) of noise to a multivariate sample.

Using models (1) and (2) to generate pseudo samples, we would like to determine which detector can yield higher power at a lesser noise level under H_1. A detection approach that can yield higher power in this context is a better detector.

Experiment 2

The goal of this simulation experiment is, in essence, to answer a question from a more specialized concern: how much noise would be required under H_0 in order to decrease power of

each testing detector, given that—by design—both detectors are able to comfortably separate samples under H_0 from those under H_1 without added noise? To answer this question, we will determine which detector can yield higher power as a function of *increasing asymmetric noise* under H_0, given that both detectors can confidently separate samples from H_0 and H_1 at *zero* noise. (In this case, asymmetric noise under H_0 relates to noise being added to only one of the samples in the detector's input sample pair, as shown in (3)).

Our conjecture is that the detection approach that can yield significantly higher power in the context of Experiment 2 is a better detector for that specialized concern, since this detector would potentially demonstrate lower sensitivity to noise under H_0. (Although the problem that Experiment 2 addresses is a specialized one, it can happen in the target application as discussed later.) Experiment 2 excludes the presence of sample mixtures in either hypothesis or in both null and alternative hypotheses.

In Experiment 2, we also selected and used two sets of real spectra representing two arbitrary classes C_1 and C_{K+2}, such that, both classes are significantly different from each other. Parameter estimates $\left(\hat{\boldsymbol{\mu}}_0, \hat{\boldsymbol{\Sigma}}_0\right)$ and $\left(\hat{\boldsymbol{\mu}}_2, \hat{\boldsymbol{\Sigma}}_2\right)$ were computed using 2500 samples from each class C_1 and C_{K+2}, where $\left(\hat{\boldsymbol{\mu}}_0, \hat{\boldsymbol{\Sigma}}_0\right)$ are described in Experiment 1 and $\left(\hat{\boldsymbol{\mu}}_2, \hat{\boldsymbol{\Sigma}}_2\right)$ are the corresponding estimates from C_{K+2}. Each spectral sample in both sets represents 120 frequency band material responses in the VNIR region. Parameter estimates $\left(\hat{\boldsymbol{\mu}}_2, \hat{\boldsymbol{\Sigma}}_2\right)$ are specified in subsection 2.4.

Null Hypothesis—H_0: n_0 multivariate samples $\mathbf{X}_{01}, \mathbf{X}_{02}, \cdots, \mathbf{X}_{0n_0}$ are drawn independently from a normal PDF $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and another n_1 samples $\mathbf{X}_{11}, \mathbf{X}_{12}, \cdots, \mathbf{X}_{1n_1}$ are drawn independently from the same PDF, such that, in both cases $(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \left(\hat{\boldsymbol{\mu}}_0, \hat{\boldsymbol{\Sigma}}_0\right)$. Both $\mathbf{X}_{0u}$ $(u = 1, \cdots, n_0)$ and $\mathbf{X}_{1h}$ $(h = 1, \cdots, n_1)$ are 120-dim column vectors. For this experiment, however, uncorrelated normally distributed random noise $\left(\eta_1, \eta_2, \cdots, \eta_{n_0}\right)$ is added to one of the sample sequences in order to observe the power behavior of both detectors as a function of noise under H_0. (Notice that in order to affect power of both detectors, noise under H_0 was applied asymmetrically to one of the sample sequences, not to both.) The null hypothesis then corresponds to

$$\begin{aligned} \mathbf{X}_0^{(w)} &= \left(\mathbf{X}_{01}^{(w)}, \mathbf{X}_{02}^{(w)}, \cdots, \mathbf{X}_{0n_0}^{(w)}\right) + \left(\eta_1^{(w)}, \eta_2^{(w)}, \cdots, \eta_{n_0}^{(w)}\right) \\ \mathbf{X}_1^{(w)} &= \left(\mathbf{X}_{11}^{(w)}, \mathbf{X}_{12}^{(w)}, \cdots, \mathbf{X}_{1n_0}^{(w)}\right) \sim \text{IID } N\left(\hat{\boldsymbol{\mu}}_0, \hat{\boldsymbol{\Sigma}}_0\right), \end{aligned} \tag{3}$$

where

$$\left(\mathbf{X}_{01}^{(w)}, \mathbf{X}_{02}^{(w)}, \cdots, \mathbf{X}_{0n_0}^{(w)}\right) \sim \text{IID } N\left(\hat{\boldsymbol{\mu}}_0, \hat{\boldsymbol{\Sigma}}_0\right),$$

$$\left(\eta_1^{(w)}, \eta_2^{(w)}, \cdots, \eta_{n_0}^{(w)}\right) \sim \text{IID } N\left(0, \hat{\boldsymbol{\Sigma}}_\eta\right),$$

$\mathbf{X}_0^{(w)}$ and $\mathbf{X}_1^{(w)}$ are independent, $\eta_h^{(w)}\ (h=1,\cdots,n_0)$ are 120-dim column vectors, $w=1,\cdots,m$ indexes independent trials, $N\left(0,\hat{\boldsymbol{\Sigma}}_\eta\right)$ is a normal PDF centered at zero in an 120-dim space where $\hat{\boldsymbol{\Sigma}}_\eta$ is a 120 by 120 diagonal matrix having non-zero components $\zeta_b^2\ (b=1,\cdots,120)$ related to the diagonal components $\lambda_b^2\ (b=1,\cdots,120)$ of $\hat{\boldsymbol{\Sigma}}_0$ by a noise level factor $p>0$ (i.e., $\zeta_b^2 = p\lambda_b^2$).

Alternative Hypothesis—H1: n_0 multivariate samples $\mathbf{X}_{01},\mathbf{X}_{02},\cdots,\mathbf{X}_{0n_0}$ are drawn independently from a normal PDF $N(\boldsymbol{\mu},\boldsymbol{\Sigma})$, and another n_1 samples $\mathbf{X}_{21},\mathbf{X}_{22},\cdots,\mathbf{X}_{2n_1}$ are drawn independently from another normal PDF $N(\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_2)$, such that, $(\boldsymbol{\mu},\boldsymbol{\Sigma})=(\hat{\boldsymbol{\mu}}_0,\hat{\boldsymbol{\Sigma}}_0)$ and $(\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_2)=(\hat{\boldsymbol{\mu}}_2,\hat{\boldsymbol{\Sigma}}_2)$. Both $\mathbf{X}_{0u}\ (u=1,\cdots,n_0)$ and $\mathbf{X}_{1h}\ (h=1,\cdots,n_1)$ are 120-dim column vectors. This alternative hypothesis corresponds to

$$\begin{aligned}\mathbf{X}_0^{(w)} &= \left(\mathbf{X}_{01}^{(w)},\mathbf{X}_{02}^{(w)},\cdots,\mathbf{X}_{0n_0}^{(w)}\right) \sim \text{IID } N\left(\hat{\boldsymbol{\mu}}_0,\hat{\boldsymbol{\Sigma}}_0\right)\\ \mathbf{X}_1^{(w)} &= \left(\mathbf{X}_{21}^{(w)},\mathbf{X}_{22}^{(w)},\cdots,\mathbf{X}_{2n_1}^{(w)}\right) \sim \text{IID } N\left(\hat{\boldsymbol{\mu}}_2,\hat{\boldsymbol{\Sigma}}_2\right)\end{aligned} \tag{4}$$

where $\mathbf{X}_0^{(w)}$ and $\mathbf{X}_1^{(w)}$ are independent, and $w=1,\cdots,m$ indexes independent trials.

Using models (3) and (4) to generate pseudo samples, we would like to determine which detector can yield higher power as a function of increasing asymmetric noise under H_0. A detection approach that can yield higher power would be considered a better detector for this particular case.

2.2 Simulation Plan

We used the following steps to conduct the simulation experiments using the data model described in section 2.1:

i. Each multivariate random sample in the three sequences $\left(\mathbf{X}_{01}^{(w)},\mathbf{X}_{02}^{(w)},\cdots,\mathbf{X}_{0n_0}^{(w)}\right)$, $\left(\mathbf{X}_{11}^{(w)},\mathbf{X}_{12}^{(w)},\cdots,\mathbf{X}_{1n_1}^{(w)}\right)$, and $\left(\mathbf{X}_{21}^{(w)},\mathbf{X}_{22}^{(w)},\cdots,\mathbf{X}_{2n_1}^{(w)}\right)$, as shown in (1), (2), (3), and (4), consists of covariates $x_b\ (b=1,\cdots,120)$ representing material radiance sensed by a HS sensor operating on *120* different frequencies. Covariates x_b are highly correlated pseudo values and will be generated every trial from multivariate normal distributions, as shown in (1), (2), (3), and (4). Parameter values for these distributions are estimated using real samples of material classes that are present in HS data. Trials are indexed by *w*, where $w=1,\cdots,m$, and the sample sizes n_0 and n_1 are fixed to *500*.

ii. We fix at once the distribution parameters of different classes to values estimated in (i), and fix also the parameters of the noise distributions in Experiment 1 and in Experiment 2 to values shown in (2) and (3), respectively, where noise samples are uncorrelated and *zero-mean* normally distributed.

iii. Attaining Cutoff Thresholds/Estimating Type I Errors: For Experiment 1, values of $\{\mathbf{X}_0^{(w)}\}_{w=1}^{m}$ and $\{\mathbf{X}_1^{(w)}\}_{w=1}^{m}$ (*m = 1500*) will be generated according to model (1) and introduced to both detectors. These detectors will yield *1500* output results each. Each set of *1500* results will be used to estimate an empirical PDF, and a cutoff threshold will be attained using the standard quantile method on the estimated PDF. The desired type I error probability (ε) for the quantile method is fixed at once for both detectors to $\varepsilon = 0.05$. The detectors' corresponding cutoff thresholds will be applied to the corresponding detectors' output results in order to estimate the type I error $(\hat{\varepsilon})$ of each detector. The type I error is estimated by counting the number of trials m_1 that satisfy the detector's output values being greater than the detector's cutoff threshold and computing the ratio $\hat{\varepsilon} = m_1 / m$.

For Experiment 2, since we are interested in measuring the effect of asymmetric noise under H_0 on the power of both detectors, the noise level factor p in (3) will be systematically increased from a virtual *zero* value to a relatively high value $(p = 5.0)$ in order to obtain cutoff thresholds as a function of varying p under H_0. Values of $\{\mathbf{X}_0^{(w)}\}_{w=1}^{m}$ and $\{\mathbf{X}_1^{(w)}\}_{w=1}^{m}$ (*m = 1500*) will be generated according to model (3) and introduced to both detectors, and we will follow the same procedure used in Experiment 1 to attain detectors' cutoff thresholds and to estimate type I errors of both detectors at a fixed type I error $\varepsilon = 0.05$. This approach will be followed for different values of p.

iv. Estimating the Power (*1.0* – Type II Error): In Experiment 1, since we are interested in observing the power behavior of both detectors as a function of decreasing noise under H_1, the noise level factor p in (2) will be systematically decreased from a relatively high value $(p = 5.0)$ to a virtual *zero* value. Values of $\{\mathbf{X}_0^{(i)}\}_{i=1}^{m}$ and $\{\mathbf{X}_1^{(i)}\}_{i=1}^{m}$ (*m = 1500*) will be generated according to model (2) and introduced to both detectors, where the detectors' corresponding cutoff thresholds will be applied to the corresponding detectors' output results in order to estimate the *power* $(\hat{\theta})$ of each detector as a function of p. The type II error at each p is estimated by counting the number m_2 of trials that satisfy the detector's output values being lower then the detector's cutoff threshold and computing the ratio m_2 / m, such that $\hat{\theta} = 1.0 - (m_2 / m)$.

In Experiment 2, values of $\{\mathbf{X}_0^{(i)}\}_{i=1}^{m}$ and $\{\mathbf{X}_1^{(i)}\}_{i=1}^{m}$ (*m = 1500*) will be generated according to model (4) and introduced to both detectors, where the detectors' corresponding cutoff thresholds will be applied to the corresponding detectors' output results in order to estimate $\hat{\theta}$ of each detector as a function of p, as shown in (3), see also (iii) under Experiment 2.

v. Estimating Performance Confidence Intervals (CI): Results from (iii) and (iv) do not provide insights into the performance variability of the two detectors. In order to check this

variability, confidence intervals for (iii) $\bar{\varepsilon} \pm z_{\alpha/2}\sqrt{\frac{1}{R-1}\sum_{r=1}^{R}\left(\hat{\varepsilon}_r - \bar{\varepsilon}\right)^2}$ and for (iv) $\bar{\theta} \pm z_{\alpha/2}\sqrt{\frac{1}{R-1}\sum_{r=1}^{R}\left(\hat{\theta}_r - \bar{\theta}\right)^2}$ will be estimated, where $z_{\alpha/2} = \Phi^{-1}\left(1-\alpha/2\right)$ is the $1-\alpha/2$ quantile $\left(\alpha = 0.05\right)$ of the standard normal distribution, random quantities simulated are indexed by *r* for a total number of simulation repetitions *R=1000*, $\bar{\varepsilon} = R^{-1}\sum_{r=1}^{R}\hat{\varepsilon}_r$, and $\bar{\theta} = R^{-1}\sum_{r=1}^{R}\hat{\theta}_r$. This approach will be used in both simulation experiments.

vi. The multivariate RX anomaly detector and the univariate AsemiP anomaly detector will be implemented as described in appendix.

2.3 Simulation Results

We present in this subsection results using the simulation plan that is outlined in section 2.2. Table 1 tabulates results using the null and alternative hypotheses, whose models are shown in equations 1 and 2 (Experiment 1), and table 2 tabulates results using models equations 3 and 4 (Experiment 2).

Table 1. Probability of correct detection (PD or Power) of each detector (*Det*) as a function of noise, *Noise (%)*, under H_1. The mark * in the *Noise (%)* column indicates *zero* noise added.

Experiment 1		$n = 500; m = 1500; R = 1000; r = 1,\cdots,R; \varepsilon = 0.05$			
		Type I $\left(\hat{\varepsilon}_r; CI\ 95\%\right)$		*PD or Power* $\left(\hat{\theta}_r; CI\ 95\%\right)$	
Det	*Noise(%)*	*Lower Bound*	*Upper Bound*	*Lower Bound*	*Upper Bound*
RX	*	0.048110241	0.051012207	0.059503804	0.081965582
	5	0.049104126	0.051960194	0.062518288	0.099420487
	10	0.049320603	0.052393682	0.137486868	0.182186600
	50	0.048363255	0.051616335	0.252472513	0.304874425
	100	0.049095127	0.051721199	0.482344791	0.545002146
	200	0.049170374	0.051523503	0.682996642	0.727003357
	300	0.048362805	0.051698418	0.837302959	0.855758264
	400	0.048295303	0.052051635	0.910798592	0.935732019
	500	0.048748037	0.051537677	1.000000000	1.000000000
AsemiP	*	0.047090192	0.051781236	0.063058074	0.088416970
	5	0.048007154	0.051890804	0.212100870	0.261572598
	10	0.048157886	0.051964562	0.661850164	0.714476365
	50	0.049380616	0.051272444	0.845217278	0.885394966
	100	0.048667252	0.051251115	0.981156962	0.992720588
	200	0.049136926	0.052116135	1.000000000	1.000000000
	300	0.048104126	0.051610159	1.000000000	1.000000000
	400	0.048756542	0.051529171	1.000000000	1.000000000
	500	0.048786787	0.052193350	1.000000000	1.000000000

Table 2. PD (Power) of each detector (*Det*) as a function of noise, *Noise (%)*, under H_0. The mark * in the *Noise (%)* column indicates *zero* noise added.

Experiment 2		$n = 500; m = 1500; R = 1000; r = 1,\cdots, R; \varepsilon = 0.05$			
		Type I $(\hat{\varepsilon}_r; CI\ 95\%)$		*PD or Power* $(\hat{\theta}_r; CI\ 95\%)$	
Det	*Noise(%)*	*Lower Bound*	*Upper Bound*	*Lower Bound*	*Upper Bound*
RX	*	0.048181716	0.051018283	1.000000000	1.000000000
	50	0.047606835	0.051993164	1.000000000	1.000000000
	70	0.046734635	0.051465364	1.000000000	1.000000000
	100	0.048601189	0.053198810	1.000000000	1.000000000
	200	0.048999652	0.052400347	1.000000000	1.000000000
	300	0.048928291	0.052871708	0.700000000	0.916848328
	400	0.048105500	0.051694499	0.600000000	0.739118276
	500	0.048435825	0.051164174	0.500000000	0.527949857
AsemiP	*	0.048700333	0.051699666	1.000000000	1.000000000
	50	0.048357151	0.052842848	1.000000000	1.000000000
	70	0.047625488	0.051774511	0.997260240	0.999139759
	100	0.049023210	0.052576789	0.952304476	0.963695523
	200	0.047972484	0.051427515	0.577642209	0.615357790
	300	0.048788093	0.051811906	0.378761022	0.415038977
	400	0.047862071	0.051337928	0.289353636	0.321646363
	500	0.048649448	0.051750551	0.233169749	0.260030250

Table 1 shows the 95% confidence intervals (*CI* 95%) of the estimated type I error $\hat{\varepsilon}_r\ (r = 1\cdots R)$ and estimated power $\hat{\theta}_r\ (r = 1\cdots R)$ corresponding to both detectors as a function of noise (by varying parameter value p) under H_1. The sample size was fixed to $n = 500$ and the total number of trials to $m = 1500$ as described in section 2.2 (*i*), the total number of simulation repetitions was fixed to *R=1000,* and the desired type I error as described in section 2.2 (*iii*) was fixed to $\varepsilon = 0.05$. The column under *Noise (%)* shows the noise added to a multivariate sample under H_1 in terms of percentage relative to the diagonal terms of a real-datum estimated covariance matrix, as described in subsection 2.1. The relationship between parameter p and *Noise (%)* has the following expression and constraint: *Noise (%)* = $p \cdot 100(\%)$, for $p > 0$ (p cannot be equal to zero because this equality would force the diagonal terms of the noise's covariance to be all zeros; instead, the mark * in the *Noise (%)* column indicates *zero* noise added).

Recall that the goal of Experiment 1 is to determine which detector can yield higher power as a function of decreasing asymmetric noise under H_1, given that—by design—neither detector is able to confidently separate samples under H_0 from those samples under H_1 at zero noise.

The *Power CI* column in table 1 shows that at zero noise, *Noise (%)* marked as *, both detectors performed poorly in power, which validates that H_1 is indeed a hard alternative at zero noise, as expected. Table 1 also shows that for a given type I error ($\varepsilon = 0.05$) and sample size ($n = n_0 = n_1 = 500$), the power of both detectors increased, as expected, as a function of increasing noise under H_1, and that the AsemiP detector yielded a significantly higher power than did the RX detector, given that both detectors shared the same pseudo sample pairs generated in Experiment 1. Notice, for instance, in table 1 that it required a noise level of 50% under H_1 for the RX detector to produce a power that is comparable to that of the AsemiP detector corresponding to 5% of noise, (see *Power CI* for *Noise (%)* equals to 50 (RX) and for *Noise (%)* equals to 5 (AsemiP)). In corresponding columns in table 1, notice also that it required a noise level between 400% and 500% under H_1 for the RX detector to produce a unity power, but only a noise level above 100% for the AsemiP detector to achieve the same performance.

Results shown in table 1, in essence, depict the detection approach that is more sensitive to material classes that are less separable in the HS space, a behavior that is desired. Since the AsemiP detector consists of a data preprocessing step that aims at separating less separable classes, we attribute the performance shown in table 1 of the AsemiP approach less to its test statistic step but more to its data preprocessing step.

The AsemiP test statistic step will be challenged in the simulation experiments described in section 3, where idealized heterogeneous samples (mixtures) will be used to simulate spatial transitions of distinct regions in HS data. In this case, the data preprocessing step of the AsemiP detector will not be of much help.

Table 2 shows tabulated results of both detectors corresponding to Experiment 2.

Recall that the goal of Experiment 2 is to determine which detector can yield higher power as a function of increasing asymmetric noise under H_0, given that—by design—both detectors are able to confidently separate samples under H_0 from those samples under H_1 at zero noise.

The PD or *Power CI* columns in table 2 for *Noise (%)* marked as * shows that both detectors can yield a unity power at *zero* noise under H_0, which validates that H_1 is indeed a relatively *easy* alternative, as expected. Table 2 also shows that for a given type I error and sample size, the power of both detectors decreased, as expected, as a function of increasing noise under H_0, and that the RX detector was able to maintain a significantly higher power than did the AsemiP detector as a function of increasing noise. Notice, for instance, in table 2 that it required a noise level between 200% and 300% under H_0 for the RX detector to decrease its power from unity. In contrast, it required a mere noise level near 70% to decrease the AsemiP detector's power from unity, (see *Power CI* for *Noise (%)* equals to 200 and 300 (RX) and for *Noise (%)* equals to 70 (AsemiP)). At noise levels between 200% and 300%, the AsemiP detector dropped its power to

about half of its peak; the same occurred for the RX detector at a much higher noise level—near 500%, which is more favorable.

We suspect that the same data preprocessing step that contributed to the favorable AsemiP detector's power behavior shown in table 1 is also responsible for the unfavorable AsemiP detector's power behavior shown in table 2. The data preprocessing step in our approach was proposed to increase sensitivity between samples belonging to different classes that happen to have similar statistical features in the original data space. In Experiment 2, this sensitivity backfired on the AsemiP power behavior, because, as a function of increasing asymmetric noise under H_0, this detector's cutoff threshold increased more significantly than the RX detector's cutoff threshold did relative to their responses under H_1.

Results shown in table 2, in essence, depict the detection approach that is more sensitive to increasing asymmetric noise under H_0, a behavior that is not desired. Although Experiment 2 symbolizes a specialized problem, this problem can actually happen in the target application, which is why we included it for testing.

To relate a practical problem to the specialized concern described in Experiment 2, we must change our sampling approach for the detector's training and testing stages. We have used thus far the sampling approach known in the target community as *inside/outside sampling* for both training and testing, where a block of data in the HS imagery and this block's surrounding data are used to form $\mathbf{X}_1$ and $\mathbf{X}_0$, respectively. For the concern expressed in Experiment 2, we assume, instead, that pre-stored HS samples are available from a database to train a detector—training is the process of obtaining a cutoff threshold, given a type I error and sample size. But to test incoming imagery, we assume that the inside/outside sampling approach is used to feed the detector with samples for power estimation.

For training, suppose that random samples $\mathbf{X}_0^{(w)}$ and $\mathbf{X}_1^{(w)}$, as shown in (3), represent spectra from a real database, such that, $\mathbf{X}_0$ depict random samples of a homogeneous background scene collected under a certain atmospheric condition ψ_a belonging to a set of A possible conditions, or $\psi_1,\cdots,\psi_A$, and $\mathbf{X}_1^{(w)}$ depict also random samples of the same homogeneous background scene under $\psi_1,\cdots,\psi_A$. Assume now that $\mathbf{X}_0^{(w)}$ and $\mathbf{X}_1^{(w)}$ are used to obtain the detector's cutoff threshold at a given type I error, such that, $\mathbf{X}_0^{(w)}$ were collected under condition ψ_a and $\mathbf{X}_1^{(w)}$ under condition ψ_c $(c=1,\cdots,A)$. In this context, conditions where $a \neq c$ can be modeled (simplistically) by adding asymmetric noise under H_0, as shown in (3). Higher effects between ψ_a and ψ_c on the same background scenes can be mimicked by increasing the noise level under H_0. To complete this scenario, we would have to assume also that $\mathbf{X}_0^{(w)}$ and $\mathbf{X}_1^{(w)}$, as shown in (4), represent the case under H_1 where $\mathbf{X}_0^{(w)}$ are random samples from the same background scene used to obtain the detector's cutoff threshold and $\mathbf{X}_1^{(w)}$ are random samples of a genuine

anomalous material in the scene. This practical scenario fits the concerned addressed in Experiment 2.

A favorable comment on the AsemiP detector's unfavorable performance shown in table 2 is that it required a relatively high noise level for this detector to show a significant power degradation, i.e., over 70%; see table 2. It was insightful to learn that the RX detector performed more favorably as the noise level increased beyond 50% in Experiment 2.

2.4 Data Model Parameter Values

For Experiment 1 and Experiment 2, we used 2,500 real spectral samples per material class to estimate these classes' means and covariance matrixes. Three material classes were chosen arbitrarily so that two of them were reasonably similar to each other, using both detectors' low power responses as criteria, and the remaining one was very distinct from the other two.

As discussed earlier, these classes' estimated means are denoted by $\hat{\boldsymbol{\mu}}_0$ (class 1) and $\hat{\boldsymbol{\mu}}_1$ (class 2)—both attained from two similar classes—and $\hat{\boldsymbol{\mu}}_2$ (class 3)—the distinct class. These classes' estimated covariance matrices are denoted by $\hat{\boldsymbol{\Sigma}}_0$ (class 1), $\hat{\boldsymbol{\Sigma}}_1$ (class 2), and $\hat{\boldsymbol{\Sigma}}_2$ (class 3).

The estimated covariance matrices of the three material classes are shown in figure 1, so that, one can visually appreciate the material radiance correlation among the 120 frequency bands by each material class. Estimated means and estimated variances corresponding to the three classes are partially shown in table 3 (first 10 components only), for a complete table see appendix).

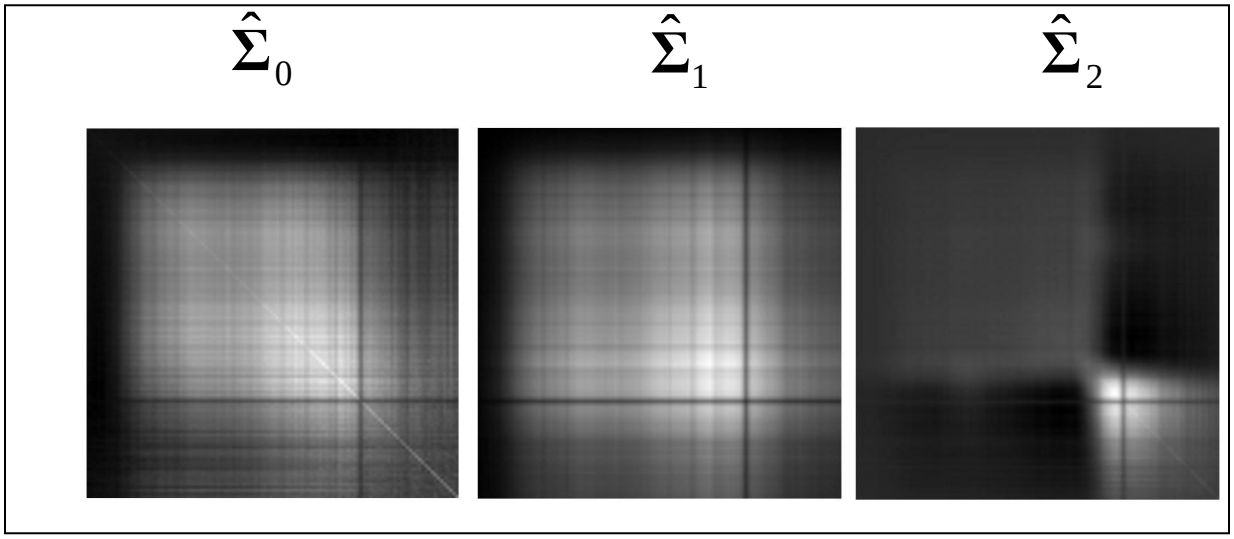

Figure 1. Estimated covariance matrices of class 1 $\left(\hat{\boldsymbol{\Sigma}}_0\right)$, class 2 $\left(\hat{\boldsymbol{\Sigma}}_1\right)$, and class 3 $\left(\hat{\boldsymbol{\Sigma}}_2\right)$. The upper left corner of each matrix represents the response variance at frequency band 1—there are 120 bands. (These images were scaled independently using their own maximum and minimum values for displaying purposes only, such that, whitest represents the high value and black represents the lowest value.)

Table 3. (Partial Table) Estimated parameter values using real HS data from three arbitrarily chosen material classes, such that, two of these classes are relatively similar and the remaining one is distinct. Estimated means $\hat{\boldsymbol{\mu}}_0$ (class 1) and $\hat{\boldsymbol{\mu}}_1$ (class 2) were attained from two similar classes, and estimated mean $\hat{\boldsymbol{\mu}}_2$ (class 3) was obtained from a distinct class. Estimates $diag\left(\hat{\boldsymbol{\Sigma}}_0\right)$, $diag\left(\hat{\boldsymbol{\Sigma}}_1\right)$, and $diag\left(\hat{\boldsymbol{\Sigma}}_2\right)$ are the diagonal terms of $\hat{\boldsymbol{\Sigma}}_0$, $\hat{\boldsymbol{\Sigma}}_1$, and $\hat{\boldsymbol{\Sigma}}_2$, respectively. These vectors are 1 by 120 representing these statistical results from 120 frequency bands; see appendix for the complete table.

$\hat{\boldsymbol{\mu}}_0$ (1.0e3)	$\hat{\boldsymbol{\mu}}_1$ (1.0e3)	$\hat{\boldsymbol{\mu}}_2$ (1.0e3)	$diag\left(\hat{\boldsymbol{\Sigma}}_0\right)$ (1.0e3)	$diag\left(\hat{\boldsymbol{\Sigma}}_1\right)$ (1.0e5)	$diag\left(\hat{\boldsymbol{\Sigma}}_2\right)$ (1.0e4)
0.204325	0.175925	0.087980	0.160109	0.004161	0.011063
0.220555	0.190895	0.094050	0.121826	0.005398	0.010110
0.244355	0.216750	0.102410	0.138863	0.007945	0.012112
0.260325	0.236395	0.106035	0.158180	0.011452	0.014576
0.282800	0.260630	0.111530	0.172251	0.016150	0.015648
0.303000	0.285110	0.114955	0.192190	0.022174	0.018835
0.306015	0.295890	0.116005	0.155532	0.028275	0.018890
0.343840	0.329070	0.128410	0.195692	0.039375	0.025303
0.400310	0.385030	0.145650	0.229260	0.063204	0.033793
0.426200	0.414335	0.152510	0.253849	0.080857	0.039194

2.5 Concluding Remarks

Section 2 focused on simulation experiments that showed relative strengths and weaknesses of two competing anomaly detection approaches, AsemiP and RX, testing idealized homogeneous sample models that included varying noise. The section's emphasis was on adding a so-called asymmetric noise to samples of a homogeneous class under the alternative hypothesis (Experiment 1) and under the null hypothesis (Experiment 2) in order to observe the power behavior (probability of correct detection) of both detectors. Other cases involving mixtures are treated in section 3. Data models were specified and shared by these detectors representing H_0 and H_1 in both experiments, where (differently from simulation experiments described in sections 3 and 4) real data were used to estimate parameters of idealized class distributions.

Experiment 1 aimed at addressing a more generalized concern, that is, to determine which one of these detectors could sustain a higher power as a function of decreasing noise under H_1, given that—by design—H_1 (with zero noise) was a hard alternative consisting of generated idealized samples representing two similar homogeneous classes. (Increasing noise in this context made these classes more separable.) The AsemiP detector showed relative strength over the RX detector in Experiment 1, as shown in table 1. We attribute the AsemiP's favorable performance shown in table 1 to this detector's data preprocessing step, which is an integral part of our anomaly detection approach. In the target community, while data preprocessing techniques are designed to promote separation among similar classes in the original data space, these classes are assumed to be homogeneous. However, if some of the observed samples belong to a mixture,

which is always the case in the target application, then one would have to rely on the test statistic step to handle the effect of mixtures on the detector's power.

Experiment 2 aimed at addressing a more specialized concern, that is, to determine which one of these detectors can sustain a higher power as a function of increasing noise under H_0, given that—by design—H_1 (with zero noise under H_0) was an easy alternative consisting of generated idealized samples representing two distinct homogeneous classes. The RX detector showed relative strength over the AsemiP detector in Experiment 2, as shown in table 2. This simulation experiment addresses a more specialized concern, but it could happen in the target application, as a special case, when samples of the same homogeneous class are collected under very distinct atmospheric conditions and these sample sets are then used later to represent the null hypothesis. The effect on the detector's power owing to atmospheric condition disparities under the null hypothesis can be modeled in a simplistic form by adding random noise to one of the random samples in the sample pair used to train the detector. Higher noise implies higher atmospheric condition disparity during data collection of the same homogenous class. And the concern here is how this increase would affect the power of these detectors. Of course, this simplistic model only used two homogenous classes between H_0 and H_1 and uncorrelated normally distributed noise, as shown in (3) and (4), but it was insightful to observe the responses of both detectors under this simulation scenario. We suspect that the same sensitivity displayed by the AsemiP detector that produced the favorable results shown in table 1 is also responsible for this detector's unfavorable results shown in table 2. A favorable comment from the results shown in table 2 for the AsemiP detector is that it required a relatively high noise to show a significant degradation of its power from unity—70% of noise was required.

In Experiment 1 and Experiment 2, we were primarily interested in presenting observable performance trends by these detectors as a function of systematically varying some parameter(s) in the models. We have designed our simulation experiments using our best interpretation of problematic cases facing anomaly detection techniques in real applications requiring automatic decisions using remotely sensed HS data.

3. Statistical Method Using Idealized Heterogeneous Multivariate Samples

This simulation design imitates the situation encountered in decision based anomaly detection applications using HS or MS imagery, where a pair of multivariate sequences $\mathbf{X}_0$ and $\mathbf{X}_1$ is presented to a detector in order to determine whether they are statistically anomalous to each other**.**

A fundamental problem with anomaly detectors, as discussed in the literature (see, for instance, (*6*)) is that they are not designed to handle mixtures of samples in a way that would be desired for their utility. In particular, since the design of popular detectors are based on class

homogeneity, independently of how samples are compared in HS imagery (e.g., samples representing blocks of data may be compared to samples surrounding these blocks), it is expected that, while testing a typical HS imagery, an anomaly detector will face the challenge of comparing samples from a mixture of multiple classes to samples belonging to one of the classes in that mixture. This challenge is widely ignored in the HS community, as noted in (*6*) using the expression *corrupted samples*, but its consequences cannot be ignored, as it will be demonstrated in this section.

The challenge occurs because, in order to test HS imagery for scene anomalies, data block sizes are chosen to be orders of magnitude smaller than the total image area and, since the locations and scales of objects in the imagery are unknown, a systematic sampling approach is customarily used to test blocks of data across the entire image area. A popular approach is to use a favorite anomaly detector to test overlapping data blocks of the same size against data surrounding each data block so that the entire spatial area of the imagery is tested. Another approach is to use pre-stored spectral data to test blocks of data across the imagery, where it is assumed that these pre-stored spectral data represent only homogeneous background classes. The criterion for selecting one approach over the other is the sensor's viewing perspective with respect to the scene, but since the concerning challenge is essentially the same independently of the sampling approach, we will refer only to the most popular approach, as needed.

Since typical HS data present different class regions across the imaged scene, spatial transitions among these classes are also expected in the data. Therefore, if we consider a block of data representing a single class *C1* and the data surrounding this block representing a mixture of classes (*C1* and one or more additional classes), then, given that a block of data is normally orders of magnitude smaller than the total spatial area of the imagery, one would expect to observe a significant number of similar cases across the imagery. The problem facing anomaly detectors then is a relative one. Let *Z1* be the outcome of a given detector at location (*x1*, *y1*) in a HS imagery when a block of data representing class *C1* is compared to surrounding samples of a mixture that includes *C1*, and let *Z2* be the outcome of this detector at location (*x2*, *y2*) when a block of data representing *C2* is compared to surrounding samples representing class *C3*. In this context, it would be desired to have $Z2 >> Z1$ (>> much greater than), since in the target community cutoff thresholds are attained empirically using, for instance, the method of quantile, and these thresholds will be influenced by the detector's higher value responses for cases representing output *Z1*. Existing anomaly detectors often produce $Z2 < Z1$, or $Z2 << Z1$, when exposed to this typical scenario. The main focus in this report was to design anomaly detectors that can perform desirably (i.e., $Z2 >> Z1$) when confronted with these sorts of relative performances on real HS imagery.

There have been some attempts (see, for instance, (*9*) and (*10*)) to model blocks of HS/MS data as mixture of densities, but they had very limited successes because these models assume normality of individual classes in the data and a prior knowledge of the maximum number of

classes in the block of data. These methods are also cumbersome to implement because the proportion of each class in the mixture is unknown and must be estimated (*10*).

The goal of this simulation is to use a standard statistical test to compare competing detectors and to relate results to what will be referred herein as *desired* and *undesired* behaviors of anomaly detectors. Relating results to desired behavior or undesired behavior should shed some light on expectations of anomaly detectors' performances in applications using real HS imagery.

The data structure of $\mathbf{X}_0$ and $\mathbf{X}_1$ will be specified and shared by competing detectors, and specific null and alternative models will be specified as mixtures of multivariate normal distributions for windows of vector (multispectral, simplified to 5 dimensions) data.

For simplicity, only two anomaly detectors will be compared, and for convenience we choose again to use the AsemiP detector and the industry standard RX detector, (see their mathematical representations in appendix). Since the AsemiP method was proposed as a two-step approach, data preprocessing and test statistic. An attempt will be made to separate performance results owing to the data preprocessing step and to the test statistic step. The RX method, as mentioned earlier, was proposed as a single step approach—a multivariate test statistic.

Using the denotation shown in section 1, we use in this simulation sample pair possibility H_{04} to represent the entire null hypothesis H_0, and H_{13} and H_{14} in different experiments to represent the entire H_1. Using H_{13} as the alternative, and modeling C_{K+1} as C_1 (see section 1) plus a bias term and a shape term, allow us to expose key sensitivity differences of both detectors and also to validate the effectiveness of our data preprocessing step.

Using H_{14} as the alternative, and systematically varying the contribution proportions of the classes in $M(C_1, C_2, C_3, C_{K+1})$ under H_1 (in this case H_{14}), allowed us to check the robustness of both detectors as a function of an alternative that became gradually harder for both detectors relative to the null hypothesis. This latter simulation experiment measured the sensitivity of a test statistic when it receives as input samples of mixtures under both H_0 and H_1. Since the AsemiP detector consists of two steps, data preprocessing and test statistic, it had to rely on its test statistic step to sustain a better power relative to the RX detector's power, because data preprocessing techniques are proposed to promote separation of homogeneous classes that happen to be less separable in their original space.

3.1 Data Structure and Representation of the Null/Alternative Hypotheses

In this session, we describe data models whose samples will be generated and used as the detector's input pair $\mathbf{X}_0$ and $\mathbf{X}_1$ in order to represent the null and alternative hypotheses.

Null Hypothesis (H_0): Fixing the number of classes in the mixture to *3*, we use an idealized case for H_0 where n_0 multivariate samples $\mathbf{X}_{01}, \mathbf{X}_{02}, \cdots, \mathbf{X}_{0n_0}$ are drawn independently from a class mixture density $g_0(\mathbf{X}|\boldsymbol{\theta})$, as shown in (5), and another n_1 samples $\mathbf{X}_{11}, \mathbf{X}_{12}, \cdots, \mathbf{X}_{1n_1}$ are drawn

independently from a class-conditional *normal* probability density $g(\mathbf{X}\,|\,\boldsymbol{\theta}_1)$, such that $g(\mathbf{X}\,|\,\boldsymbol{\theta}_1)$ represents one of the classes in $g_0(\mathbf{X}\,|\,\boldsymbol{\theta})$ as follows:

$$g_0(\mathbf{X}\,|\,\boldsymbol{\theta}) = \sum_{k=1}^{3} g(\mathbf{X}\,|\,\boldsymbol{\theta}_k)\pi_k \ , \tag{5}$$

where $\mathbf{X}$ is a *p*-dimensional column vector, $g(\mathbf{X}\,|\,\boldsymbol{\theta}_k)$ is a class-conditional *normal* PDF of unknown parameters $\boldsymbol{\theta}_k$ $(k=1,2,3)$, $\boldsymbol{\theta}_k$ represents the k^{th} parameter set having the *p*-dimensional mean vector $\boldsymbol{\mu}_k$ and the *p* x *p* variance-covariance matrix $\boldsymbol{\Sigma}_k$, π_k is the unknown proportion of class contribution to the mixture $\left(\pi_k : \sum_{k=1}^{3}\pi_k = 1\right)$, and $\boldsymbol{\theta}$ is the full parameter set $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3)$.

Hypothesis H_0, in the framework discussed in section 1 and in this section, corresponds to setting

$$\begin{aligned} \mathbf{X}_0^{(w)} &= \left(\mathbf{X}_{01}^{(w)}, \mathbf{X}_{02}^{(w)}, \cdots, \mathbf{X}_{0n_0}^{(w)}\right) \ \sim \text{IID} \ \ g_0(\mathbf{X}\,|\,\boldsymbol{\theta}) = \sum_{k=1}^{3} g(\mathbf{X}\,|\,\boldsymbol{\theta}_k)\pi_k \\ \mathbf{X}_1^{(w)} &= \left(\mathbf{X}_{11}^{(w)}, \mathbf{X}_{12}^{(w)}, \cdots, \mathbf{X}_{1n_1}^{(w)}\right) \ \sim \ \text{IID} \ \ g(\mathbf{X}\,|\,\boldsymbol{\theta}_1) \end{aligned} \tag{6}$$

where, $w = 1, \cdots, m$ indexes independent trials, $\mathbf{X}_{0u}^{(w)}$ $(u = 1, \cdots, n_0)$ and $\mathbf{X}_{1h}^{(w)}$ $(h = 1, \cdots, n_1)$ are *B*-dim column vectors, and $\mathbf{X}_0^{(w)}$ and $\mathbf{X}_1^{(w)}$ are independent.

To relate (6), for instance, to the top view anomaly detection problem, consider instead $w = 1, \cdots, m$ indexing independent spatial locations in the HS imagery, where model (6) would be reasonably satisfied, such that, $\mathbf{X}_1^{(w)}$ would represent a reformatted block of data W_1 (centered at *w*) and $\mathbf{X}_0^{(w)}$ would represent reformatted data from W_1's outer ring W_0. (The term *reformatted* indicates that pixels representing a spatial area, e.g., a block, is represented as a sequence as shown in (5) from 1 to the maximum number of pixels within that spatial area.)

In essence, H_0 using (6) represents a detector being introduced to a multivariate sample belonging to a mixture of classes and a multivariate sample belonging to one of the classes in that mixture.

Alternative Hypothesis (H_1): In order to gain enough flexibility in H_1, i.e., be able to use the same model to represent different levels of difficulty for this hypothesis, H_1 will be represented by setting $\mathbf{X}_0^{(w)}$ to the same mixture in (6) and, adding a 4th class to that mixture, by setting $\mathbf{X}_1^{(w)}$ to a new four-class mixture, or

$$\begin{aligned} \mathbf{X}_0^{(w)} &= \left(\mathbf{X}_{01}^{(w)}, \mathbf{X}_{02}^{(w)}, \cdots, \mathbf{X}_{0n_0}^{(w)}\right) \ \sim \text{IID} \ \ g_0(\mathbf{X}\,|\,\boldsymbol{\theta}) = \sum_{k=1}^{3} g(\mathbf{X}\,|\,\boldsymbol{\theta}_k)\pi_k \\ \mathbf{X}_1^{(w)} &= \left(\mathbf{X}_{41}^{(w)}, \mathbf{X}_{42}^{(w)}, \cdots, \mathbf{X}_{4n_1}^{(w)}\right) \ \sim \ \text{IID} \ \ g_1(\mathbf{X}\,|\,\boldsymbol{\theta}) = \sum_{k=1}^{4} g(\mathbf{X}\,|\,\boldsymbol{\theta}_k)\rho_k \end{aligned} \tag{7}$$

where $\mathbf{X}_0^{(w)}$ and $\mathbf{X}_1^{(w)}$ are independent, $w = 1, \cdots, m$ indexes independent trials, $\mathbf{X}_{0u}^{(w)}$ $(u = 1, \cdots, n_0)$ was defined in equation 6, *B-dim* column vectors $\mathbf{X}_{4h}^{(w)}$ $(h = 1, \cdots, n_1)$ are drawn independently from the class mixture density $g_1(\mathbf{X} \mid \boldsymbol{\theta})$, such that all of the class component are normal densities as in (1), and that the normal density $g(\mathbf{X} \mid \boldsymbol{\theta}_4)$ does *not* represent one of the classes in the mixture density $g_0(\mathbf{X} \mid \boldsymbol{\theta})$; the parameter $\boldsymbol{\theta}_4$ represents $(\boldsymbol{\mu}_4, \boldsymbol{\Sigma}_4)$, where the *B*-dim vector $\boldsymbol{\mu}_4$ is the unknown mean and the *B* x *B* matrix $\boldsymbol{\Sigma}_4$ is the unknown covariance matrix of a normal multivariate distribution; the mixture proportions ρ_k are defined such that $\sum_{k=1}^{4} \rho_k = 1$.

Hypothesis H_1, using equation 7, then represents a detector being introduced to a multivariate sample belonging to a mixture of classes and to another multivariate sample belonging to that mixture plus an additional class. Notice that the proportion parameters $(\rho_1, \rho_2, \rho_3, \rho_4)$ will determine the level of difficulty imposed on the detector by H_1 with respect to H_0. For instance, H_1 would correspond to H_0 by setting $(\rho_1, \rho_2, \rho_3, \rho_4) = (1.0, 0.0, 0.0, 0.0)$, and by using a small variation of this setting, e.g., $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.9, 0.0, 0.0, 0.1)$ with $\boldsymbol{\mu}_4 \neq \boldsymbol{\mu}_k$ $(k = 1, \cdots, 3)$, H_1 would represent a *hard* alternative hypothesis for a detector to detect; alternatively, by setting $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.0, 0.0, 0.0, 1.0)$ with $\boldsymbol{\mu}_4 \neq \boldsymbol{\mu}_k$ $(k = 1, \cdots, 3)$, H_1 would represent an *easy* alternative hypothesis using the same detector.

3.2 Simulation Plan

The following steps will be used in order to conduct the simulation experiments using the data model described in subsection 3.1:

i. Setting sample sizes to be equal $(n = n_0 = n_1)$, multivariate samples $\mathbf{X}_{0j}^{(w)}$, $\mathbf{X}_{1j}^{(w)}$, and $\mathbf{X}_{4j}^{(w)}$ $(j = 1, \cdots, n)$, as shown in (6) and (7), consist of covariates b_p representing material radiances at *p* different frequency bands, where $1 \leq p \leq 5$ in this simulation. Covariates b_p will be highly correlated by design in this simulation, which is the case for real HS imagery. These covariates are pseudo values simulated every trial from multivariate normal distributions or mixture of these distributions, as shown in (6) and (7). Parameter values for these distributions are not based on spectroscopy physics, since in real HS imagery $1 \leq p \leq B$ and $B >> 5$. Trials are indexed by *w*, where $w = 1, \cdots, m$, and the sample size n is fixed to *500*.

ii. Since the industry standard RX detector assumes that its input sample pair is composed of IID multivariate normal samples of unknown means and unknown but equal covariance matrices, for this simulation we fix at once covariance matrices $\boldsymbol{\Sigma}_k$ $(k = 1, \cdots, 4)$ to be equal to $\boldsymbol{\Sigma}$, as shown in subsection 3.3.3, where the correlation parameters used in $\boldsymbol{\Sigma}$ are all fixed to *1.0*. Values of the mean parameter vectors μ_k $(k = 1, \cdots, 3)$ in the mixture density in the null hypothesis equation 6 and in the alternative hypothesis equation 7 are also fixed, as shown in subsection 3.3.3, while values of $\boldsymbol{\mu}_4$ in the alternative hypothesis and mixture

proportions $\rho_k\,(k=1,\cdots,4)$ in equation 7 will vary as shown later. For analysis purposes, the value of $\boldsymbol{\mu}_4$ is composed of three parts, or $\boldsymbol{\mu}_4 = \boldsymbol{\mu} + \Delta_B + \Delta_S$, where $\boldsymbol{\mu}$ will be fixed, but Δ_B (a magnitude or bias vector term) and Δ_S (a shape vector term) will vary, see subsection 3.3.1 and subsection 3.3.3 for parameter settings and justification.

iii. Attaining Cutoff Thresholds: Setting $(\pi_1, \pi_2, \pi_3) = (0.4, 0.3, 0.3)$, once and for all, values of $\{\mathbf{X}_0^{(w)}\}_{w=1}^{m}$ and $\{\mathbf{X}_1^{(w)}\}_{w=1}^{m}$ (*m = 1500*) will be generated according to model equation 6 representing samples under the null hypothesis and introduced to both detectors. Note that the setting of $\pi_k\,(k=1,\cdots,3)$ in equations 6 or 7 is independent from the setting of $\rho_k\,(k=1,\cdots,4)$ in equation 7, see (*v*) in this subsection. Both detectors will yield *1500* output results each. Each set of *1500* results will be used to estimate an empirical PDF, and a cutoff threshold will be attained using the standard quantile method on the estimated PDF. The desired type I error probability (ε) for the quantile method is fixed at once for both detectors to $\varepsilon = 0.05$.

iv. Estimating the Type I Error: Additional values of $\{\mathbf{X}_0^{(w)}\}_{w=1}^{m}$ and $\{\mathbf{X}_1^{(w)}\}_{w=1}^{m}$ (*m = 1500*) will be generated according to (*iii*) and introduced to both detectors, where the detectors' corresponding cutoff thresholds will be applied to the corresponding detectors' output results in order to estimate the type I error $(\hat{\varepsilon})$ of each detector. The type I error is estimated by counting the number of trials m_1 that satisfy the detector's output values being greater than the detector's cutoff threshold and computing the ratio $\hat{\varepsilon} = m_1 / m$. (Additional samples will be generated for this step to ensure that the sample size is adequate for the estimation of ε.)

v. Estimating the Power (*1.0* – Type II Error): Using $(\pi_1, \pi_2, \pi_3) = (0.4, 0.3, 0.3)$ and combinations of $\rho_k\,(k=1,\cdots,4)$ and $\boldsymbol{\mu}_4$ (mean of the 4th class in the mixture shown in (7)), values of $\{\mathbf{X}_0^{(w)}\}_{w=1}^{m}$ and $\{\mathbf{X}_1^{(w)}\}_{w=1}^{m}$ (*m = 1500*) will be generated according to model (7) and introduced to both detectors, where the detectors' corresponding cutoff thresholds will be applied to the corresponding detectors' output results in order to estimate the *power* $(\hat{\theta})$ of each detector for a given set of combinations $\rho_k\,(k=1,\cdots,4)$ and $\boldsymbol{\mu}_4$. (Justification for the values of these combinations will be discussed later.) The type II error is estimated by counting the number m_2 of trials that satisfy the detector's output values being lower then the detector's cutoff threshold and computing the ratio m_2 / m, such that $\hat{\theta} = 1.0 - (m_2 / m)$.

vi. Estimating Performance Confidence Intervals (CI): Results from (*iv*) and (v) do not provide insights into the performance variability of the two detectors. In order to check this variability, confidence intervals for (iv) $\bar{\varepsilon} \pm z_{\alpha/2}\sqrt{\frac{1}{R-1}\sum_{r=1}^{R}(\hat{\varepsilon}_r - \bar{\varepsilon})^2}$ and for (v)

$\bar{\theta} \pm z_{\alpha/2}\sqrt{\frac{1}{R-1}\sum_{r=1}^{R}\left(\hat{\theta}_r - \bar{\theta}\right)^2}$ will be estimated, where $z_{\alpha/2} = \Phi^{-1}(1-\alpha/2)$ is the $1-\alpha/2$ quantile $(\alpha = 0.05)$ of the standard normal distribution, random quantities simulated will be indexed by *r* for a total number of simulation repetitions *R=2000*, $\bar{\varepsilon} = R^{-1}\sum_{r=1}^{R}\hat{\varepsilon}_r$, and $\bar{\theta} = R^{-1}\sum_{r=1}^{R}\hat{\theta}_r$.

3.3 Simulation Results

In order to separate subtle differences between the AsemiP anomaly detector and the RX anomaly detector, we decomposed the mean vector $\boldsymbol{\mu}_4$ (4^{th} class in the alternative hypothesis) into three parts: a baseline vector $\boldsymbol{\mu}$, a bias vector Δ_B, and a shape vector Δ_S. The 4^{th}-class mean vector then is represented by

$$\boldsymbol{\mu}_4 = \boldsymbol{\mu} + \Delta_B + \Delta_S , \tag{8}$$

where $\Delta_B = \begin{bmatrix} a \\ a \\ a \\ a \\ a \end{bmatrix}$; $a \geq 0$ (scalor), $\Delta_S = \begin{bmatrix} s \\ -s \\ s \\ -s \\ s \end{bmatrix}$; $s \geq 0$ (scalor), and $\boldsymbol{\mu}$ is fixed.

Notice that Δ_B changes only the bias (also interpreted as magnitude) of $\boldsymbol{\mu}$ and Δ_S changes only the shape of $\boldsymbol{\mu}$. This decomposition will allow us to show the difference between *desired* and *undesired* performances of anomaly detectors through the influence of bias and shape of the 4^{th}-class mean vector in the alternative hypothesis equation 7.

***Justification for specifying parameter values*:** To show the effect of different combinations in our model's high dimensional parameter space, we divided the exposition into two parts: (a) to show the effect of spectral bias/shape in the power performances of both detectors and (b) to show the effect of a difficult H_0 (class mixtures) in the power of these detectors as they test different complexity levels under H_1 (easy, moderate, and hard) relative to H_0.

The criterion for exploring different parameter combinations will not be aimed at favoring one detector relative to the other, but instead at measuring power performance degradation of the AsemiP detector and observing the RX detector's behavior. In order to achieve this goal, we used arbitrary values for some of the parameters to initiate the process and followed with systematic selections of other parameters, as shown in subsection 3.3.1 and subsection 3.3.2. Subsection 3.3.1 discusses the effects of spectral bias/shape differences on both detectors, using the simulation plan outlined in subsection 3.2 for $(\pi_1, \pi_2, \pi_3) = (0.4, 0.3, 0.3)$ in equations 6 and

7, $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.0, 0.0, 0.0, 1.0)$ in equation 7, $\boldsymbol{\mu} = \boldsymbol{\mu}_1$ (see subsection 3.3.1) and three combinations of parameter values (a, s). Subsection 3.3.2 discusses the effect of mixture proportions using the same simulation plan over eight combinations of parameter values $(\rho_1, \rho_2, \rho_3, \rho_4)$ and seven combinations of parameter values (a, s) for $\boldsymbol{\mu} = \boldsymbol{\mu}_1$.

3.3.1 Bias and Shape Influence in Detectors' Performances

This subsection contains results from an experiment using the simulation plan outlined in subsection 3.2 and fixing the values of $\boldsymbol{\mu} = \boldsymbol{\mu}_1$, $(\pi_1, \pi_2, \pi_3) = (0.4, 0.3, 0.3)$, and $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.0, 0.0, 0.0, 1.0)$.

Recall that the mean $\boldsymbol{\mu}_1$ and its relationship to $\boldsymbol{\mu}_2$ and $\boldsymbol{\mu}_3$ were chosen arbitrarily, although their general shapes follow the concave shape of typical (unprocessed) background spectra. The parameters labeled as A, B, and C in (9) were chosen to show the performance difference of both detectors facing bias and/or shape differences under H_1 so that we can discuss what is *desired* versus *undesired* behavior of these detectors in the context of our application using HS VNIR data.

Table 4 shows the 95% confidence intervals (*CI* 95%) for the estimated type I error $\hat{\varepsilon}_r \ (r = 1 \cdots R)$ and estimated power $\hat{\theta}_r \ (r = 1 \cdots R)$ corresponding to both detectors for three combinations of parameter values (a, s):

$$\begin{array}{cccc} & \text{A} & \text{B} & \text{C} \\ (a, s) = & (20, 5) & (20, 0) & (0, 5) \end{array}. \tag{9}$$

The sample size was fixed to $n = 500$ and the total number of trials to $m = 1500$ as described in subsection 3.2 (*i*). The total number of simulation repetitions was fixed to *R=2000*, and the desired type I error as described in subsection 3.2 (*iii*) was fixed to $\varepsilon = 0.05$. The non-zero parameter choices in equation 9 reflect relative high differences (magnitude and/or shape) between $\boldsymbol{\mu}_4$ and $\boldsymbol{\mu}_1$.

Results shown in table 4 shed some light on separating the contribution of the data preprocessing stage of the AsemiP approach from its test statistic, and they also help to determine whether the detectors' performances are desired or undesired once related to the actual application.

Notice in table 4 that the bias and shape of the 4$^{\text{th}}$-class mean vector in the alternative hypothesis equation 7 do not interfere with the type I error estimates, as shown in the appropriate column in table 4 for both detectors, as it is expected since this class—using the parameter settings as shown in table 4—is not included in the null hypothesis equation 6.

Table 4. Illustration of bias and shape influence in the power of each detector (*Det*). Parameter (*Par*) combinations (a, s) labeled A, B, C according to the scheme of (5) are shown in the column under *Par*.

$n = 500$ $m = 1500$ $R = 2000$ $r = 1, \cdots, R$ $\varepsilon = 0.05$		$(\pi_1, \pi_2, \pi_3) = (0.4, 0.3, 0.3)$ $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.0, 0.0, 0.0, 1.0)$			
		Type I $(\hat{\varepsilon}_r; CI\ 95\%)$		*PD or Power* $(\hat{\theta}_r; CI\ 95\%)$	
Det	*Par*	*Lower Bound*	*Upper Bound*	*Lower Bound*	*Upper Bound*
AsemiP	A	0.049689992	0.054310007	1.000000000	1.000000000
	B	0.048396724	0.053763275	0.049344224	0.053815775
	C	0.048724038	0.053715961	1.000000000	1.000000000
RX	A	0.048546493	0.052733506	1.000000000	1.000000000
	B	0.048646478	0.052673521	1.000000000	1.000000000
	C	0.044824983	0.051990016	0.047146621	0.056852153

On the other hand, for the 4th-class mixture proportion $\rho_4 = 1.0$, the parameter combination labeled A yields a sufficient separation between H_0 and H_1, allowing both detectors to demonstrate a perfect power using the simulation plan outlined in subsection 3.2, (see column under *Power* and row A for both detectors in table 4).

Parameter combination labeled B completely removes the shape difference between $(\mathbf{X}_{11}, \mathbf{X}_{12}, \cdots, \mathbf{X}_{1n})$ and $(\mathbf{X}_{41}, \mathbf{X}_{42}, \cdots, \mathbf{X}_{4n})$, (see equations 6 and 7), but preserves a significant bias between both samples. Since the AsemiP data preprocessing step was designed to remove the influence of bias and preserve shape differences, the AsemiP anomaly detector cannot distinguish between the simulated sample pair in equation 6 and the simulated sample pair in equation 7. Therefore, its CI corresponding to power is equivalent to its CI corresponding to type I error, as shown in table 4. This performance behavior is, in fact, desired for anomaly detectors because, in the context of our application, robustness also implies that the response of a detector toward an imaged object should be invariant to changes in spectral magnitude. For example, if the detector's response to an object is invariant independently of whether this object is in shadow or outside shadow, in a cloudy day or in a sunny day, then this detector would be considered more robust than a competing detector that is sensitive to spectral magnitude changes imposed in the same object. Paradoxically, HS anomaly detectors that are found in the engineering literature (see for instance, (*6*), (*11*)) are sensitive to spectral magnitude differences, including the industry standard detector (*2*), RX, see table 4, under label B and *Power* under the RX detector.

Parameter combination labeled C, in the contrary, completely removes the bias (magnitude) while holding a significant shape difference between $(\mathbf{X}_{11}, \mathbf{X}_{12}, \cdots, \mathbf{X}_{1n})$ and $(\mathbf{X}_{41}, \mathbf{X}_{42}, \cdots, \mathbf{X}_{4n})$.

Since the AsemiP detector emphasis is on shape differences between multivariate samples, it yielded a perfect power, see table 4, while the RX detector failed to distinguish this important difference between sample pairs using model (6) and sample pairs using model (7), as this detector is sensitive only to spectral magnitude differences. The target community will be interested in seeing these results, which were made possible by decomposing the 4^{th} class mean vector as shown in equation 8.

The main impact of making a detection approach sensitive to spectral shape differences and insensitive to spectral magnitude differences is that of gaining robustness, e.g., gaining the ability to detect objects in different illumination conditions and most importantly under certain types of camouflaging materials. For instance, materials composing sniper camouflage suits provide an average spectral magnitude that is comparable to the average spectral magnitude of natural canopy, but most of these camouflage suits have spectral shapes that are different from natural canopy in other regions of the spectrum outside the visible region—these findings were reported in (*12*). A data preprocessing technique that can augment the detector's sensitivity to spectral shape differences is then desired for certain types of camouflaged object detection applications.

Unfortunately, the detector's sensitivity to spectral shape differences and insensitivity to spectral magnitude differences are not sufficient to deal with an additional important issue that often prevents the desired effectiveness of an anomaly detection approach: the detector's inability to deal with mixtures of spectral samples in HS imagery. A detector that can deal with mixtures will be able to maintain cutoff thresholds relatively low under the challenging null hypothesis in equation 6, which is why we used a mixture of densities to conduct performance analyses via simulations in this document. Thus, to complement the benefit of using the AsemiP data preprocessing stage, we proposed the AsemiP test statistic to handle samples from mixture of densities. Subsection 3.3.2 presents performance results of both detectors under changes of proportion parameters in the alternative hypothesis in (7) over a baseline combination set of (a, s).

3.3.2 Mixture Proportions' Influence in Detectors' Performances

In this subsection, we aim at showing the benefit of using anomaly detectors that can maintain a relatively low cutoff threshold under the challenging null hypothesis in equation 6, as they test the alternative hypothesis in equation 7 for different parameter combinations. This assessment can be made by recording the estimated power of these detectors, as the proportion of the 4^{th} class in equation 7 varies from unity to smaller values. Our conjecture is that a detector that yields a relatively low cutoff threshold from the null hypothesis in equation 6 will yield a superior power than otherwise for smaller proportions of the 4^{th} class in the alternative hypothesis in equation 7.

Since both anomaly detectors (RX and AsemiP) are sensitive to different spectral sample features (magnitude or shape), we will ignore in this part of the simulation whether the detectors'

outcomes for the alternative hypothesis are desired or undesired, and calibrate instead their performances to a baseline, i.e., we will find a combination of parameters that will cause the detectors to perform about the same and name it: *calibrated performances*. We explored various combinations of parameter values $(\rho_1, \rho_2, \rho_3, \rho_4, a, s)$ and attained calibrated performances using parameter combination $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.0, 0.0, 0.0, 1.0)$ and seven labeled combinations of (a, s):

$$\begin{array}{cccc} & \text{A1} & \text{A2} & \text{A3} \\ (a, s) = & (20.0, 5.0) & (17.6, 4.20) & (15.1, 3.60) \end{array}$$

$$\begin{array}{cccccc} & \text{A4} & & \text{A5} & \text{A6} & \text{A7} \\ (a, s) = & (11.8, 2.30) & (8.3, 1.53) & & (5.2, 0.90) & (0.00, 0.00) \end{array} \quad , \qquad (10)$$

where, for the given parameter combination $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.0, 0.0, 0.0, 1.0)$, the label combination A1 presents the *easiest* alternative hypothesis for both detectors to detect, and A7 presents the *hardest* alternative hypothesis for both detectors.

Parameters labeled as A1 through A7 in (10) and other parameters that led to the results presented in table 4 were chosen to show a gradual deterioration of both detectors (using the degradation of AsemiP as a reference, observe degradation of RX) and be marked as a *calibrated* benchmark of results so that we could precede showing further deterioration of these detectors' performances as a function of mixture proportions $\rho_k \ (k = 1, \cdots, 4)$ under H_1 for a fixed mixture proportions $\pi_k \ (k = 1, \cdots, 3)$ under H_0. The $\pi_k \ (k = 1, \cdots, 3)$ were arbitrarily set to be approximately equal to each other, and ρ_4 was gradually decreased (while the other $\rho_k \ (k = 1, \cdots, 3)$ were set approximately equal to each other satisfying the unity requirement $\sum_{k=1}^{4} \rho_k = 1.0$ in order to observe performance degradation of both detectors for different levels of complexities under H_1.

Performance results that were not shown in the document using different combinations of these parameters were discarded because they were either redundant or did not show significant differences from the results shown in the document.

The results shown in table 5 are the calibrated performances. Both detectors perform approximately the same using the simulation plan outlined in subsection 3.2 and parameter combinations shown in equation 6 for $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.0, 0.0, 0.0, 1.0)$. Tables 6 through 12 show results using the same simulation plan and seven parameter combinations of $(\rho_1, \rho_2, \rho_3, \rho_4)$ over the seven labeled combinations according to scheme in equation 6. Each table, from table 5 through table 12, corresponds to a fixed parameter combination of $(\rho_1, \rho_2, \rho_3, \rho_4)$ over the same seven labeled combinations of (a, s).

Table 5. Detectors' calibrated oerformances.

$n = 500$ $m = 1500$ $R = 2000$ $r = 1, \cdots, R$ $\varepsilon = 0.05$		$(\pi_1, \pi_2, \pi_3) = (0.4, 0.3, 0.3)$ $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.0, 0.0, 0.0, 1.0)$			
		Type I $(\hat{\varepsilon}_r; CI\ 95\%)$		*Power* $(\hat{\theta}_r; CI\ 95\%)$	
Det	*Par*	*Lower Bound*	*Upper Bound*	*Lower Bound*	*Upper Bound*
AsemiP	A1	0.049689992	0.054310007	1.000000000	1.000000000
	A2	0.048649929	0.053230070	0.947422544	0.958537455
	A3	0.047847406	0.052952593	0.764110188	0.782769811
	A4	0.046972776	0.051427223	0.510582052	0.529857947
	A5	0.047973420	0.052466579	0.286606531	0.300793468
	A6	0.047019216	0.051540783	0.068192684	0.073687315
	A7	0.047060977	0.051939022	0.046682036	0.051597963
RX	A1	0.049315748	0.053444251	1.000000000	1.000000000
	A2	0.047081616	0.052158383	0.842231671	0.916848328
	A3	0.047855700	0.052984299	0.449845730	0.517754269
	A4	0.049995794	0.054604205	0.256048622	0.279871377
	A5	0.050133742	0.054546257	0.170833174	0.187046825
	A6	0.049551476	0.053568523	0.071805233	0.080074766
	A7	0.047149983	0.051530016	0.048550856	0.053129143

Table 5 shows that the type I error CI estimates using both detectors included the desired type I error $(\varepsilon = 0.05)$, as expected. The actual values of these detectors' power performances shown in table 5 are unimportant, but what is important is the fact that they decrease gradually using labeled combinations A1 through A7. Notice that the power performances of both detectors using *Par* A7 correspond in essence to these detectors' type I error performances, since the shape difference and bias were removed from the 4th class mean vector, implying that the alternative hypothesis corresponds to the null hypothesis, or $\boldsymbol{\mu}_4 = \boldsymbol{\mu}_1$, since $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.0, 0.0, 0.0, 1.0)$.

Now, we will increase the difficulty level of the alternative hypothesis and measure its effect on these detectors. The difficulty level can be increased by changing the values of $(\rho_1, \rho_2, \rho_3, \rho_4)$. From table 6 to table 12, the parameter combinations $(\rho_1, \rho_2, \rho_3, \rho_4)$ were set as follows:

Table 6 – $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.02, 0.02, 0.01, 0.95)$,

Table 7 – $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.04, 0.03, 0.03, 0.90)$,

Table 8 – $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.07, 0.07, 0.06, 0.80)$,

Table 9 – $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.14, 0.13, 0.13, 0.60)$,

Table 10 – $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.20, 0.20, 0.20, 0.40)$,

Table 11 – $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.27, 0.27, 0.26, 0.20)$, and

Table 12 – $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.32, 0.32, 0.31, 0.05)$,

where these tables were organized, such that, for a fixed labeled combination (e.g., A4), the alternative hypothesis ranges from the easiest (see table 5) to the hardest (table 12), and, for all of the results presented in these tables, the parametric combination *Par* A1 in table 5 represents the absolute easiest alternative hypothesis in this simulation and *Par* A7 in table 12 represents the absolute hardest alternative hypothesis.

The estimated type I error CI presented in table 5 for both detectors are replicated for table 6 through table 12, since the parameter combinations $(\rho_1, \rho_2, \rho_3, \rho_4)$ do not affect the type I error estimates. The parameter changes shown in table 6 do not significantly affect the estimated power CI of both detectors, relative to their calibrated performances, as the contribution of the anomalous class (4^{th} class, which is in H_1 but not in H_0) is still relatively high $(\rho_4 = 0.95)$.

Table 6. Performances using anomalous-class contribution $\rho_4 = 0.95$.

$n = 500$ $m = 1500$ $R = 2000$ $r = 1, \cdots, R$ $\varepsilon = 0.05$		$(\pi_1, \pi_2, \pi_3) = (0.4, 0.3, 0.3)$ $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.02, 0.02, 0.01, 0.95)$			
		Type I $(\hat{\varepsilon}_r; CI\ 95\%)$		*Power* $(\hat{\theta}_r; CI\ 95\%)$	
Det	*Par*	*Lower Bound*	*Upper Bound*	*Lower Bound*	*Upper Bound*
AsemiP	A1	0.049689992	0.054310007	1.000000000	1.000000000
	A2	0.048649929	0.053230070	0.942281875	0.954998124
	A3	0.047847406	0.052952593	0.751184110	0.770335889
	A4	0.046972776	0.051427223	0.518196795	0.537163204
	A5	0.047973420	0.052466579	0.287309548	0.302330451
	A6	0.047019216	0.051540783	0.074645887	0.080154112
	A7	0.047060977	0.051939022	0.053755197	0.058604802
RX	A1	0.049315748	0.053444251	1.000000000	1.000000000
	A2	0.047081616	0.052158383	0.825180772	0.903339227
	A3	0.047855700	0.052984299	0.455990555	0.521049444
	A4	0.049995794	0.054604205	0.263348966	0.290131033
	A5	0.050133742	0.054546257	0.174690052	0.191149947
	A6	0.049551476	0.053568523	0.074816120	0.082583879
	A7	0.047149983	0.051530016	0.047682679	0.053237320

Table 7. Performances using anomalous-class contribution $\rho_4 = 0.90$.

$n = 500$ $m = 1500$ $R = 2000$ $r = 1, \cdots, R$ $\varepsilon = 0.05$		$(\pi_1, \pi_2, \pi_3) = (0.4, 0.3, 0.3)$ $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.04, 0.03, 0.03, 0.90)$			
		Type I $(\hat{\varepsilon}_r; CI\ 95\%)$		*Power* $(\hat{\theta}_r; CI\ 95\%)$	
Det	*Par*	*Lower Bound*	*Upper Bound*	*Lower Bound*	*Upper Bound*
AsemiP	A1	0.049689992	0.054310007	1.000000000	1.000000000
	A2	0.048649929	0.053230070	0.854184598	0.872335401
	A3	0.047847406	0.052952593	0.642609642	0.662750357
	A4	0.046972776	0.051427223	0.408920288	0.427239711
	A5	0.047973420	0.052466579	0.234175038	0.249224961
	A6	0.047019216	0.051540783	0.066402824	0.072517175
	A7	0.047060977	0.051939022	0.048011480	0.053188519
RX	A1	0.049315748	0.053444251	0.075253189	0.080506810
	A2	0.047081616	0.052158383	0.000000000	0.000000000
	A3	0.047855700	0.052984299	0.000000000	0.000000000
	A4	0.049995794	0.054604205	0.000000000	0.000000000
	A5	0.050133742	0.054546257	0.000000000	0.000000000
	A6	0.049551476	0.053568523	0.000000000	0.000000000
	A7	0.047149983	0.051530016	0.000000000	0.000000000

In tables 7 and 8, notice that the power of the AsemiP detector decreases noticeably relative to its calibrated power shown in table 5 for some of the labeled combinations in the *Par* column. Decreasing the proportion of the 4th class in the alternative hypothesis equation 7 from 1.0 to 0.9 or 0.8 was sufficient to cause this degradation of the AsemiP detector; however, under *Par* A1 in tables 7 and 8, this detector shows that it can preserve its calibrated power performance. The reason for this performance is that the AsemiP detector produces a relatively low cutoff threshold under the challenging null hypothesis in equation 4, relative to this detector's response to weaker contributions—although still strong at $\rho_4 = 0.90$ or $\rho_4 = 0.80$—of the anomalous 4th class to the alternative hypothesis.

For the RX detector, results in tables 7 and 8 show a significant degradation of its power, down to a virtual *zero*, as the parameter combination $(\rho_1, \rho_2, \rho_3, \rho_4)$ changes from the one shown in table 5 to the ones shown in tables 7 and 8.

Table 8. Performances using anomalous-class contribution $\rho_4 = 0.80$.

$n = 500$ $m = 1500$ $R = 2000$ $r = 1, \cdots, R$ $\varepsilon = 0.05$		$(\pi_1, \pi_2, \pi_3) = (0.4, 0.3, 0.3)$ $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.07, 0.07, 0.06, 0.80)$			
		Type I $(\hat{\varepsilon}_r; CI\ 95\%)$		*Power* $(\hat{\theta}_r; CI\ 95\%)$	
Det	*Par*	*Lower Bound*	*Upper Bound*	*Lower Bound*	*Upper Bound*
AsemiP	A1	0.049689992	0.054310007	1.000000000	1.000000000
	A2	0.048649929	0.053230070	0.718690650	0.736829349
	A3	0.047847406	0.052952593	0.517011600	0.536228399
	A4	0.046972776	0.051427223	0.330848485	0.346991514
	A5	0.047973420	0.052466579	0.191038123	0.203401876
	A6	0.047019216	0.051540783	0.061276485	0.066363514
	A7	0.047060977	0.051939022	0.047990578	0.053209421
RX	A1	0.049315748	0.053444251	0.000000000	0.000000000
	A2	0.047081616	0.052158383	0.000000000	0.000000000
	A3	0.047855700	0.052984299	0.000000000	0.000000000
	A4	0.049995794	0.054604205	0.000000000	0.000000000
	A5	0.050133742	0.054546257	0.000000000	0.000000000
	A6	0.049551476	0.053568523	0.000000000	0.000000000
	A7	0.047149983	0.051530016	0.000000000	0.000000000

This degradation is dramatic owing to this detector's inability to produce a cutoff threshold that is relatively low under the challenging null hypothesis in equation 6 with respect to this detector's responses under the alternative hypothesis in equation 7 using the labeled parameter combinations shown in (10) and the class proportion combination $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.04, 0.03, 0.03, 0.90)$, or $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.07, 0.07, 0.06, 0.80)$. Tables 7 and 8 show that the RX detector is too sensitive under H_0.

For the anomalous 4th class contributions decreasing even further, tables 9 through 12 present the power degradation of the AsemiP detector, as the alternative hypothesis overlaps significantly more into the null hypothesis, especially for parameter combinations A2 through A5. Under a high shape difference (A1), the AsemiP detector could hold power to its corresponding calibrated performance, including for parameter combination $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.27, 0.27, 0.26, 0.20)$, which in essence shows this detector's relatively low sensitivity under H_0, under which the cutoff threshold was attained from, and relatively high sensitivity to the 4th class' 20% contribution of shape difference from the null hypothesis' 1st class. This behavior is owed to the data preprocessing step and to the AsemiP test statistic's ability to hold the cutoff threshold relatively low under the null hypothesis. Further power deterioration of the AsemiP detector is observed in table 12.

The RX detector's performances using parameter settings, as shown in tables 9 through 12, were unchanged relative to its performance shown in table 8, hence, those results were not included in tables 9 through 12. The RX detector's relatively high sensitivity under H_0 is responsible for this detector's degraded performances.

Table 9. Performances using anomalous-class contribution $\rho_4 = 0.60$.

$n = 500$ $m = 1500$ $R = 2000$ $r = 1, \cdots, R$ $\varepsilon = 0.05$		$(\pi_1, \pi_2, \pi_3) = (0.4, 0.3, 0.3)$ $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.14, 0.13, 0.13, 0.60)$			
		Type I $(\hat{\varepsilon}_r; CI\ 95\%)$		*Power* $(\hat{\theta}_r; CI\ 95\%)$	
Det	*Par*	*Lower Bound*	*Upper Bound*	*Lower Bound*	*Upper Bound*
AsemiP	A1	0.049689992	0.054310007	1.000000000	1.000000000
	A2	0.048649929	0.053230070	0.409792374	0.427767625
	A3	0.047847406	0.052952593	0.285322165	0.300317834
	A4	0.046972776	0.051427223	0.188144397	0.199695602
	A5	0.047973420	0.052466579	0.123314105	0.132245894
	A6	0.047019216	0.051540783	0.057298891	0.062181108
	A7	0.047060977	0.051939022	0.047194843	0.052365156

Table 10. Performances using anomalous-class contribution $\rho_4 = 0.40$.

$n = 500$ $m = 1500$ $R = 2000$ $r = 1, \cdots, R$ $\varepsilon = 0.05$		$(\pi_1, \pi_2, \pi_3) = (0.4, 0.3, 0.3)$ $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.20, 0.20, 0.20, 0.40)$			
		Type I $(\hat{\varepsilon}_r; CI\ 95\%)$		*Power* $(\hat{\theta}_r; CI\ 95\%)$	
Det	*Par*	*Lower Bound*	*Upper Bound*	*Lower Bound*	*Upper Bound*
AsemiP	A1	0.049689992	0.054310007	1.000000000	1.000000000
	A2	0.048649929	0.053230070	0.194112671	0.206447328
	A3	0.047847406	0.052952593	0.143623596	0.152816403
	A4	0.046972776	0.051427223	0.104153482	0.111926517
	A5	0.047973420	0.052466579	0.078379089	0.084820910
	A6	0.047019216	0.051540783	0.052904455	0.058015544
	A7	0.047060977	0.051939022	0.048669129	0.053450870

Table 11. Performances using anomalous-class contribution $\rho_4 = 0.20$.

$n = 500$ $m = 1500$ $R = 2000$ $r = 1, \cdots, R$ $\varepsilon = 0.05$		$(\pi_1, \pi_2, \pi_3) = (0.4, 0.3, 0.3)$ $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.27, 0.27, 0.26, 0.20)$			
		Type I $(\hat{\varepsilon}_r; CI\ 95\%)$		*Power* $(\hat{\theta}_r; CI\ 95\%)$	
Det	*Par*	*Lower Bound*	*Upper Bound*	*Lower Bound*	*Upper Bound*
AsemiP	A1	0.049689992	0.054310007	1.000000000	1.000000000
	A2	0.048649929	0.053230070	0.078764146	0.084755853
	A3	0.047847406	0.052952593	0.068028364	0.073491636
	A4	0.046972776	0.051427223	0.060796030	0.066163969
	A5	0.047973420	0.052466579	0.056403217	0.060676782
	A6	0.047019216	0.051540783	0.048726758	0.053753247
	A7	0.047060977	0.051939022	0.049206628	0.054313371

Table 12. Performances using anomalous-class contribution $\rho_4 = 0.05$.

$n = 500$ $m = 1500$ $R = 2000$ $r = 1, \cdots, R$ $\varepsilon = 0.05$		$(\pi_1, \pi_2, \pi_3) = (0.4, 0.3, 0.3)$ $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.32, 0.32, 0.31, 0.05)$			
		Type I $(\hat{\varepsilon}_r; CI\ 95\%)$		*Power* $(\hat{\theta}_r; CI\ 95\%)$	
Det	*Par*	*Lower Bound*	*Upper Bound*	*Lower Bound*	*Upper Bound*
AsemiP	A1	0.049689992	0.054310007	0.272279737	0.289680262
	A2	0.048649929	0.053230070	0.049244815	0.053755184
	A3	0.047847406	0.052952593	0.049208689	0.054351310
	A4	0.046972776	0.051427223	0.048885476	0.053274523
	A5	0.047973420	0.052466579	0.048577670	0.052662329
	A6	0.047019216	0.051540783	0.045402912	0.049197087
	A7	0.047060977	0.051939022	0.048428880	0.053051119

Results presented in tables 5 through 12 depict detectors' robustness as a function of increased complexity under H_1. In this simulation, the complexity level under H_1 can be significantly increased by gradually decreasing ρ_4 and adjusting ρ_1, ρ_2, and ρ_3, accordingly. Following this approach is equivalent to assessing the detector's ability to maintain a rather uniformly low cutoff threshold, and our conjecture is that, using this approach for comparison, better performers will produce better power.

Table 5 shows results from smaller disparities between H_0 and H_1, using $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.0, 0.0, 0.0, 1.0)$ and parameter combinations A1, A2,…, A7. Tables 6

through 12 show additional results for the reader to assess the robustness of both detectors as a function of A1,…,A7 and increased complexity under H_1.

Results in those tables show that the competing RX detector is not as robust as that of our favorite AsemiP detector as they test difficult alternative hypothesis, and these results are consistent with results obtained by testing these detectors on real HS imagery for the application discussed in this application.

3.3.3 Data Model Parameter Values and Justification

The mean vector of class 1 $(\boldsymbol{\mu}_1)$ was chosen to form a concave shape, which resembles the response of natural terrain in the visible to near infrared region of the spectrum. (The covariate values in this mean were arbitrarily chosen.) The mean vector of class 2 $(\boldsymbol{\mu}_2)$ and the mean vector of class 3 $(\boldsymbol{\mu}_3)$ are offsets of $(\boldsymbol{\mu}_1)$, or

$$\boldsymbol{\mu}_1 = \begin{bmatrix} 10 \\ 60 \\ 100 \\ 80 \\ 30 \end{bmatrix}; \quad \boldsymbol{\mu}_2 = \boldsymbol{\mu}_1 + \begin{bmatrix} 10 \\ 10 \\ 10 \\ 10 \\ 10 \end{bmatrix}; \quad \boldsymbol{\mu}_3 = \boldsymbol{\mu}_1 + \begin{bmatrix} 20 \\ 20 \\ 20 \\ 20 \\ 20 \end{bmatrix}; \quad \boldsymbol{\mu}_2 = \begin{bmatrix} 20 \\ 70 \\ 110 \\ 90 \\ 40 \end{bmatrix}; \quad \boldsymbol{\mu}_3 = \begin{bmatrix} 30 \\ 80 \\ 120 \\ 100 \\ 50 \end{bmatrix}.$$

We fixed at once the values of $\boldsymbol{\Sigma}_i = \boldsymbol{\Sigma} \ \ (i = 1, \cdots, 4)$, where

$$\boldsymbol{\Sigma} = \begin{bmatrix} 10.0000 & 14.1421 & 20.0000 & 14.1421 & 10.000 \\ 14.1421 & 20.0000 & 28.2843 & 20.0000 & 14.1421 \\ 20.0000 & 28.2843 & 40.0000 & 28.2843 & 20.0000 \\ 14.1421 & 20.0000 & 28.2843 & 20.0000 & 14.1421 \\ 10.0000 & 14.1421 & 20.0000 & 14.1421 & 10.0000 \end{bmatrix};$$

for simplicity, all of the correlations imbedded in $\boldsymbol{\Sigma}$ are all equal to *1.0*, and the variances shown in the diagonal of $\boldsymbol{\Sigma}$ were chosen arbitrarily, although their values have a general trend observed in MS or HS data—variability of natural clutter background in the visible to near infrared region of the spectrum tends to increase toward the frequency band center and to decrease thereafter.

We decomposed the mean vector $\boldsymbol{\mu}_4$ (4^{th} class in the alternative hypothesis) into three parts: a baseline vector $\boldsymbol{\mu}$, a bias vector Δ_B, and a shape vector Δ_S. The 4^{th}-class mean vector then is represented by $\boldsymbol{\mu}_4 = \boldsymbol{\mu} + \Delta_B + \Delta_S$, where $\Delta_B = [a \ a \ a \ a \ a]^t; \ a \geq 0$ (scalor), $\Delta_B = [s \ \ -s \ \ s \ \ -s \ \ s]^t; \ s \geq 0$ (scalor), and $\boldsymbol{\mu}$ was fixed at once to $\boldsymbol{\mu} = \boldsymbol{\mu}_1$.

Notice that, by setting $\boldsymbol{\mu} = \boldsymbol{\mu}_1$, Δ_B changes only the bias (also interpreted as magnitude) of $\boldsymbol{\mu}_1$, and Δ_S changes only the shape of $\boldsymbol{\mu}_1$. This decomposition allows us to show the difference between desired and undesired performances of anomaly detectors through the influence of bias and shape of the 4^{th}-class mean vector in the alternative hypothesis equation 7.

3.4 Concluding Remarks

This section focused on analyzing the effect of idealized sample mixtures on performances of two distinct anomaly detection techniques. As in section 2, the AsemiP detector and the RX detector were chosen for this section. The AsemiP detector consists of two steps, data preprocessing and univariate test statistic. The RX detector was proposed as a single step multivariate test statistic, which is also the case for alternative anomaly detection techniques in the literature. Data models were specified and shared by both detectors representing null and alternative hypotheses. Simulated samples were generated according to the models, and a standard statistical test was used to compare detectors' performances as a function of varying parameters.

The RX technique showed high sensitivity to a challenging null hypothesis that is found in real HS data (involving samples of a mixture and samples of a class belonging to that mixture), and showed also high sensitivity to magnitude (bias) differences among generated multivariate samples; both are highly undesirable sensitivities for anomaly detection applications because this detector's power does depend on the detector's null-hypothesis based cutoff threshold and on the illumination (which influences spectral magnitude) on objects in the scene.

The AsemiP technique, on the other hand, showed significantly less sensitivity to the same null hypothesis (owing to its data-combining test statistic), and insensitivity to magnitude differences between multivariate samples (owing to its data preprocessing step). The proposed data preprocessing step also promotes a higher sensitivity of the front-end test statistic to spectral shape differences between multivariate samples.

The results shown in table 4, in essence, verified a desired goal by the AsemiP detector's developer of removing spectral magnitude (bias) differences among spectra through data preprocessing. Table 4 showed two surprising outcomes: first, high effectiveness of the AsemiP data preprocessing step removing bias differences, which prompted the results produced by the AsemiP test statistic step shown in table 4; and, second, high insensitiveness of the RX detector to shape differences among spectra, which is undesired using HS VNIR data, but possibly desired in other regions of the spectrum (e.g., HS LWIR data).

A key point to notice in the sequence of tables after table 4 is that even if sensitivity to spectral bias were desired, RX showed lack of robustness testing harder alternative hypotheses. Results presented in table 5 show *calibrated* power of both detectors for parameter combinations A1-A7, while tables 6 through 12 show deterioration of power of both detectors, as a function of A1-A7

and increased complexity under H_1, relative to the detectors' corresponding (calibrated) power shown in table 5.

Given the results shown in table 4, it is worth noting that if spectral magnitude difference turns out to be an important feature for effective anomaly detection outside the VNIR region of the spectrum, the developers of the AsemiP detector would have to consider a different data preprocessing technique and continue to use a robust test statistic (e.g., AsemiP).

Results presented in tables 5 through 12 depict detectors' robustness as a function of increased complexity under H_1. In this simulation, the complexity level under H_1 can be significantly increased by gradually decreasing ρ_4 and adjusting ρ_1, ρ_2, and ρ_3, accordingly. Following this approach is equivalent to assessing the detector's ability to maintain a rather uniformly low cutoff threshold, and our conjecture is that, using this approach for comparison, better performers will produce better power.

Table 5 shows results from smaller disparities between H_0 and H_1, using $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.0, 0.0, 0.0, 1.0)$ and parameter combinations A1, A2,…, A7. Tables 6 through 12 show additional results for the reader to assess the robustness of both detectors as a function of A1,…,A7 and increased complexity under H_1.

Results in those tables show that the AsemiP detector is more robust than the RX detector as they test difficult alternative hypothesis. These results are consistent with results obtained by testing these detectors on real HS imagery for the application discussed in this application.

4. Statistical Method Using Idealized Multivariate Cubes

This simulation design imitates the situation encountered in decision based anomaly detection applications using HS or MS imagery, where a pair of multivariate sequences $\mathbf{X}_0$ and $\mathbf{X}_1$ is presented to a detector in order to determine whether they are statistically anomalous to each other.

In this section, we seek to demonstrate, under controlled top-view background configuration scenarios, the differences in performances among different anomaly detection approaches using idealized data.

The simulation experiments discussed in section 2 (which used idealized homogeneous samples) and section 3 (which used idealized sample mixture) differs from the simulation experiment discussed in this section, while uses idealized pseudo cubes representing smaller versions of real HS (cubes) imagery. The main goal in this section is to assess the power (correct probability of detection, PD) of competing anomaly detectors as they test idealized multivariate *imagery*, and to use a standard statistical method to assess their power estimates at given type I errors and

equal sample size. To achieve this goal, we start by generating a large number of idealized *training* cubes using three different background configurations, and test these cubes with the various anomaly detectors so that cutoff thresholds and type I errors can be estimated from these detectors' output surfaces. (Each pixel in a given output surface corresponds to a trial outcome using the corresponding detector.) We then generate another large set of idealized *test* cubes using the same three background configurations and added *targets* and test these cubes using the same detectors and their corresponding cutoff thresholds so that their powers can be estimated from their output surfaces. Tabulated results and ROC curves will be shown for this simulation experiment.

Anomaly detection approaches that can maintain a relatively low cutoff threshold using idealized realizations of increasingly more complex background configurations will produce better power—this is our **conjecture**. For this expectation to be true, desired anomaly detection approaches would have to be less sensitive to local background configurations relative to competing approaches.

The detectors chosen for this simulation are three univariate detectors (AsemiP, AFT, and AVT) and five alternative detectors (ANOVA, RX, FLD, EST, and DPC). These detectors' mathematical formulations are shown in the appendix.

This simulation is designed to estimate detectors' performances over a rather large number of experiment trials.

4.1 Problem Formulation

Let B_1, B_2, and B_3 represent three simulated multispectral cubes of increasing background complexity having the same size r x c x b, where r is the number of rows, c is the number of columns, and b is the number of spectral bands. Let B_1 consist of highly correlated multivariate random samples of a single homogeneous class C_1, and B_2 and B_3 consist of equally highly correlated multivariate random samples of multiple homogeneous classes C_k ($k = 1, \ldots, n_c$), where these classes form different stripe patterns between B_2 and B_3, see figure 2. (Statistical details on these classes are presented later).

The configuration of the backgrounds in B_1, B_2, and B_3 are purposively formed so that the null hypothesis become more difficult for detectors to keep a relatively low cutoff threshold, where B_1 represents an easy background configuration, B_2 represents a moderately difficult configuration, and B_3 represents the most difficult configuration.

Now consider the inside-outside window sampling mechanism described section 1 and depicted in figure 2 as dotted boxes at specific positions labeled as a through l, separating the local area into two regions—the inner window region (W_{in} or W_0) and the outer window region (W_{out} or W_1)—sliding concentrically across the area r x c in each simulated cube, such that, at each discrete position in the imagery, multivariate vector samples $\mathbf{x}_{0p} = \left[x_{0p1}, x_{0p2}, \cdots, x_{0pb}\right]^t$ ($p = 1,$

… n_0) that are viewed within W_{out} will be compared in some form to multivariate vector samples $\mathbf{x}_{1q} = \left[x_{1q1}, x_{1q2}, \cdots, x_{1qb}\right]^t$ ($q = 1, \ldots n_1$) that are viewed within W_{in}. The size of the dual window is set such that the W_{in} encloses a target sized region and the W_{out} includes its surrounding region. If the dual window is placed within a spatially homogeneous region consisting of similar types of materials, such as natural backgrounds, the statistical characteristics of samples that are observed within W_{in} and W_{out} will be similar to each other. Samples within W_{in} and W_{out} will contain significantly different statistical features if the dual window is centered on a region where a target, for instance, is surrounded by its local background. Use of appropriate cutoff thresholds on anomaly detectors' outputs would allow most targets to be detected as local anomalies.

A proportionally sized dual rectangular window with respect to the cubes' sizes is shown at different positions on B_2, and B_3, (see figure 2 (top)). Depending on the detection technique being used, these multivariate samples $\mathbf{x}_{0p}$ and $\mathbf{x}_{1q}$ will be transformed (preprocessed) into two sequences $y_0 = \left(y_{01},\ y_{02},\ \cdots,\ y_{0n_0}\right)$ and $y_1 = \left(y_{11},\ y_{12},\ \cdots,\ y_{1n_0}\right)$ for comparison. (Notice that both sequences have the same size, i.e., n_0, which is independent of n_1. Details on this transformation are discussed in appendix.)

The main objective in this simulation is twofold: (i) to apply the detectors on background-only imagery in order to obtain critical values (cutoff thresholds)—these background only data will be referred to in this document as training cubes, and (ii) to test a large number of simulated data using these critical values in order to estimate confidence intervals for type I and type II errors—simulated data mentioned in this fold will be referred to as test cubes. To obtain these critical values, detectors will test realizations of B_1, B_2, and B_3 and have their output surfaces results *sliced* to form binary images, where the number of pixel values above slicing thresholds can be used to yield predetermined error probabilities—in this case, type I errors. A single critical value will be dependent on three factors: the dual window size, which will be constant; the specific detector being applied; and the requested type I error. To estimate error confidence intervals, a large number of B_1, B_2, and B_3 will be realized, *targets* will be added in these cubes, and these target-background simulated data will be tested using the different detectors and their corresponding background specific critical values. Simulated cubes that include targets Tw ($w=1,\ldots, n_t$), or testing cubes, will be denoted by BT_1, BT_2, BT_3, and BT_4 corresponding to backgrounds B_1, B_2, and B_3, respectively. (Notice that both BT_3, and BT_4 have the same background B_3.) Targets are represented by t x t spatial areas consisting of multivariate random samples that are significantly different from their surrounding backgrounds (see figures 2 and 3).

Using the denotations introduced in this subsection, we conjecture that anomaly detection approaches that can maintain a relatively low cutoff threshold using idealized realizations of B_1, B_2, and B_3, as input imagery, will produce better power. For this expectation to be true, desired anomaly detection approaches would have to be less sensitive to local background configurations relative to competing approaches.

4.2 Simulation Plan and Multivariate Cube Structures

In this subsection, we provide a simulation plan and describe the structures of data cubes, which will be used to carry out the simulation experiment described herein. We start by describing the up level simulation plan and then follow it with the details pertained to the constructions of idealized pseudo imagery (data cubes).

4.2.1 Simulation Plan

The following steps were used to conduct the simulation experiments using the data model specified in subsection 4.2.2 through subsection 4.2.3:

i. Since HS or MS imagery can be interpreted as a set of 2-dim images (pixel co-registered) collected at different frequency bands forming a cube, two sets of idealized cubes will be generated: training $\mathbf{B}_{\mathrm{h}}^{(g)}$ and testing $\mathbf{BT}_{z}^{(g)}$ (sometimes a test cube may be referred to as $\mathbf{BT}z^{(g)}$), where $\mathbf{B}_{\mathrm{h}}^{(g)}$ and $\mathbf{BT}_{\mathrm{z}}^{(g)}$ have volume r x c x b, r x c = 256 x 256 represents the total spatial area of a scene and b = 5 represents the total number of spectral bands, $h = 1, \cdots, 3$ indexes different kinds of training background configurations, $z = 1, \cdots, 4$ indexes different kinds of target-background configurations, and $g = 1, \cdots, 1500$ indexes independent repetition.

ii. Training cubes $\mathbf{B}_{\mathrm{h}}^{(g)}$, which is under the null hypothesis, will be constructed by generating independent, identically distributed (i.i.d.) multivariate samples of background classes $\mathbf{C}_{\mathrm{k}}^{(g,i,j)}$, or $\mathbf{B}_{\mathrm{h}}^{(g)} = \left\{ \left\{ \mathbf{C}_{\mathrm{k}}^{(g,i,j)} \right\}_{i=1}^{r} \right\}_{j=1}^{c}$, where (i, j) indexes a spatial (pixel) location in $\mathbf{B}_{\mathrm{h}}^{(g)}$ and $k = 1, \cdots 6$ indexes a particular background class (a distribution, which will be specified later; i.i.d. samples will be generated from single classes, and the samples are independent between the classes). A particular kind of background configuration can be attained by determining where samples of similar or different classes are placed with respect to each other throughout the imagery spatial area, as it will be shown later. In order to mimick a known characteristic of real MS imagery, each sample $\mathbf{C}_{\mathrm{k}}^{(g,i,j)}$ will consist of 5 highly correlated covariates. These covariates are pseudo values generated from multivariate normal distributions, as will be shown later.

Parameter specifications for these distributions are not based on real material spectroscopy, since we do not have meaningful MS imagery and the HS imagery that we have has the number of bands equal to or greater than 120, which is much greater than 5—computational time made us settle for 5 rather than a larger number. Parameter specification will be made to make sure that each one of the 6 classes in the background is significantly different from the other classes.

iii. Cubes $\mathbf{BT}_z^{(g)}$, which is under the alternative hypothesis, will be constructed by first generating independent realizations of $\mathbf{B}_{\mathrm{h}}^{(g)}$, as described in (ii), and then by replacing some (a few) of the background samples $\mathbf{C}_{\mathrm{k}}^{(g,i,j)}$ with i.i.d. multivariate samples of target classes $Tw^{(g,y,x)}$, where $w = 1,\cdots,5$ indexes different target classes (target classes will be also significantly different from any one of the background classes) and (x,y) indexes predetermined spatial locations where theses targets will be placed in $\mathbf{BT}_z^{(g)}$, as it will be shown later. (The term predetermined means that the locations where targets are found are stored so that we can determine later whether a detection is a correct one or a false positive.) Target class samples are independently generated for each location where they replace background class samples. This process is repeated $g = 1,\cdots,1500$ times.

In this simulation experiment, $\mathbf{BT}_1^{(g)}$ will use the background configuration of $\mathbf{B}_1^{(g)}$, $\mathbf{BT}_2^{(g)}$ will use the background configuration of $\mathbf{B}_2^{(g)}$, $\mathbf{BT}_3^{(g)}$, and $\mathbf{BT}_4^{(g)}$ will use the background configuration of $\mathbf{BT}_3^{(g)}$. The number of targets and their locations in the scene background are the same for $\mathbf{BT}_1^{(g)}$ and $\mathbf{B}_2^{(g)}$, but different from the ones in $\mathbf{BT}_3^{(g)}$ and $\mathbf{BT}_4^{(g)}$. Target locations were selected to challenge corresponding anomaly detectors with respect to the complexity of background configuration, target spatial area, and testing window area, as it will be shown later.

iv A detector will produce multiple results testing a cube because it only tests an area at a time, and this area is orders of magnitude smaller than the entire image area of that cube.

A detector trial in this simulation experiment corresponds to comparing samples representing a block of data in the imagery to samples representing data in the outer ring of this block of data. If we denote $\mathbf{X}_1$ as a sequence $1,\cdots,n_1$ of multivariate samples observed in a block of data and $\mathbf{X}_0$ as a sequence $1,\cdots,n_0$ of multivariate samples observed in this block's outer ring, then comparing $\mathbf{X}_1$ to $\mathbf{X}_0$ represents a trial. Samples $\mathbf{X}_1$ and $\mathbf{X}_0$ will be shared by all the detectors used in this simulation experiment.

Making similar comparisons across the entire imagery for overlapping blocks of data produce multiple trial results. From these results a cutoff threshold, type I error or power of the test can be estimated for the given detector, depending on whether it is a training or testing activity.

Attaining cutoff thresholds and estimating type I errors will be done during a training activity using training idealized cubes, and the power of the test will be estimated during

testing activities using testing idealized cubes. Sample sizes n_1 and n_0, which depends on the sizes of the inside window and outside window, respectively, are fixed at once to $n_1 = 81$ (from a 9 x 9 area) and $n_0 = 208$ (from a extension of 4 pixels beyond the 9 x 9 block).

Note that depending on which cube (imagery) is used and where in the imagery the block of data is located, $\mathbf{X}_1$ and/or $\mathbf{X}_0$ may have pure or mixtures of background samples $\mathbf{C}_k$ and/or target samples Tw, see figure 2.

v. Since the industry standard RX detector assumes that $\mathbf{X}_1$ and $\mathbf{X}_0$ are composed of i.i.d. multivariate normal samples of unknown means and unknown but equal covariance matrices, for this simulation we fix at once covariance matrices of different background classes to Σ (see subsection 4.2.2), and fixed at once covariance matrices of different target classes to Ξ (see subsection 4.2.3), where the correlation parameters in this given covariance are all fixed to 1.0.

vi. Attaining Cutoff Thresholds: A training cube will be generated as described in (ii) and tested using a detector as described in (iv). This procedure will produce 57,121 trial results [(r – 17) x (c – 17) = (256 – 17) x (256 – 17)] for each detector. Each set of 57,121 trial results will be used to estimate an empirical PDF, and a cutoff threshold will be attained using the standard quantile method on the estimated PDF. The desired type I error probability (α_t) for the quantile method will vary, where $t = 1, \cdots, 4$ indexes a set of type I error probabilities. In this simulation, we will use $\alpha_1 = 10^{-1}, \alpha_2 = 10^{-2}, \alpha_3 = 10^{-3}, \alpha_4 = 10^{-4}$.

vii. Estimating the Type I Error: An additional training cube will be generated according to (ii) and introduced to each detector, where the detector's corresponding cutoff threshold will be applied to the corresponding detector's output trial results in order to estimate the type I error $\hat{\alpha}_t^{(g)}$ of each detector for a given repetition *g*. The type I error is estimated by counting the number of trials $m_1^{(g)}$ that satisfy the detector's output values being greater than the corresponding detector's cutoff threshold $\varepsilon_t^{(g)}$ and computing the ratio $\hat{\alpha}_t^{(g)} = m_1^{(g)} / m$, where m = 57,121. (Notice that m is sufficiently large for the estimation of the lowest type I error probability $\alpha_4 = 10^{-4}$.)

viii. Estimating the Power (*1.0* – Type II Error): A test cube will be generated according to (iii) and introduced to each detector, where each detector's corresponding cutoff threshold will be applied to the corresponding detector's output results in order to estimate the *power* $\hat{\beta}_t^{(g)}$ for a given repetition *g*. The type II error is estimated by counting the number $m_2^{(g)}$ of trials that satisfy the detector's output values being lower

then the corresponding detector's cutoff threshold $\varepsilon_t^{(g)}$, and satisfy also that the trial location in the test cube corresponds to the location of a target sample (recall that the inserted locations of target samples will be stored). The type II error then can be estimated by computing the ratio $m_2^{(g)} / m_3$, such that power $\hat{\beta}_t^{(g)} = 1.0 - \left(m_2^{(g)} / m_3\right)$, where m_3 is the total number of target samples that is present in the test cube, which will be known for each test cube.

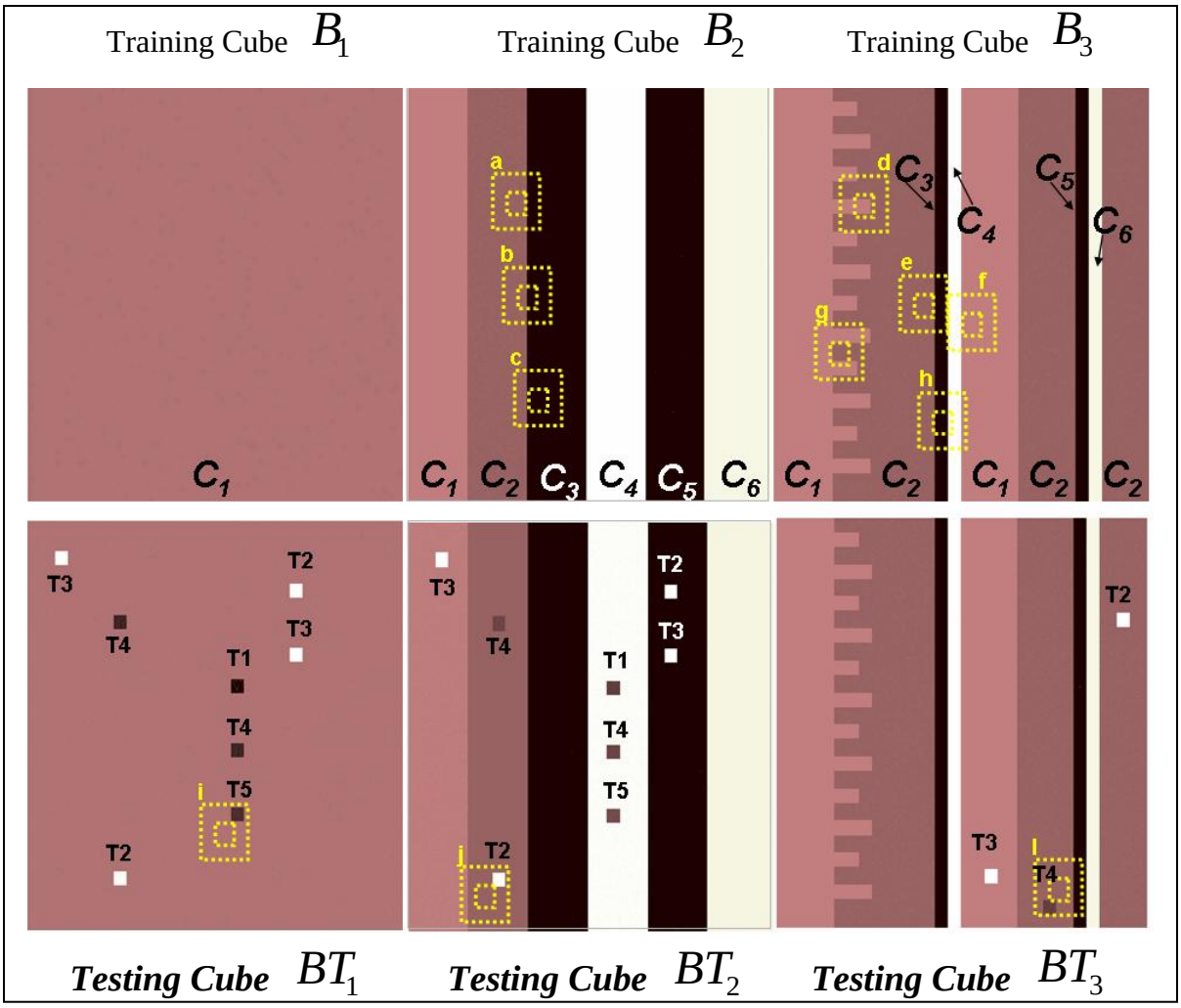

Figure 2. Training cubes B_1, B_2, and B_3, shown as the average of five planes, were used to obtain cutoff thresholds for multiple simulated realizations of cubes BT_1, BT_2, and BT_3, also shown as the average of five planes, for testing. The dotted boxes show approximately the size of a dual rectangular window in proportion to targets' size, background stripes' sizes, and cubes' size. Targets are labeled according to their statistical characteristics—discussed in text. Notice that depending on the window's position in a cube, outside window samples of *1* to *3* classes may be compared to inside window samples of a single class—see, for instance, positions *d* and *h*. Different background classes are denoted by different labels, e.g., C_1.

ix. Estimating Performance Confidence Intervals (CI): In order to check variability using results from (vii) and (viii), confidence intervals for (vii) $\overline{\alpha}_t \pm z_{\nu/2}\sqrt{\frac{1}{R-1}\sum_{g=1}^{R}\left(\hat{\alpha}_t^{(g)} - \overline{\alpha}_t\right)^2}$ and for (viii) $\overline{\beta}_t \pm z_{\nu/2}\sqrt{\frac{1}{R-1}\sum_{g=1}^{R}\left(\hat{\beta}_t^{(g)} - \overline{\beta}_t\right)^2}$ will be estimated, where $z_{\nu/2} = \Phi^{-1}(1-\nu/2)$ is the $1-\nu/2$ quantile $(\nu = 0.05)$ of the standard normal distribution, random quantities simulated will be indexed by *g* for a total number of simulation repetitions *R=1500,* $\overline{\alpha}_t = R^{-1}\sum_{g=1}^{R}\hat{\alpha}_t^{(g)}$, $\overline{\beta}_t = R^{-1}\sum_{g=1}^{R}\hat{\beta}_t^{(g)}$, and $t = 1,\cdots,4$. Tabulated results will be shown for each detector using $\alpha_1 = 10^{-1}, \alpha_2 = 10^{-2}, \alpha_3 = 10^{-3}, \alpha_4 = 10^{-4}$.

4.2.2 Clutter Background Cube Structures (B_1, B_2, and B_3)

This subsection presents the in-depth description of construction of idealized training imagery (pseudo cubes). In addition to parameter specifications, the information contained in this subsection and in follow-on subsections includes computer programming perspective details for estimating power of detectors and other metrics discussed in this subsection.

As discussed in subsection 2.1, a background (training) cube indexed by *h* is denoted by $\mathbf{B}_h^{(g)} = \left\{\left\{\mathbf{C}_k^{(g,i,j)}\right\}_{i=1}^{r}\right\}_{j=1}^{c}$, where $\mathbf{C}_k^{(g,i,j)}$ is a *5-dim* column vector, $k = 1,\cdots,6$ indexes different background classes, (*i,j*) indexes the spatial location in the *r* x *c* area of the cube, and *g* indexes the repetition number. For convenience, the indexes (*i,j*) and *g* will be dropped for the discussions that follow.

Background cube B_1 consists of i.i.d. multivariate samples of a single class and was constructed using highly correlated, normally distributed multivariate samples; a sample C_1 from this class is specified as follows

$$C_1 \sim N(\mu_1, \Sigma), \tag{11}$$

where,

$$\mu_1 = \begin{bmatrix} 630 \\ 640 \\ 720 \\ 660 \\ 650 \end{bmatrix}; \quad \Sigma = \begin{bmatrix} 10.0000 & 14.1421 & 20.0000 & 14.1421 & 10.000 \\ 14.1421 & 20.0000 & 28.2843 & 20.0000 & 14.1421 \\ 20.0000 & 28.2843 & 40.0000 & 28.2843 & 20.0000 \\ 14.1421 & 20.0000 & 28.2843 & 20.0000 & 14.1421 \\ 10.0000 & 14.1421 & 20.0000 & 14.1421 & 10.0000 \end{bmatrix}.$$

B_1 was constructed to form a volume of *256* x *256* x *5* using simulated realizations of C_1. For simplicity, the correlations imbedded in Σ were all equal to *1*, and the variances are of course the diagonal values in Σ.

Background cube B_2 consists of six classes, and multivariate samples $C_1, C_2, C_3, C_4, C_5,$ and C_6 from these classes are specified as follows

$$\begin{aligned} &C_1 \sim N(\mu_1, \Sigma),\ C_2 \sim N(\mu_2, \Sigma),\ C_3 \sim N(\mu_3, \Sigma), \\ &C_4 \sim N(\mu_4, \Sigma),\ C_5 \sim N(\mu_5, \Sigma),\ C_6 \sim N(\mu_6, \Sigma), \end{aligned} \tag{12}$$

where, μ_1 and Σ are defined in equation 11 and

$$\mu_2 = \mu_1 - \begin{bmatrix} 300 \\ 300 \\ 300 \\ 300 \\ 300 \end{bmatrix},\ \mu_3 = \mu_1 - \begin{bmatrix} 820 \\ 780 \\ 820 \\ 780 \\ 820 \end{bmatrix},\ \mu_4 = \mu_1 + \begin{bmatrix} 1400 \\ 1400 \\ 1400 \\ 1400 \\ 1400 \end{bmatrix};\ \mu_5 = \mu_1 - \begin{bmatrix} 800 \\ 800 \\ 800 \\ 800 \\ 800 \end{bmatrix},\ \mu_6 = \mu_3 + \begin{bmatrix} 2000 \\ 2000 \\ 2000 \\ 2000 \\ 2000 \end{bmatrix},$$

or

$$\mu_2 = \begin{bmatrix} 330 \\ 340 \\ 420 \\ 360 \\ 350 \end{bmatrix}, \mu_3 = \begin{bmatrix} -190 \\ -140 \\ -100 \\ -120 \\ -170 \end{bmatrix}, \mu_4 = \begin{bmatrix} 2030 \\ 2040 \\ 2120 \\ 2060 \\ 2050 \end{bmatrix}, \mu_5 = \begin{bmatrix} -170 \\ -160 \\ -80 \\ -140 \\ -150 \end{bmatrix}, \mu_6 = \begin{bmatrix} 1810 \\ 1860 \\ 1900 \\ 1880 \\ 1830 \end{bmatrix}.$$

Background cube B_2 was also constructed to form a total volume of *256* x *256* x *5* using simulated realizations of these six classes, such that, each class covered a sub-volume of *256* x *42* x *5*, as shown in figure 2.

Background cube B_3 was constructed to form a total volume of *256* x *256* x *5* using simulated realizations of the same six classes in B_2, but it displays very different spatial configurations from those in B_2. In reference to figure 2 and denoting *i1* and *i2* as the *start* column and *end* column, respectively, C_1 in B_3 covered columns *i1 = 1* to *i2 = 40*, having horizontal extensions of two sizes (*9* x *18*) and (*9* x *27*) located at rows (*10, 30, 50, 70, 90, 110, 130, 150, 170, 190, 210,* and *230*) for a total of *12* horizontal extensions of C_1, see B_3 in figure 2; C_2 covered columns *i1 = 41* to *i2 = 109*; C_3 covered columns *i1 = 110* to *i2 = 118*; C_4 covered columns *i1 = 119* to *i2 = 127*; C_1, being used again, covered columns *i1 = 128* to *i2 = 166*; C_2, being used again, covered columns *i1 = 129* to *i2 = 205*; C_5 covered columns *i1 = 206* to *i2 = 214*; C_6

covered columns *i1 = 215* to *i2 = 223*; C_2, being used again, covered columns *i1 = 224* to *i2 = 256*.

The column widths of classes $C_3, C_4, C_5,$ and C_6 in B_3 were chosen to match the column width of W_{in} (inside window), see figure 2. The size of W_{in} was arbitrarily chosen to cover a *9* x *9* spatial area, and the size of W_{out} was chosen to cover a *17* x *17* spatial area minus the concentric area of W_{in} within W_{out}. Recall that both windows slide concentrically across the imagery.

4.2.3 Background/Target Cube Structures (BT_1, BT_2, BT_3, and BT_4)

Five different multivariate random variables were specified to represent *targets*, $T1, T2, T3, T4,$ and $T5$; their specifications are as follows:

$$\begin{gathered} T1 \sim N(\tau_1, \Xi), \;\; T2 \sim N(\tau_2, \Xi), \;\; T3 \sim N(\tau_3, \Xi), \\ T4 \sim N(\tau_4, \Xi), \;\; T5 \sim N(\tau_5, \Xi), \end{gathered} \tag{13}$$

where,

$$\tau_1 = \mu_1 - \begin{bmatrix} 620 \\ 580 \\ 620 \\ 580 \\ 620 \end{bmatrix}, \; \tau_2 = \tau_1 + \begin{bmatrix} 2000 \\ 2000 \\ 2000 \\ 2000 \\ 2000 \end{bmatrix}, \; \tau_3 = \tau_1 + \begin{bmatrix} 2050 \\ 2050 \\ 2050 \\ 2050 \\ 2050 \end{bmatrix}, \; \tau_4 = \tau_1 + \begin{bmatrix} 50 \\ 50 \\ 50 \\ 50 \\ 50 \end{bmatrix}, \; \tau_5 = \tau_1 + \begin{bmatrix} 100 \\ 100 \\ 100 \\ 100 \\ 100 \end{bmatrix}$$

or

$$\tau_1 = \begin{bmatrix} 10 \\ 60 \\ 100 \\ 80 \\ 30 \end{bmatrix}, \tau_2 = \begin{bmatrix} 2010 \\ 2060 \\ 2100 \\ 2080 \\ 2030 \end{bmatrix}, \tau_3 = \begin{bmatrix} 2060 \\ 2110 \\ 2150 \\ 2130 \\ 2080 \end{bmatrix}, \tau_4 = \begin{bmatrix} 60 \\ 110 \\ 150 \\ 130 \\ 80 \end{bmatrix}, \tau_5 = \begin{bmatrix} 110 \\ 160 \\ 200 \\ 180 \\ 130 \end{bmatrix},$$

and

$$\Xi = \begin{bmatrix} 100 & 100 & 100 & 100 & 100 \\ 100 & 100 & 100 & 100 & 100 \\ 100 & 100 & 100 & 100 & 100 \\ 100 & 100 & 100 & 100 & 100 \\ 100 & 100 & 100 & 100 & 100 \end{bmatrix}.$$

Targets were constructed to form sub-volumes of constant space size *9* x *9* x *5* using simulated realizations as specified in the third dimension. For simplicity, the correlations imbedded in Ξ were all equal to *1*, and the variances were all equal to *100*.

Cube BT_1 was formed by simulating realizations of B_1 and imbedding (*9* x *9*) boxes of simulated realizations of $T1, T2, T3, T4,$ and $T5$ as shown in figure 2. For convenience, those boxes were labeled using the same denotations of target random variables. In BT_1, the accuracy of targets' spatial locations are unimportant, as long as they are significantly apart from each other (e.g., *35* pixels apart from their centroids, or greater.)

Samples of cube BT_2 were formed by simulating realizations of B_2 and imbedding (*9* x *9*) boxes of target random sample realizations, as shown in figure 2. The targets were spatially collocated sufficiently apart from each other and apart from any one of the transitions of background classes (e.g., greater than *18* pixels between target centroids and a transition). For convenience, the same targets and their locations in BT_1 were used for BT_2, see figure 2.

Samples of BT_3 cubes were formed by simulating realizations of B_3 and imbedding (*9* x *9*) boxes of simulated realizations of $T2, T3,$ and $T4$, as shown in figure 2. The motivation here was to measure performance of anomaly detectors on A challenging background configuration—challenging with respect to class transitions and opportunities to have information from a single class compared to information from two or three classes, (see, for instance, window location *l* in figure 2 [BT_3]). Targets were spatially collocated significantly apart from each other and from class transitions (e.g., greater than *25* pixels apart between their centroids and background transitions, as shown in figure 2.)

Samples of BT_4 cube were formed by simulating realizations of B_3 and adding (*9* x *9*) boxes of simulated realizations of $T1, T2, T3, T4,$ and $T5$, as shown in figure 3. The motivation here was to increase the challenge by putting some targets in transitions of classes, where these targets would be compared to two background classes, and by putting some targets in locations where they would be compared to three background classes (see, for instance, window location *a* in figure 3 [BT_4]) and potential problem areas similar to window location *b* in figure 3 [BT_4]). These targets were collocated significantly apart from each other (i.e., greater than *35* pixels apart from their centroids), and for targets that were embedded in narrow background stripes, they were put perfectly to match the width of those stripes, (see figure 3). Recall that the spatial size of these targets matches the spatial size of W_{in}. There are some challenging areas in BT_4 cubes, for instance, the window location *b* in figure 3 (BT_4) shows a case where inside window samples of a single class will be compared to outside window samples of *4* classes.

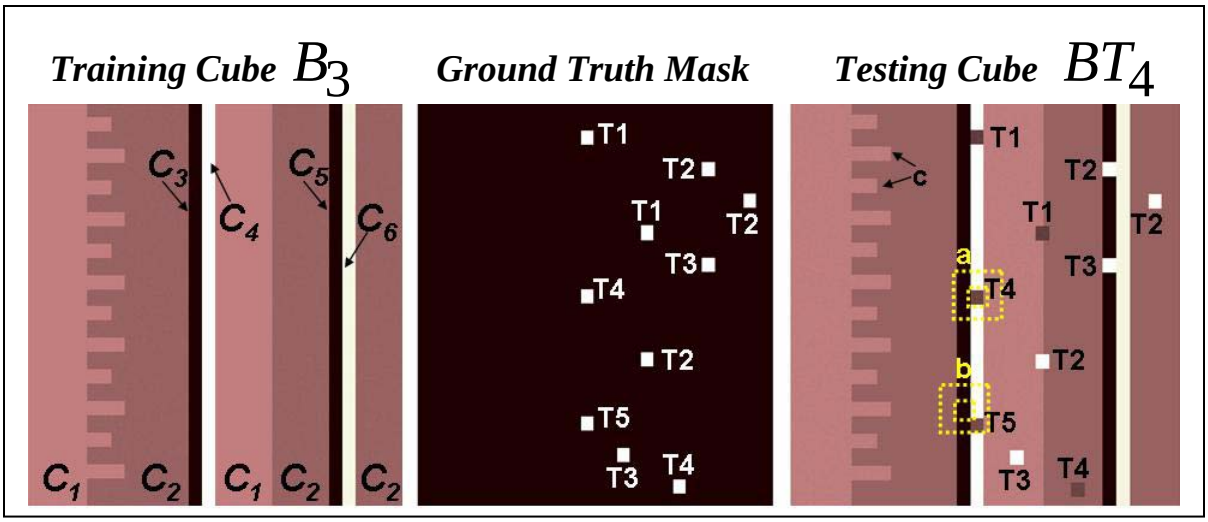

Figure 3. Training cube B_3, shown as the average of five planes, was also used to obtain cutoff thresholds for multiple simulated realizations of testing cube BT_4, also shown as the average of five planes. This testing cube is considered in this simulation as the most challenging target-background configuration for existing anomaly detectors because some of background stripes' sizes correspond to the size of the inside window, and additionally, as shown in the figure, there are horizontal background extensions (see positions *c* in BT_4) of vertical sizes corresponding to inside window's vertical size. The *ground truth mask* is a binary image, where bright square rectangles representing values of *1* validate target locations. Similar *truth masks* are available for the other testing cubes shown in figure 2. Targets labeled differently (e.g., T1 versus T3) have different statistical characteristics.

4.2.4 Type I and Type II Errors

In order to estimate type I and type II errors, a 2 dimensional (2D) *mask* was required to verify the spatial location of targets in the simulated imagery. These masks are binary (i.e., the locations of target pixels are *1* and locations of background pixels are *0*); these masks are often referred to in the target community as *ground truth*. Figure 3 shows the ground truth for BT_4. For this simulation, three ground truth masks were generated, one for BT_1 and BT_2, as they have the same target locations; a second one for BT_3; and a third one for BT_4, which is shown in figure 3.

Thus a detector tests a simulated cube producing a 2D output surface of real values. A detector-corresponding cutoff threshold, which is based on a specified type I error and which is relevant to the cube's background excluding targets, is applied to that surface, such that, pixel values that are above the threshold and which fall within target regions, as validated through a corresponding ground truth mask, are considered a correct target detection; otherwise, they are considered a false detection. These measures can be converted into type I and type II errors by estimating the probability of correct target detection, which is equivalent to *1* minus type II error, and by estimating the probability of false detections, which is equivalent to type I error.

The type II error depends on the sample size, on the detector being used, and on the desired type I error. The sample size depends on the window size used for sampling. Since the existing HS

anomaly detectors were designed to exploit certain multivariate characteristic of HS data, and the ones proposed in the report were designed to operate in an univariate feature domain, we felt that the fairest way to compare these different types of detectors was to test each detector using the same test procedure on the same simulated data, using the same window size, and the same desired type I error.

The nonparametric HS anomaly detectors FLD, EST, and DPC use multivariate information from W_{out} and W_{in} as input to yield a scalor response. The parametric HS anomaly detectors RX uses multivariate information from W_{out} to estimate parameters of a Normal distribution, and test whether the average multivariate information from W_{in} belong to that population, yielding a scalor response. The ANOVA anomaly detector uses as input the same transformed information that AsemiP, AFT, and AVT detectors do.

4.2.5 Obtaining Cutoff Thresholds

A single simulated realization of the three background configurations B_1, B_2, and B_3 were used to obtain cutoff thresholds based on the following set of chosen type I errors α:

$$\begin{aligned} \alpha &= (\alpha_1, \alpha_2, \alpha_3, \alpha_4) \\ &= \left(10^{-1},\ 10^{-2},\ 10^{-3},\ 10^{-4}\right). \end{aligned} \tag{14}$$

(For the purpose of anomaly detection, Type I errors that are higher than 10^{-1} have no practical value). Using fixed sizes W_{out} and W_{in} (i.e., W_{out} yielding $n_0 = 208$ samples and W_{in} yielding $n_1 = 81$ samples), the eight detectors were applied to these simulated realizations, using the data transformation, or no transformation, yielding *3* sets of cutoff thresholds per detector. A corresponding set of thresholds then per detector was obtained for B_1, B_2, and B_3 based on the desired type I errors shown in (14). For a given type I error, one would expect these thresholds to increase depending on the background cube reflecting the level of increasing background complexity among B_1, B_2, and B_3. For illustration, table 13 shows these sets corresponding to detectors RX (multivariate) and AsemiP (univariate).

Table 13. Cutoff thresholds yielded by the multivariate **RX** anomaly detector using *57,121* trial results per simulated background cube $\mathbf{B}_{\mathrm{h}}$ $(h = 1, \cdots, 3)$.

Type I Error	**Background-Only Simulated Cubes**		
α	B_1	B_2	B_3
10^{-1}	2.73586756020100	17.14910687789102	30.75813296377958
10^{-2}	6.69480804231700	29.87601546556457	110.93641256094044
10^{-3}	10.81913096696389	34.61656634779293	164.77893286133680
10^{-4}	16.23242322101338	41.66426583843549	284.76027446709963

In reference to tables 13 and 14, it is worth mentioning that a detector that is less sensitive to these three background cubes would be more desired for the purpose of anomaly detection. This sensitivity can be noticed by observing the rate of increase among the cutoff thresholds for a given type I error among detectors.

Table 14. Cutoff thresholds yielded by the univariate **AsemiP** anomaly detector using *57,121* trial results per simulated background cube $\mathbf{B}_{\mathrm{h}}\ (h = 1, \cdots, 3)$.

Type I Error	**Background-Only Simulated Cubes**		
α	B_1	B_2	B_3
10^{-1}	0.10714023252450	4.04119537549038	5.41993634110348
10^{-2}	0.65071118087721	15.31063919196472	19.20618525917164
10^{-3}	1.66063008313870	15.97474497597062	43.95133931745743
10^{-4}	2.60835114791086	16.31042775997816	44.96548124318737

Sets of thresholds per detector were used to estimate the type I and type II errors as described next.

4.2.6 Estimating Type I and Type II Errors

This subsection presents a computer programming implementation version for estimating type I and type II errors in the context of this simulation experiment.

Type I errors were estimated for each detector using their corresponding sets of cutoff thresholds on their output surfaces after testing each detector on *R* simulated realizations of $BT_1, BT_2, BT_3,$ and BT_4, such that, cutoff thresholds obtained using B_1 were only used to estimate type I errors on B_1; cutoff thresholds obtained using B_2 were only used to estimate type I errors on BT_2; and cutoff thresholds obtained using B_3 were only used to estimate type I errors on BT_3 and BT_4, as these latter target-background cubes shared the same background configuration of B_3.

A *generic* null hypothesis H_0 can be stated for this simulation as follows: At any given location in a simulated cube, samples observed in W_{in} belong to the same class of samples observed in W_{out}. This test will be repeated across the simulated cube generating an output surface for each detector, and the whole process repeated for *M* simulated cubes so that confidence intervals could be estimated for the types I and type II errors.

The specific steps taken to obtain empirical results for the type I and type II errors are shown below for a test cube, in this case BT_4, using a detector.

- Set sizes of the dual rectangular window described in this section to cover an area of 9 x 9—the area of targets—for W_{in} and to cover an area of 17 x 17 minus the concentric area of W_{in} for W_{out}.
- Reduce the area of the corresponding ground truth mask (see figure 3)—this process is known as downsampling—from 256 x 256 to (256–17) x (256–17) to coincide with the detector's output surface size. Denote downsampled mask as TRUTH.
- For g = 1 to R = 1500 (R, maximum number of repetitions)
  - Generate a simulated realization of BT_4, as described earlier
  - For i = 1 to 256
    - For j = 1 to 256
      - Using the dual rectangular window's left-upper corner at (i,j), test the local region using the detector according to its input requirements
      - This process will generate an output surface OUTPUT
  - Intermediate Result: OUTPUT of size (256 – 17) x (256 – 17)
  - For t = 1 to 4 (the maximum number of cutoff thresholds per detector)
    - Let $\varepsilon = (\varepsilon_1, \cdots, \varepsilon_4)$ be the detector's cutoff thresholds corresponding to the desired type I errors shown in (4), where ε_t corresponds to the desired type I error α_t. Apply ε_t to OUTPUT
    - Let $v(k,w)$ represent the value of a pixel located at (k,w) in OUTPUT, Θ the set of target pixel locations in TRUTH, and Θ^c the mutually exclusive set representing background pixel locations in TRUTH; notice that the set of all pixel locations in TRUTH is $\Theta \bigcup \Theta^c$, where $\bigcup$ is the union of sets. Denote N_{Θ^c} the total number of locations in Θ^c.
    - A **type I error** is committed when $v(k,w) > \varepsilon_t$ and $(k,w) \in \Theta^c$, where $\in$ denotes *belongs to*. Add all instances when the type I error was committed, denoting this sum N_{fa}.
      - Estimate type I error for trial g and for α_t, denote this estimate

$$\hat{\alpha}_t^{(g)} = \frac{N_{fa}}{N_{\Theta^c}} \tag{15}$$

- A **type II error** is committed when $v(k,w) < \varepsilon_t$ and $(k,w) \in \Theta$. Add all instances when the type II error was committed, such that, multiple instances of the same target will be counted **only once** to avoid redundancies. (For instance, if among ten targets, one of them did not have at least a portion of its spatial area detected, the estimated type II error would be 1 divided by 10, or 0.1; on the other hand, if a target yields a response resembling a relatively wide peak, this target would be counted as a single detection, as long as a portion of the peak's footprint coincides with the target's expected spatial location. This procedure is widely practiced in the target community because targets often produce adjacent artifact responses owing to their presences in clutter backgrounds. Output surfaces will be shown later to clarify this point.)
- Denote the sum of targets that committed type II errors as N_{miss}, and the total number of individual targets in TRUTH as N_{total} (for BT_4, $N_{total} = 10$, see figure 3); notice that $N_{miss} \leq N_{total}$, where $\leq$ is *less or equal to*
    - Estimate the *power* for trial *g* and for α_t, denote it $\hat{\beta}_t^{(g)}$, which represents *1.0* minus the estimated type II error, or

$$\hat{\beta}_t^{(g)} = 1.0 - \frac{N_{miss}}{N_{total}} \tag{16}$$

- Intermediate Results for a single trial *g*:

$$\hat{\alpha}^{(g)} = \left(\hat{\alpha}_1^{(g)}, \hat{\alpha}_2^{(g)}, \hat{\alpha}_3^{(g)}, \hat{\alpha}_4^{(g)}\right)$$
$$\hat{\beta}^{(g)} = \left(\hat{\beta}_1^{(g)}, \hat{\beta}_2^{(g)}, \hat{\beta}_3^{(g)}, \hat{\beta}_4^{(g)}\right)$$

(17)

- Intermediate Results for *R = 1,500* trials:

$$\hat{\alpha} = \left(\hat{\alpha}^{(1)}, \hat{\alpha}^{(2)}, \cdots, \hat{\alpha}^{(1500)}\right)$$
$$\hat{\beta} = \left(\hat{\beta}^{(1)}, \hat{\beta}^{(2)}, \cdots, \hat{\beta}^{(1500)}\right) \tag{18}$$

- For *t* = 1 to 4 (total number of desired type I error values)
    - Estimate the means using results from 1,500 trials for the t^{th} cutoff threshold:

$$\hat{\mu}_t^{(0)} = \frac{1}{1500}\sum_{g=1}^{1500} \hat{\alpha}_t^{(g)}$$
$$\hat{\mu}_t^{(1)} = \frac{1}{1500}\sum_{g=1}^{1500} \hat{\beta}_t^{(g)} \tag{19}$$

- Estimate variances using results from 1,500 repetitions for the t^{th} cutoff threshold:

$$\hat{\sigma}_t^{(0)} = \frac{1}{1500}\sum_{g=1}^{1500}\left(\hat{\alpha}_t^{(g)} - \hat{\mu}_t^{(0)}\right)^2$$
$$\hat{\sigma}_t^{(1)} = \frac{1}{1500}\sum_{g=1}^{1500}\left(\hat{\beta}_t^{(g)} - \hat{\mu}_t^{(1)}\right)^2 \quad (20)$$

- Estimate 95% confidence interval using results from the 1,500 trials for the t^{th} cutoff threshold:

$$\left(\hat{\mu}_t^{(0)} - 1.96\sqrt{\frac{\hat{\sigma}_t^{(0)}}{1500}},\ \ \hat{\mu}_t^{(0)} + 1.96\sqrt{\frac{\hat{\sigma}_t^{(0)}}{1500}}\right)$$
$$\left(\hat{\mu}_t^{(1)} - 1.96\sqrt{\frac{\hat{\sigma}_t^{(1)}}{1500}},\ \ \hat{\mu}_t^{(1)} + 1.96\sqrt{\frac{\hat{\sigma}_t^{(1)}}{1500}}\right) \quad (21)$$

- Final Results: 95% confidence intervals for the power of the test, i.e.,
  - Type I Error Confidence Intervals

$$\left(\hat{\mu}_1^{(0)} - 1.96\sqrt{\frac{\hat{\sigma}_1^{(0)}}{1500}},\ \ \hat{\mu}_1^{(0)} + 1.96\sqrt{\frac{\hat{\sigma}_1^{(0)}}{1500}}\right)$$
$$\left(\hat{\mu}_2^{(0)} - 1.96\sqrt{\frac{\hat{\sigma}_2^{(0)}}{1500}},\ \ \hat{\mu}_2^{(0)} + 1.96\sqrt{\frac{\hat{\sigma}_2^{(0)}}{1500}}\right) \quad (22)$$
$$\vdots$$
$$\left(\hat{\mu}_4^{(0)} - 1.96\sqrt{\frac{\hat{\sigma}_4^{(0)}}{1500}},\ \ \hat{\mu}_4^{(0)} + 1.96\sqrt{\frac{\hat{\sigma}_4^{(0)}}{1500}}\right)$$

  - (1.0 – Type II Error) Confidence Intervals

$$\begin{gathered}\left(\hat{\mu}_1^{(1)}-1.96\sqrt{\frac{\hat{\sigma}_1^{(1)}}{1500}},\ \hat{\mu}_1^{(1)}+1.96\sqrt{\frac{\hat{\sigma}_1^{(1)}}{1500}}\right)\\ \left(\hat{\mu}_2^{(1)}-1.96\sqrt{\frac{\hat{\sigma}_2^{(1)}}{1500}},\ \hat{\mu}_2^{(1)}+1.96\sqrt{\frac{\hat{\sigma}_2^{(1)}}{1500}}\right)\\ \vdots\\ \left(\hat{\mu}_4^{(1)}-1.96\sqrt{\frac{\hat{\sigma}_4^{(1)}}{1500}},\ \hat{\mu}_4^{(1)}+1.96\sqrt{\frac{\hat{\sigma}_4^{(1)}}{1500}}\right)\end{gathered} \tag{23}$$

The results shown in equations 22 and 23 were computed for each detector as it tested simulated data cubes. Their results are shown next.

4.3 Simulation Results

This subsection presents the 95% confidence intervals that were computed for each detector as they tested cubes BT_1, BT_2, BT_3, and BT_4. The tables are organized such that the first column shows the detector's name, followed by the desired type I error α_t $(t=1,\cdots,4)$ using (14), followed by the type I error's 95% confidence intervals using equation 22, followed by the power's 95% confidence intervals using (23)). This organization was applied to tables 15 through 22. Table 15 shows the simulation results of competing detectors testing *1,500* simulated realizations of BT_1, targets on a *easy* background configuration. The computation of individual power and type I error estimates used 57,121 trial results per repetition.

From tables 15 and 16, it is quite evident that in many of these results the desired type I errors α_t $(t=1,\cdots,4)$ do not fall within their corresponding confidence intervals. The reason is that since the cutoff thresholds were obtained from a background-only simulated cube (in this case B_1), it produced lower threshold values compared to detectors' artifact responses that can be relatively high owing to the presence of targets in a test cube (in this case BT_1). These artifacts occur when homogenous samples in the inside window are compared to a mixture of target-background samples in the outside window—see, for instance, window location *i* in figure 2 (BT_1), and examples of 3D surfaces in figure 4.

Table 16 shows the simulation results of our detectors testing *1,500* simulated realizations of BT_1, where estimations of power and type I error per simulation repetition used *57,121* trial results.

Table 15. Multivariate detectors' type I error and power performances using *57,121* trial results per simulated background cube $B_1^{(g)}$, where *g* indexes repetitions $\{B_1^{(g)}\}_{g=1}^{1500}$, and *57,121* trial results per target-background cube $BT_1^{(g)}$, $\{B_1^{(g)}\}_{g=1}^{1500}$.

Detectors	α_t	**Single Homogeneous Background Region Plus *8* Targets Easy background configuration, targets in homogeneous areas**			
		Type I Error 95% Confidence Interval		**(1.0 – Type II Error) 95% Confidence Interval**	
		Lower Bound	**Upper Bound**	**Lower Bound**	**Upper Bound**
RX[1,2]	10^{-1}	0.155322	0.158816	1.000000	1.000000
	10^{-2}	0.038951	0.039978	1.000000	1.000000
	10^{-3}	0.012227	0.012741	1.000000	1.000000
	10^{-4}	0.002292	0.002528	1.000000	1.000000
FLD[2]	10^{-1}	0.170777	0.173506	1.000000	1.000000
	10^{-2}	0.087272	0.087644	1.000000	1.000000
	10^{-3}	0.078813	0.078926	1.000000	1.000000
	10^{-4}	0.077659	0.077706	1.000000	1.000000
EST[2]	10^{-1}	0.170015	0.172612	1.000000	1.000000
	10^{-2}	0.077576	0.078367	1.000000	1.000000
	10^{-3}	0.069210	0.069504	1.000000	1.000000
	10^{-4}	0.067697	0.067767	1.000000	1.000000
DPC[2]	10^{-1}	0.171855	0.175635	1.000000	1.000000
	10^{-2}	0.087120	0.088003	1.000000	1.000000
	10^{-3}	0.078594	0.078854	1.000000	1.000000
	10^{-4}	0.077440	0.077517	1.000000	1.000000

Table 16. Univariate detectors' type I error and power performances using *57,121* trial results per simulated background cube $B_1^{(g)}$, where g indexes repetitions $\left\{B_1^{(g)}\right\}_{g=1}^{1500}$, and *57,121* trial results per target-background cube $BT_1^{(g)}$, $\left\{BT_1^{(g)}\right\}_{g=1}^{1500}$.

Detectors	α_t	**Single Homogeneous Background Region Plus *8* Targets Easy background configuration, targets in homogeneous areas**			
		Type I Error 95% Confidence Interval		**(1.0 – Type II Error) 95% Confidence Interval**	
		Lower Bound	**Upper Bound**	**Lower Bound**	**Upper Bound**
AsemiP	10^{-1}	0.171441	0.174740	1.000000	1.000000
	10^{-2}	0.072887	0.073796	1.000000	1.000000
	10^{-3}	0.048219	0.048483	1.000000	1.000000
	10^{-4}	0.037595	0.037807	1.000000	1.000000
AFT	10^{-1}	0.022617	0.022708	1.000000	1.000000
	10^{-2}	0.001783	0.001813	1.000000	1.000000
	10^{-3}	0.000579	0.000589	1.000000	1.000000
	10^{-4}	0.000579	0.000589	1.000000	1.000000
AVT	10^{-}	0.156213	0.159702	1.000000	1.000000
	10^{-2}	0.055914	0.056730	1.000000	1.000000
	10^{-3}	0.039220	0.039370	1.000000	1.000000
	10^{-4}	0.039220	0.039370	1.000000	1.000000
ANOVA	10^{-1}	0.121679	0.125269	1.000000	1.000000
	10^{-2}	0.010971	0.011928	1.000000	1.000000
	10^{-3}	0.001161	0.001412	1.000000	1.000000
	10^{-4}	0.000086	0.000147	1.000000	1.000000

Relatively high artifact responses can be observed in figure 4, where 3D views of some of the detectors' output surfaces are exhibited. (Those sample surfaces are examples of the intermediate result denoted as OUTPUT in the steps discussed earlier for estimating type I and type II errors.)

A final observation from results shown in table 16 is the repetition of lower and upper bounds for some of the detectors' confidence intervals at different α_t. This repetition happens because of the resolution's discrete nature in histogram generation. Our detectors often produce discretely isolated relatively high responses, which sometimes cannot be resolved for relatively lower type I errors (e.g., $\alpha_t = 10^{-3}$). So, when this resolution limitation occurs at some type I error, lower type I errors will produce the same confidence interval estimate.

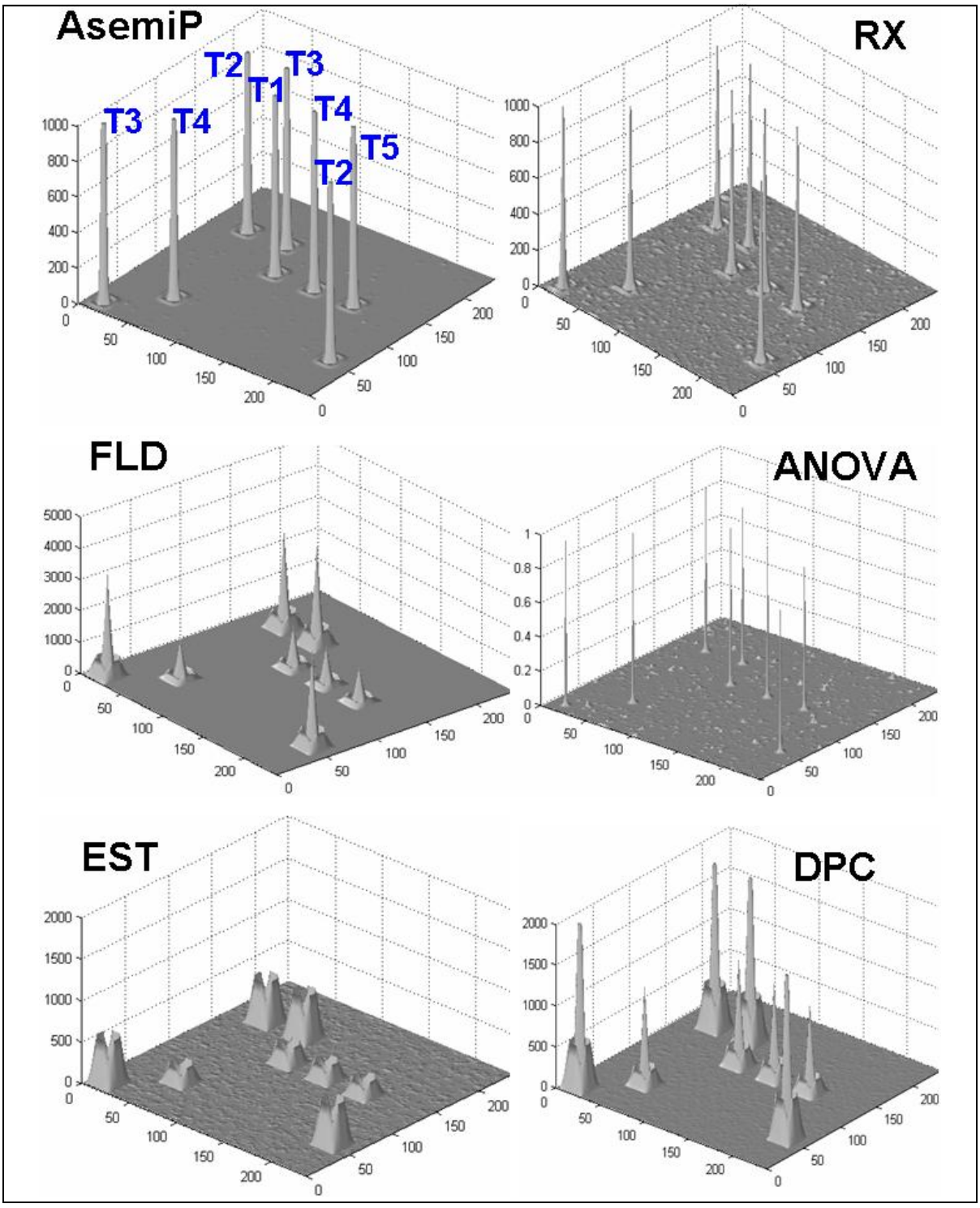

Figure 4. Examples of intermediate result OUTPUT, as described in text. The peaks are responses from the eight targets as *seen* by the different detectors testing a single simulated realization of BT_1. Notice the artifact responses in the vicinity of these peaks. Those artifacts contribute to the type I error, thus, increasing its estimate in respect to the desired type I error.

Now that we have calibrated detectors' performances in a *benign* target-background configuration, let us take a look at their performances in a more challenging background configuration—a moderate target-background configuration. Table 17 shows the simulation results using existing detectors to test *1,500* simulated realizations of BT_2—a moderate target-background configuration.

Results in table 17 already show some signs of performance losses (higher type II error) for three of the algorithms, they are: EST, DPC, and FLD. The fundamental cause for these losses is the detectors' inability to handle transitions of regions. Recall that the cutoff thresholds used to test BT_2 were obtained using a moderate background configuration B_2, which is a six-class background-only cube. As seen in figure 5, the transitions among these regions yield relatively high responses using these detectors, which in turn yield relatively high cutoff thresholds to approximate the type I error requested. Figure 5 depicts, for instance, target examples that can respond comfortably above some of these detectors' responses on transitions of background regions, and it also shows target examples that cannot respond as high. The EST detector suffered the worst performance loss testing BT_2. Figure 6 shows two 3D views of the same output surface produced by the EST detector testing a single simulated realization of BT_2. Notice that the same eight targets that could be detected in BT_1 (see figure 3, figure 4 and table 15), responded below the five cutoff thresholds corresponding to the desired type I errors in (14) for this detector on BT_2, producing *zero* target detection—or equivalently a type II error of *unity*, as shown in table 17. The FLD and DPC detectors missed *1* out of *8* targets. Although the detectors FLD, EST, and DPC already show signs of performance losses (higher type II errors) in table 17—in contrast to their performances in table 15, the industry standard RX detector, as well as the ANOVA detector, handled well this moderately challenging target-background configuration.

Table 17. Multivariate detectors' type I error and power performances using *57,121* trial results per simulated background cube $B_2^{(g)}$, where *g* indexes repetitions $\left\{B_2^{(g)}\right\}_{g=1}^{1500}$, and *57,121* trial results per target-background cube $BT_2^{(g)}$, $\left\{BT_2^{(g)}\right\}_{g=1}^{1500}$.

Detectors	α_t	**Six Homogeneous Background Regions Plus *8* Targets Moderate background configuration, targets in homogeneous areas**			
		Type I Error 95% Confidence Interval		**(1.0 – Type II Error) 95% Confidence Interval**	
		Lower Bound	**Upper Bound**	**Lower Bound**	**Upper Bound**
RX	10^{-1}	0.105908	0.106425	1.000000	1.000000
	10^{-2}	0.011534	0.011852	1.000000	1.000000
	10^{-3}	0.001980	0.002107	1.000000	1.000000
	10^{-4}	0.000258	0.000346	1.000000	1.000000
FLD	10^{-1}	0.107706	0.107727	0.875000	0.875000
	10^{-2}	0.005343	0.005670	0.875000	0.875000
	10^{-3}	0.001772	0.001783	0.875000	0.875000
	10^{-4}	0.001767	0.001768	0.875000	0.875000
EST	10^{-1}	0.106490	0.106511	0.000000	0.000000
	10^{-2}	0.005344	0.005670	0.000000	0.000000
	10^{-3}	0.001772	0.001783	0.000000	0.000000
	10^{-4}	0.001768	0.001768	0.000000	0.000000
DPC	10^{-1}	0.107706	0.107727	0.875000	0.875000
	10^{-2}	0.005344	0.005670	0.875000	0.875000
	10^{-3}	0.001772	0.001783	0.875000	0.875000
	10^{-4}	0.001768	0.001768	0.875000	0.875000

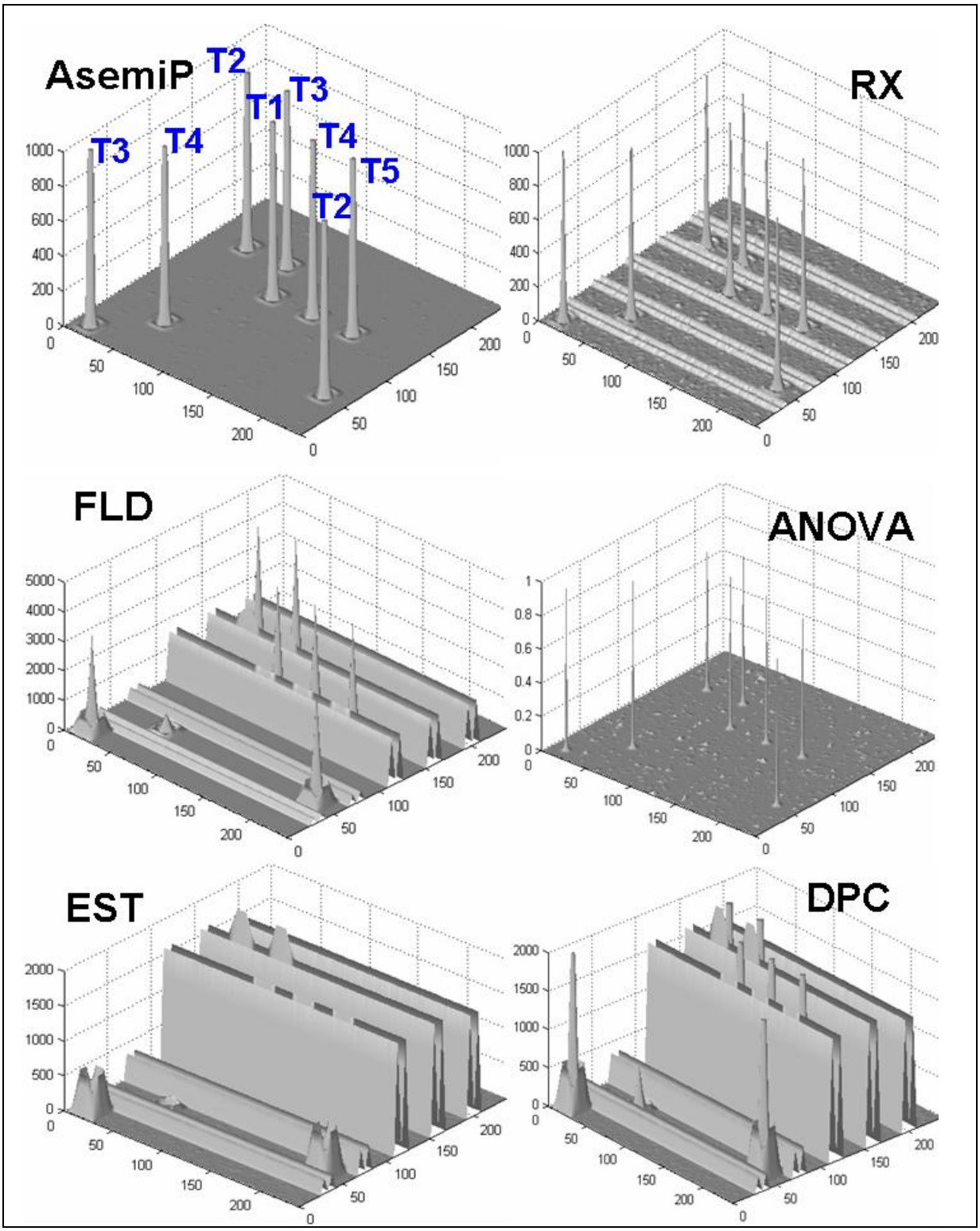

Figure 5. Examples of intermediate result OUTPUT, as described in text. The peaks are responses from the eight targets as *seen* by the different detectors testing a single simulated realization of BT_2. Notice, in some of these surfaces, *M* shaped row responses owing to transitions of different background classes—see, for instance, window locations *a*, *b*, and *c* in figure 2 (B_2). Location *a* yields a local peak to the left of *b*, *b* yields a local valley, and *c* yields a local peak to the right of *b*.

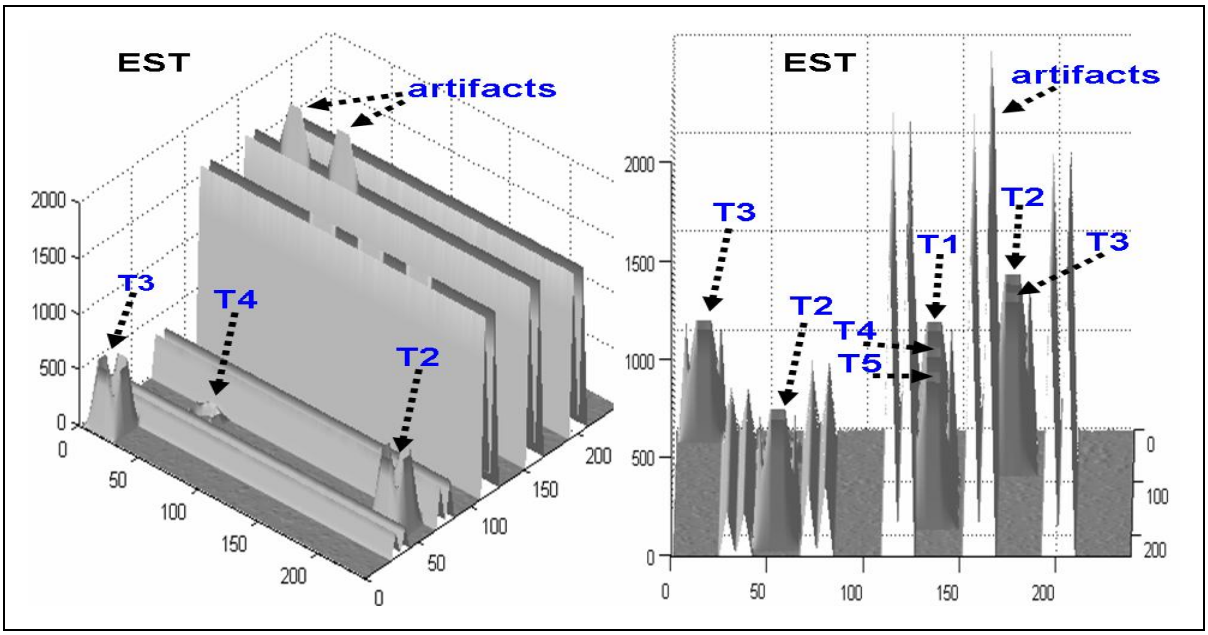

Figure 6. An intermediate result OUTPUT, EST detector testing BT_2. Both surfaces are the same, but shown at different viewing perspectives. Some of the artifacts shown for the view at the right hand side were due to the presence of targets T2 and T3 in the outside window W_{out} —these two targets are shown immediately to the right of these artifacts. A similar case is shown for window location *j* in figure 2 (BT_2). The responses of all *8* targets using the EST detector were below the cutoff thresholds corresponding to this detector for this background configuration.

Table 18 shows the simulation results using the univariate detectors to test *1,500* simulated realizations of BT_2.

Table 18. Univariate detectors' type I error and power performances using *57,121* trial results per simulated background cube $B_2^{(g)}$, where *g* indexes repetitions $\left\{B_2^{(g)}\right\}_{g=1}^{1500}$, and *57,121* trial results per target-background cube $BT_2^{(g)}$, $\left\{BT_2^{(g)}\right\}_{g=1}^{1500}$.

Detectors	α_t	Six Homogeneous Background Regions Plus 8 Targets Moderate background configuration, targets in homogeneous areas			
		Type I Error 95% Confidence Interval		(1.0 – Type II Error) 95% Confidence Interval	
		Lower Bound	Upper Bound	Lower Bound	Upper Bound
AsemiP	10^{-1}	0.010186	0.010240	1.000000	1.000000
	10^{-2}	0.002429	0.002429	1.000000	1.000000
	10^{-3}	0.002429	0.002429	1.000000	1.000000
	10^{-4}	0.002429	0.002429	1.000000	1.000000
AFT	10^{-1}	0.010248	0.010305	1.000000	1.000000
	10^{-2}	0.002450	0.002473	1.000000	1.000000
	10^{-3}	0.002450	0.002473	1.000000	1.000000
	10^{-4}	0.002449	0.002470	1.000000	1.000000
AVT	10^{-1}	0.135670	0.198334	1.000000	1.000000
	10^{-2}	0.008885	0.016029	1.000000	1.000000
	10^{-3}	0.001067	0.001435	1.000000	1.000000
	10^{-4}	0.000236	0.000259	1.000000	1.000000
ANOVA	10^{-1}	0.117101	0.119892	1.000000	1.000000
	10^{-2}	0.011954	0.012952	1.000000	1.000000
	10^{-3}	0.001187	0.001442	1.000000	1.000000
	10^{-4}	0.000154	0.000230	1.000000	1.000000

Some of the confidence intervals shown in tables 17 and 18 do not include the values of α_t, the reason for these apparent discrepancies were explained in the text discussion for tables 15 and 16.

Detectors AsemiP, AFT, and AVT handled well this moderately challenging target-background configuration, see table 18 and examples of output surfaces in figure 5. The output surfaces corresponding to AsemiP and ANOVA detectors in figure 5 depict how insensitive these detectors are to transitions of regions. The same holds true for detectors AFT and AVT, although their surfaces are not shown in figure 5.

Table 19 shows the simulation results of competing detectors testing *1,500* simulated realizations of BT_3, a difficult background configuration having targets in homogenous areas, see figure 2 (BT_3).

Table 19. Multivariate detectors' type I error and power performances using *57,121* trial results per simulated background cube $B_3^{(g)}$, where *g* indexes repetitions $\left\{B_3^{(g)}\right\}_{g=1}^{1500}$, and *57,121* trial results per target-background cube $BT_3^{(g)}$, $\left\{BT_3^{(g)}\right\}_{g=1}^{1500}$.

Detectors	α_t	**Nine Homogeneous Background Regions Plus *3* Targets Difficult background configuration, targets in homogeneous areas**			
		Type I Error 95% Confidence Interval		**(1.0 – Type II Error) 95% Confidence Interval**	
		Lower Bound	**Upper Bound**	**Lower Bound**	**Upper Bound**
RX	10^{-1}	0.099516	0.099962	1.000000	1.000000
	10^{-2}	0.010100	0.010353	1.000000	1.000000
	10^{-3}	0.001470	0.001488	1.000000	1.000000
	10^{-4}	0.000009	0.000032	1.000000	1.000000
FLD	10^{-1}	0.098667	0.098757	0.666667	0.666667
	10^{-2}	0.010641	0.010773	0.666667	0.666667
	10^{-3}	0.001663	0.001768	0.666667	0.666667
	10^{-4}	0.000918	0.000954	0.666667	0.666667
EST	10^{-1}	0.095054	0.095144	0.000000	0.000000
	10^{-2}	0.010278	0.010411	0.000000	0.000000
	10^{-3}	0.001663	0.001767	0.000000	0.000000
	10^{--}	0.000918	0.000954	0.000000	0.000000
DPC	10^{-1}	0.098667	0.098757	0.666667	0.666667
	10^{-2}	0.010641	0.010773	0.666667	0.666667
	10^{-3}	0.001663	0.001767	0.666667	0.666667
	10^{-4}	0.000918	0.000954	0.666667	0.666667

As mentioned immediately after table 17, for a different set of targets and background configuration, table 19 shows that the EST detector using its corresponding cutoff thresholds missed all *3* targets in this difficult background configuration. Similarly, using their corresponding cutoff thresholds for this difficult background configuration, table 19 shows that detectors FLD and DPC missed *1* out of *3* targets. On the other hand, detectors RX and ANOVA handled well this relatively difficult target-background configuration.

Table 20 shows the simulation results of our detectors testing *1,500* simulated realizations of BT_3.

Table 20. Univariate detectors' type I error and power performances using *57,121* trial results per simulated background cube $B_3^{(g)}$, where *g* indexes repetitions $\left\{B_3^{(g)}\right\}_{g=1}^{1500}$, and *57,121* trial results per target-background cube $BT_3^{(g)}$, $\left\{BT_3^{(g)}\right\}_{g=1}^{1500}$.

Detectors	α_t	**Nine Homogeneous Background Regions Plus *3* Targets Difficult background configuration, targets in homogeneous areas**			
		Type I Error 95% Confidence Interval		**(1.0 – Type II Error) 95% Confidence Interval**	
		Lower Bound	**Upper Bound**	**Lower Bound**	**Upper Bound**
AsemiP	10^{-1}	0.107317	0.107851	1.000000	1.000000
	10^{-2}	0.010891	0.011159	1.000000	1.000000
	10^{-3}	0.001047	0.001141	1.000000	1.000000
	10^{-4}	0.000214	0.000230	1.000000	1.000000
AFT	10^{-1}	0.107367	0.107903	1.000000	1.000000
	10^{-2}	0.010893	0.011158	1.000000	1.000000
	10^{-3}	0.001047	0.001142	1.000000	1.000000
	10^{-4}	0.000216	0.000234	1.000000	1.000000
AVT	10^{-1}	0.100379	0.101131	1.000000	1.000000
	10^{-2}	0.009416	0.010360	1.000000	1.000000
	10^{-3}	0.001003	0.001247	1.000000	1.000000
	10^{-4}	0.000235	0.000282	1.000000	1.000000
ANOVA	10^{-1}	0.098873	0.102113	1.000000	1.000000
	10^{-2}	0.008414	0.009312	1.000000	1.000000
	10^{-3}	0.000853	0.001049	1.000000	1.000000
	10^{-4}	0.000069	0.000105	1.000000	1.000000

Table 20 shows that univariate detectors can also maintain a relatively low type I error and detect all *3* targets. As explained earlier, adjacent locations to targets often yield high responses, which will contribute to the type I error. These contributions explain the reason why some values of α_t do not fall within estimated confidence intervals—see, for instance, estimated confidence intervals for $\alpha_5 = 10^{-5}$ in table 20.

The gradual increase of target-background configuration difficulty from BT_1 to BT_3 has shown some performance losses and some performance comparability between multivariate and univariate detectors (e.g., detectors RX and ANOVA have performed comparably with detectors AsemiP, AFT, and AVT.) We present next the detectors' performances on cubes representing the most difficult target-background configuration in this simulation—BT_4.

Table 21 shows the simulation results of competing detectors testing *1,500* simulated realizations of BT_4.

Table 21. Multivariate detectors' type I error and power performances using *57,121* trial results per simulated background cube $B_3^{(g)}$, where *g* indexes repetitions $\left\{B_3^{(g)}\right\}_{g=1}^{1500}$, and *57,121* trial results per target-background cube $BT_4^{(g)}$, $\left\{BT_4^{(g)}\right\}_{g=1}^{1500}$.

Detectors	α_t	**Nine Homogeneous Background Regions Plus *10* Targets Difficult background configuration, *7* targets in transition areas**			
		Type I Error 95% Confidence Interval		**(1.0 – Type II Error) 95% Confidence Interval**	
		Lower Bound	**Upper Bound**	**Lower Bound**	**Upper Bound**
RX	10^{-1}	0.101381	0.101805	1.000000	1.000000
	10^{-2}	0.009608	0.009831	1.000000	1.000000
	10^{-3}	0.000851	0.000861	0.700000	0.700000
	10^{-4}	0.000921	0.000923	0.500000	0.500000
FLD	10^{-1}	0.101444	0.101535	0.666667	0.666667
	10^{-2}	0.010374	0.010522	0.500000	0.500000
	10^{-3}	0.001120	0.001279	0.500000	0.500000
	10^{-4}	0.000072	0.000112	0.500000	0.500000
EST	10^{-1}	0.101303	0.101394	0.700000	0.700000
	10^{-2}	0.010374	0.010522	0.300000	0.300000
	10^{-3}	0.001120	0.001279	0.300000	0.300000
	10^{-4}	0.000072	0.000112	0.300000	0.300000
DPC	10^{-1}	0.101444	0.101535	0.666667	0.666667
	10^{-2}	0.010374	0.010522	0.500000	0.500000
	10^{-3}	0.001120	0.001279	0.500000	0.500000
	10^{-4}	0.000072	0.000112	0.500000	0.500000

The results shown in table 21 for *10* targets included the same *3* targets used to obtain results shown in table 19. Table 21 shows noticeable target detection degradations compared to previous tables, with an apparent exception—the EST detector's performance between table 19 and table 21. The *3* targets that were undetected by the EST detector as shown in table 19 were again missed by this detector as shown in table 21 (see row for EST, $\alpha_1 = 10^{-1}$). The EST detector missed those *3* out of *10* targets using its correspondent cutoff threshold for α_1, but—ironically—it could detect other targets found in more difficult locations, see locations characterized by transitions of regions in figure 3 (window locations *a* and *b* in BT_4). This irony is what motivates some users to use the EST detector in real HS data.

As mentioned earlier, anomaly detectors are known for producing relatively high responses adjacent to target locations because, at those adjacent locations, samples of homogeneous backgrounds in the inside window W_{in} are compared to a mixture of samples in the outside

window W_{out}—this mixture may consist of samples of targets, samples of the same background in W_{in} and from other backgrounds (see, for instance, window position *b* in figure 3 [BT_4].)

Table 22 shows the simulation results using proposed detectors to test *1,500* simulated realizations of BT_4.

Table 22. Univariate detectors' type I error and power performances using *57,121* trial results per simulated background cube $B_3^{(g)}$, where *g* indexes repetitions $\left\{B_3^{(g)}\right\}_{g=1}^{1500}$, and *57,121* trial results per target-background cube $BT_4^{(g)}$, $\left\{BT_4^{(g)}\right\}_{g=1}^{1500}$.

Detectors	α_t	**Nine Homogeneous Background Regions Plus *10* Targets Difficult background configuration, *7* targets in transition areas**			
		Type I Error 95% Confidence Interval		**(1.0 – Type II Error) 95% Confidence Interval**	
		Lower Bound	**Upper Bound**	**Lower Bound**	**Upper Bound**
AsemiP	10^{-1}	0.111715	0.112103	1.000000	1.000000
	10^{-2}	0.011173	0.011399	1.000000	1.000000
	10^{-3}	0.001400	0.001496	1.000000	1.000000
	10^{-4}	0.000802	0.000817	1.000000	1.000000
AFT	10^{-1}	0.111758	0.112148	1.000000	1.000000
	10^{-2}	0.011171	0.011395	1.000000	1.000000
	10^{-3}	0.001402	0.001499	1.000000	1.000000
	10^{-4}	0.000805	0.000823	1.000000	1.000000
AVT	10^{-1}	0.103537	0.104030	1.000000	1.000000
	10^{-2}	0.026165	0.026788	1.000000	1.000000
	10^{-3}	0.019580	0.019728	1.000000	1.000000
	10^{-4}	0.018893	0.018938	1.000000	1.000000
ANOVA	10^{-1}	0.100254	0.103467	1.000000	1.000000
	10^{-2}	0.009011	0.009827	0.500000	0.500000
	10^{-3}	0.000978	0.001151	0.500000	0.500000
	10^{-4}	0.000077	0.000107	0.500000	0.500000

Despite the difficult target-background configuration in BT_4, table 22 shows that the univariate detectors can maintain a relatively low type I error while detecting all *10* targets in this simulation. To better appreciate the difference in performance among different detectors, see examples output surfaces shown in figure 7.

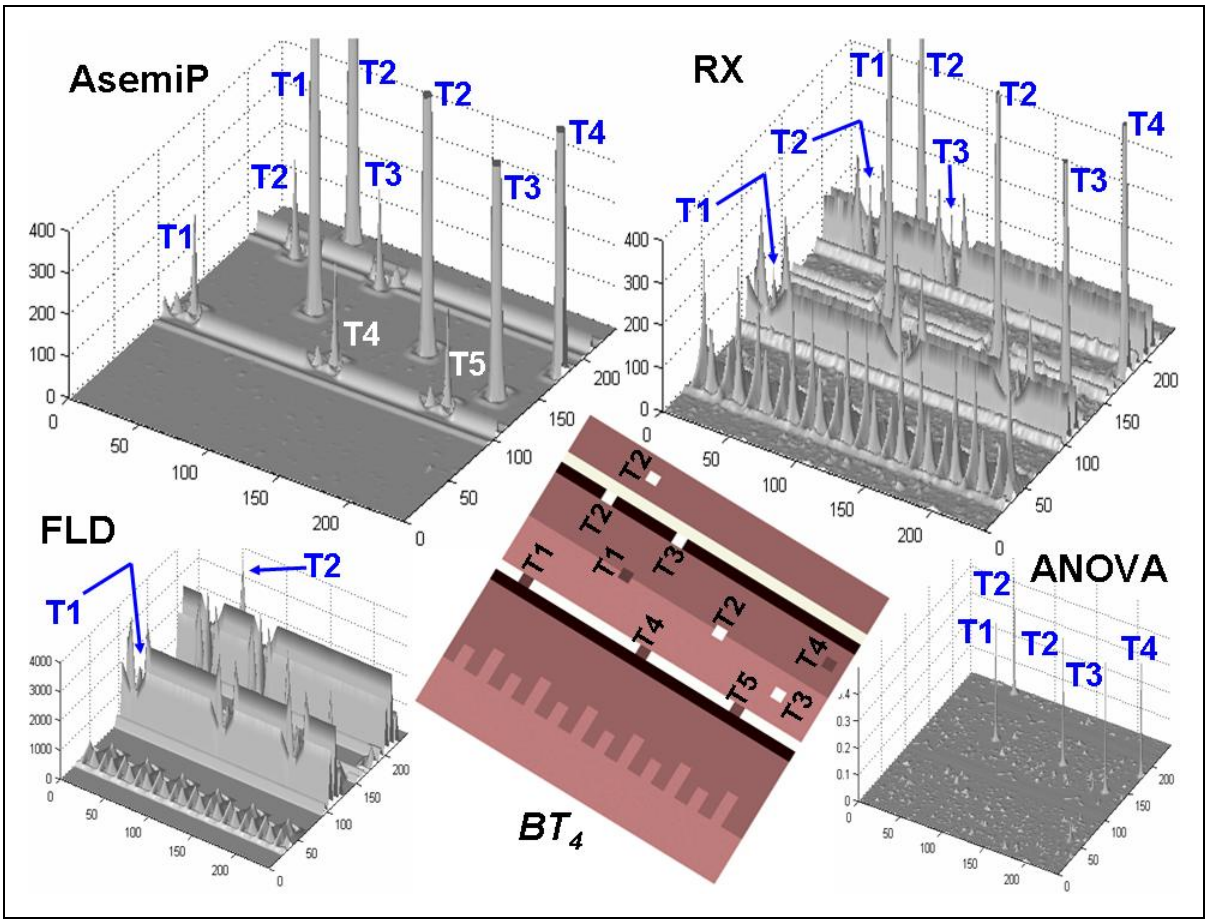

Figure 7. Examples of intermediate result OUTPUT, as described in text. The peaks are responses from the *10* targets as *seen* by the different detectors testing a single simulated realization of BT_4. Notice the artifact responses in the vicinity of some of these peaks. Some times these artifact responses are more accentuated then targets' responses—see, for instance, the responses of targets T1, T4, T5, T2, and T3 (embedded in narrow background stripes) in the RX output surface. Those artifacts also contribute to the type I error, thus, increasing its estimate in respect to the desired type I error.

Figure 7 shows that the ANOVA detector can suppress very well the challenges presented by the difficult background configuration in BT_4, but unfortunately it also suppresses the most challenging positions where targets are located. All the five targets that were located in narrow background stripes were undetected by the ANOVA detector (see targets T1, T4, T5, T2, and T3 in figure 7 [BT_4] and relate where in the ANOVA output surface their responses would be positioned).

Similar trend can be observed for the industry standard RX detector's performance, see table 21 and figure 7. As the desired type I error for the RX detector decreases, its estimated type II error increases. For instance, the responses of targets T1, T4, and T5 in one of the narrow background stripes were undetected at the cutoff threshold corresponding to the desired type I error $\alpha_3 = 10^{-3}$ and lower, and the responses of targets T2 and T3 in another narrow background

stripe were also undetected at the cutoff threshold corresponding to the desired type I error $\alpha_4 = 10^{-4}$ and lower. Thus, at $\alpha_4 = 10^{-4}$ $\alpha_4 = 10^{-4}$, the RX detector missed *5* out of *10* targets.

Similar performance loss trends were already observed for the other multivariate detectors, as shown in tables 17, 19, and 21. (Incidentally, the outputs surfaces shown for AsemiP and RX in figure 7 were clipped at *400* for displaying purposes only, some of those peaks continue to much higher values, as we will see shortly.)

Table 22 shows that the proposed detectors handled well the most difficult target-background configuration in this simulation. The AsemiP output surface shown in figure 7 illustrates in a genetic form the reason why the proposed detectors did not decrease their performance levels testing simulated realizations of BT_4. In other words, with respect to the ten target responses, these detectors could (i) suppress window-size-related background region extensions from a major background class, (ii) suppress challenging narrow background regions (see, for instance, window location *b* in figure 3), and (iii) suppress—relatively speaking—the local transitions of distinct background regions.

It is worth mentioning from results in figure 7 that the AsemiP detector's signal to noise ratios between target responses (signal) and background responses (noise) are not necessarily the same for all targets, or for that matter for the same target type located in different local background configurations (see, for instance, target responses of T1, T2, and T4 in figure 7, as they are spatially located in two different local areas, inside a narrow background stripe and outside this stripe). The same observation can be made for the other detectors' performances as well. This issue is related to signal to noise ratio, which will be addressed later. Our detectors are not completely insensitive to the local background configuration and dual window size, but they demonstrate a significant amount of insensitivity testing difficult background scenarios, which is in contrast to the performances of alternative techniques. More importantly, this contrast in performances has been consistent using these detectors on real hyperspectral imagery, see, for instance, (*13*).

4.4 Concluding Remarks

The simulation experiment results presented in section 4 were intended only as an illustration of what types of backgrounds univariate detectors perform better than multivariate detectors. In this illustration, the background complexity gradually increased from a relatively simple homogeneous class configuration, to a moderately difficult five region class configuration, to a difficult nine region class configuration having some of the local transitions of regions corresponding to the dual window size. Targets then were strategically introduced into multiple realizations of these background cubes and tested so that we could measure the effects of these changes on different anomaly detection techniques using their corresponding set of calibrated cutoff thresholds.

Results presented in subsection 4.3 demonstrated that the univariate anomaly detectors' performances are significantly less dependent on the background configuration than the performances of the multivariate detectors discussed in this report. These univariate detectors also work well accentuating the presence of targets in difficult background configurations (see figure 7 and table 22). This kind of robustness is highly desired in the target community, since real life scenarios present all kinds of background configurations.

The univariate anomaly detectors' effectiveness is less sensitive to difficult target-background configurations because they are designed to perform an *indirect comparison* between samples that are observed, in this particular case, via the inside window and samples viewed from the outside window, with both windows centered at the same spatial location in the imagery, (see (*13*) for details on the indirect comparison approach).

In doing so, these univariate detectors are better equipped to handle mixtures of classes that may be observed through the outside window, since they compare estimated means and/or variances of outside samples to corresponding estimates using the union of the outside and inside samples.

However, the univariate anomaly detectors are not completely insensitive to the local background configuration and window sizes, but they appear to be significantly less sensitive while testing difficult background scenarios, which is in contrast to the performances of the multivariate detectors.

We argued in section 1 that the deficiencies built into multivariate anomaly detectors, i.e., inability to properly handle transitions of regions, which increases the false alarm rate, is the main drive for their lack of robustness. Alternative techniques that were not investigated in this report were not proposed to outperform the industry standard RX anomaly detector in terms of yielding a higher PD and a lower false alarm rate, but to outperform the RX detector's computational time.

Since the results presented in section 4 were obtained from conducting controlled simulations, estimated PD differences for these detectors could also be controlled by designed. For instance, if we increased the number of targets in local areas characterized by a benign background configuration, the PD of the multivariate detectors—especially the RX detector—could be made to correspond to PD results yielded by our detectors. Conversely, if we increased the number of targets in difficult local background configurations, the RX detector and the other multivariate techniques would have produced significantly lower PD results.

5. Summarized Conclusions

The statistical methods and data specifications discussed in this report are intended to serve as a guide for analyzing more rigorously the strengths and weaknesses of HS anomaly detectors. Summarized conclusions on the simulation experiments discussed in sections 2, 3, and 4 are found below in subsections 5.1, 5.2, and 5.3, respectively. Additional and more detailed concluding remarks on these simulation experiments are found in subsections 2.5, 3.4, and 4.4.

5.1 Simulation Tests Using Idealized Homogeneous Samples

The simulation experiments (Experiment 1 and Experiment 2) presented in section 2 focused on showing relative strengths and weaknesses of two competing anomaly detection techniques, an univariate based anomaly detector (AsemiP) and the multivariate based industry standard (RX). Tailored hypothesis tests were devised using idealized homogeneous samples, such that, both detectors shared the same input samples.

Experiment 1 addressed a more generalized concern, i.e., the sensitivity of these detectors as a function of decreasing noise under an alternative hypothesis H_1, given that—by design—H_1 (with zero noise) was a hard alternative to be detected because it consisted of randomly generated idealized samples representing two similar homogeneous classes. (Decreasing noise in this context made these classes less separable.) The AsemiP detector showed relative strength over the RX detector in Experiment 1, see table 1. As discussed in detail in subsection 2.3, the AsemiP's favorable performance shown in table 1 is attributed to this detector's data preprocessing step, which is an integral part of this anomaly detection approach.

Experiment 2 addressed a more specialized concern, i.e., the sensitivity of these detectors as a function of increasing noise under H_0, given that—by design—H_1 (with zero noise under H_0) was an *easy* alternative to be detected because it consisted of randomly generated idealized samples representing two distinct homogeneous classes. The RX detector showed relative strength over the AsemiP detector in Experiment 2, as shown in table 2. The specialized concern addressed in this experiment could happen in the target application, as a special case, when samples of the same homogeneous class are collected under distinct atmospheric conditions and these sample sets are then used later on to represent the null hypothesis, as explained in more detail in subsection 2.3.

5.2 Simulation Tests Using Heterogeneous Samples

The simulation experiments presented in section 3 measured the sensitivity of two anomaly detectors (AsemiP and RX) to idealized sample mixtures under both null and alternative hypothesis tests, H_0 and H_1, respectively. These simulation experiments first emphasized testing the sensitivity of both detectors as a function of varying spectral magnitude in a specified data

model, and independently testing the sensitivity of both detectors as a function of varying spectral shape in the specified data model. These simulation experiments then focused on testing these detectors' abilities to suppress transitions of different regions in the imagery relative to genuine anomalies in the same experiment—this is a well known but often ignored problem in the HS image processing community. Data models consisting of free parameters were introduced to specify mixtures of difference object classes, and a set of tailored null and alternative hypotheses were introduced to mimic testing these detectors over spatial transitions of regions.

The RX technique showed high sensitivity to transitions of different regions, which can be modeled as the comparison between samples of a mixture and samples of a class belonging to that mixture. The RX technique showed also high sensitivity to magnitude (bias) differences among idealized multivariate samples. Both behaviors are highly undesirable for anomaly detection applications because this detector's PD does depend on the detector's ability to keep relatively low null-hypothesis based cutoff thresholds (transitions of regions under the null hypothesis) and on the illumination (which influences spectral magnitude) on objects in the scene.

The AsemiP technique, on the other hand, showed significantly less sensitivity to sample mixtures in the null hypothesis (owing to its sample-combining test statistic), and insensitivity to magnitude differences among input samples (owing to its data preprocessing step, which by design removes sample bias).

5.3 Simulation Tests Using Idealized Cubes

The simulation experiments presented in section 4 showed results from testing idealized imagery (cubes) with multiple univariate-based and multivariate-based anomaly detectors. This simulation experiment used exactly the same testing mechanism that these detectors would have used to test real nadir looking HS data in the target application. The emphasis of this simulation experiment was on evaluating these detectors as a function of increasing complexity level on the local configuration of clutter background. These simulation experiments used 1,500 randomly generated cubes.

Target locations in the test cubes were selected to challenge these anomaly detectors with respect to a complex interaction among the local background configuration, target spatial area, and local sampling window, which is known for affecting detectors' PD performances on real HS data. The target and window sizes (inside and outsize) were kept constant, but the local background configurations varied from easy, moderate, to difficult.

Results from these simulation experiments demonstrated that the performances of sample-combining univariate based anomaly detectors are significantly less dependent on the local background configuration than the performances of multivariate detectors that are discussed in appendix. This insensitivity to local background configuration augments the practical value of

anomaly detection in real operational scenarios, since real life scenarios present all kinds of background configurations, and different target sizes. We argued in section 1 and throughout this report that this inherent deficiency of most multivariate anomaly detectors found in the literature, i.e., inability to properly handle transitions of regions, is a main cause for producing a high false alarm rate.

6. References

1. Schowengerdt, R. A. *Remote Sensing, Models and Methods for Image Processing;* 2nd ed. San Diego, CA: Academic Press, 1997.

2. Yu, X.; Hoff, L.; Reed, I.; Chen, A.; Stotts, L. Automatic target detection and recognition in multiband imagery: A unified ML detection and estimation approach. *IEEE Tran. Image Processing* **1999**, *6*, 143–156.

3. Kwon, H.; Nasrabadi, N. Kernel RX-algorithm: a nonlinear anomaly detector for hyperspectral. *IEEE Trans. on Geoscience and Remote Sensing* **2005**, *43* (2).

4. Stocker, A. Stochastic expectation maximization (SEM) algorithm. in *Proc. DARPA Adaptive Spectral Reconnaissance Algorithm Workshop*, Jan. 1999.

5. Masson, P.; Pieczynski, W. SEM algorithm and unsupersivsed statistical segmentation of satellite images. *IEEE Trans. Geosci. Remote Sensing* **1993**, *31*, 618–633.

6. Manolakis, D.; Marden, D.; Shaw, G. Target detection algorithms for hyperspectral imaging applications. *Lincoln Laboratory Journal* **2003**, *14* (1).

7. Rosario, D. Highly Effective Logistic Regression Model for Signal Detection. *Proc. IEEE ICASSP*, Montreal, Canada, May 17–21, 2004.

8. Lehmann, E. L. *Testing Statistical Hypotheses*, New York: transferred to Chapman & Hall, 2nd Edition, pp 68, 74, Chapter 3, 1993.

9. Beaven, S.; Stein, D.; Hoff, L. Comparison of Gaussian mixture and linear mixture classification of hyperspectral data. in *Proc. IGARSS 2000* Honolulu, HI, July 24-28.

10. Duda, R. O.; Hart, P. E. *Pattern Classification Scene Anal.*, Second Edition, New York: J. Wiley & Sons, 2004.

11. Schweizer, S. M.; Moura, J. M. F. Hyperspectral imagery: clutter adaptation in anomaly detection. *IEEE Trans. Information Theory* **2000**, *46*, (5), 1855–1871.

12. Rosario, D.; Rauss, P. Experimental Results Detecting Camouflaged Soldiers in VNIR Hyperspectral Imagery. in *Proc. 2005 Military Sensing Symposia (MSS)*, 14–18 Feb 2005.

13. Rosario, D. *Alternative Asymmetric Hypothesis Tests for Hyperspectral Imagery*; ARL-TR-3712; U.S. Army Research Laboratory: Adelphi, MD, Feb 2006.

INTENTIONALLY LEFT BLANK

Appendix

In this appendix, details are discussed on miscellaneous results that are mentioned throughout the report, e.g., the data transformation as the first step prior to the application of univariate anomaly detectors, the mathematical representation of all of the detectors used for this report (see also (*13*) for additional details), full table of parameter estimations used for one of the data models.

Data Preprocessing (Feature Space)

Figure A-1 depicts the dual rectangular window that was used to sample local information from the simulated data discussed in this document.

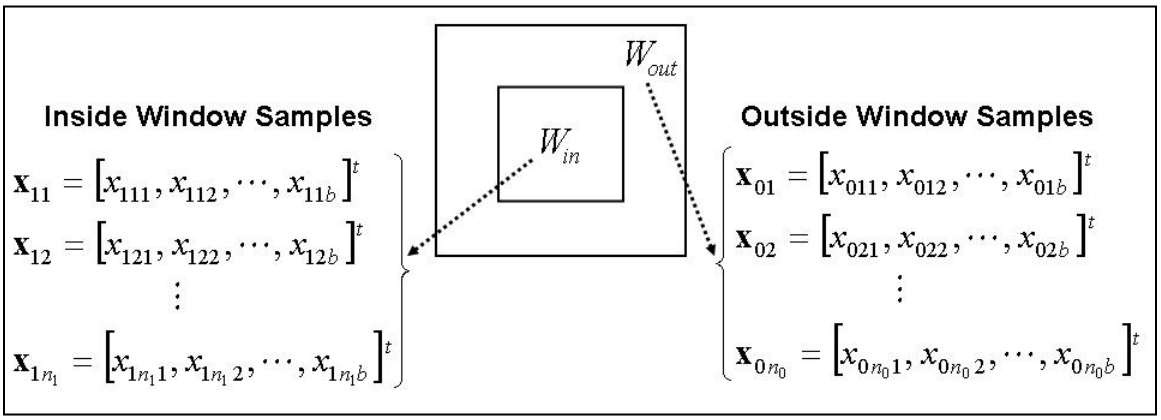

Figure A-1. Dual rectangular window.

In the literature (see references in the report), one would be hard pressed to find hyperspectral (or multispectral) anomaly detectors that are based on univariate methods—an unnatural choice given that such a data type is in fact multivariate. We proposed, for reasons discussed in the report, to transform or to preprocess the data into a univariate space and to apply our detectors in that space. Details follow.

From the inside and outside samples shown in figure 1, apply a first order differentiation—an approximation—and compute their averages after differentiating, i.e.,

$$\begin{aligned}
\Delta_{11} &= \left[x_{112} - x_{111},\ x_{113} - x_{112},\ \cdots,\ x_{11b} - x_{11(b-1)} \right]^t \\
\Delta_{12} &= \left[x_{122} - x_{121},\ x_{123} - x_{122},\ \cdots,\ x_{12b} - x_{12(b-1)} \right]^t \\
&\vdots \\
\Delta_{1n_1} &= \left[x_{1n_1 2} - x_{1n_1 1},\ x_{1n_1 3} - x_{1n_1 2},\ \cdots,\ x_{1n_1 b} - x_{1n_1 (b-1)} \right]^t
\end{aligned} \tag{A-1}$$

and

$$\begin{aligned}
\Delta_{01} &= \left[x_{012} - x_{011},\ x_{013} - x_{012},\ \cdots,\ x_{01b} - x_{01(b-1)} \right]^t \\
\Delta_{02} &= \left[x_{022} - x_{021},\ x_{023} - x_{022},\ \cdots,\ x_{02b} - x_{02(b-1)} \right]^t \\
&\vdots \\
\Delta_{0n_0} &= \left[x_{0n_0 2} - x_{0n_0 1},\ x_{0n_0 3} - x_{0n_0 2},\ \cdots,\ x_{0n_0 b} - x_{0n_0 (b-1)} \right]^t
\end{aligned} \tag{A-2}$$

where $[\cdot]^t$ denotes the vector transpose operator.

Compute the averages of resulting vectors in equation A-1 and A-2, or

$$\begin{aligned}
\bar{\Delta}_0 &= \frac{1}{n_0} \sum_{i=1}^{n_0} \Delta_{0i} \\
\bar{\Delta}_1 &= \frac{1}{n_1} \sum_{i=1}^{n_1} \Delta_{1i}
\end{aligned} \quad . \tag{A-3}$$

Compute features as

$$\begin{aligned}
y_{0i} &= \frac{180}{\pi} \arccos\left(\frac{\Delta_{0i}^t \bar{\Delta}_0}{\|\Delta_{0i}\| \|\bar{\Delta}_0\|} \right) \\
y_{1i} &= \frac{180}{\pi} \arccos\left(\frac{\Delta_{0i}^t \bar{\Delta}_1}{\|\Delta_{0i}\| \|\bar{\Delta}_1\|} \right)
\end{aligned} \tag{A-4}$$

where the operator $\|\mathbf{x}\|$ denotes the squared root of $\mathbf{x}^t\mathbf{x}$, and $i = 1, \cdots, n_0$.

From equation A-4, the outside (reference) feature sequence is

$$y_0 = \left(y_{01},\ y_{02},\ \cdots,\ y_{0n_0} \right) \tag{A-5}$$

and the inside (test) feature sequence is

$$y_1 = \left(y_{11},\ y_{12},\ \cdots,\ y_{1n_0} \right). \tag{A-6}$$

Notice that both the reference and test feature sequences have size n_0.

Detectors AsemiP, AFT, AVT, and ANOVA use (5A) and (6A) as input data.

Using the denotations shown in figure 1, detectors RX, FLD, EST, and DPC use matrices $\mathbf{X}_{\text{in}}$ (test) and $\mathbf{X}_{\text{out}}$ (reference) shown below to extract input data,

$$\mathbf{X_{in}} = \begin{bmatrix} x_{111},\ x_{121},\ \cdots,\ x_{1n_1 1} \\ x_{112},\ x_{122},\ \cdots,\ x_{1n_1 2} \\ \vdots \\ x_{11b},\ x_{12b},\ \cdots,\ x_{1n_1 b} \end{bmatrix} \tag{A-7}$$

and

$$\mathbf{X_{out}} = \begin{bmatrix} x_{011},\ x_{021},\ \cdots,\ x_{0n_1 1} \\ x_{012},\ x_{022},\ \cdots,\ x_{0n_1 2} \\ \vdots \\ x_{01b},\ x_{02b},\ \cdots,\ x_{0n_1 b} \end{bmatrix}. \tag{A-8}$$

The detectors used in this simulation are discussed next.

Approximation to Semiparametric (AsemiP) Anomaly Detector

Our interest to pursue a *principle of indirect comparison* led us to the semiparametric test *statistic*—discussed in the report, which in turn motivated us to propose a significantly simpler and computationally faster detector, the AsemiP anomaly detector. We showed in the report that the AsemiP detector can approximately yield the performance level achieved by the semiparametric detector testing real HS data.

The AsemiP anomaly detector is defined as

$$Z_{AsemiP} = \frac{\hat{\beta}^2}{\left(\frac{1}{n_0}+\frac{1}{n_0}\right)} \frac{S_t^2}{S^4}\ , \tag{A-9}$$

where from (5A) and (6A) using $n = 2n_0$

$$t = \left(y_{11},\cdots,y_{1n_0},y_{01},\cdots,y_{0n_0}\right) = \left(t_1,...,t_n\right)\ , \tag{A-10}$$

$$S_t^2 = (n-1)^{-1}\sum_{i=1}^{n}(t_i-\bar{t})^2,\ \ \bar{t} = n^{-1}\sum_{i=1}^{n}t_i\ , \tag{A-11}$$

$$S^4 = \left(\frac{(n_0-1)S_1^2+(n_0-1)S_0^2}{(n_0-1)+(n_0-1)}\right)^2, \tag{A-12}$$

$$S_1^2 = \frac{\sum_{i=1}^{n_0} \left(y_{1i} - \overline{y}_1\right)^2}{n_0 - 1}, \qquad \text{(A-13)}$$

$$\overline{y}_1 = n_0^{-1} \sum_{i=1}^{n_0} y_{1i}$$

$$S_0^2 = \frac{\sum_{i=1}^{n_0} \left(y_{0i} - \overline{y}_0\right)^2}{n_0 - 1}, \qquad \text{(A-14)}$$

$$\overline{y}_0 = n_0^{-1} \sum_{i=1}^{n_0} y_{0i}$$

and

$$\hat{\beta}^2 = \left(\overline{y}_1 - \overline{y}_0\right)^2. \qquad \text{(A-15)}$$

Asymptotic F Test (AFT) Anomaly Detector

Our interest to pursue a principle of indirect comparison using as reference the one-way ANOVA approach motivated us to propose the AFT anomaly detector, which is defined as follows

$$Z_{AFT} = \rho \frac{\hat{\beta}^2}{\hat{\beta}_t^2} \frac{S_t^2}{S^2}, \qquad \text{(A-6A)}$$

where, from equations A-5 and A-6A, $n = 2n_0$, $\rho = \left(n^{-1}\right)\left(n_0^{-1} + n_0^{-1}\right)^{-1}$, $\hat{\beta}^2$ was defined in equation (A-15), $\hat{\beta}_t = \overline{t}_1 - \overline{y}_0$ [using equation A-10 and the mean averages in equations A-11 and A-14], S_t^2 was defined in equation A-11, and S^2 is the square root of equation A-12.

Asymmetric Variance Test (AVT) Anomaly Detector

Our interest to pursue a *principle of indirect comparison* by merely exploiting differences of estimated variances motivated us to propose the AVT anomaly detector, which is defined as follows

$$Z_{AVT} = n_0 \frac{\left(S_0^2 - S_t^2\right)^2}{\hat{\zeta}_0^2} \qquad \text{(A-17)}$$

where, S_t^2 was defined in equation A-11, S_0^2 was defined in equation A-14, and using equation A-5 and the mean average in equation A-14

$$\hat{\zeta}_0^2 = \sum_{j=1}^{n_0} \frac{\left[\left(y_{0j} - \bar{y}_0\right)^2 - S_0^2\right]^2}{n_0 - 1}. \tag{A-18}$$

Reed-Xi (RX) Algorithm

This detector is a generalized version of the adaptive spectral matched filter, where the problem was formulated to detect objects of a known spatial pattern but unknown spectral distribution against a clutter background with unknown spectral distribution against a clutter background with unknown spectral covariance. This detector has been claimed to be one of the most robust detection techniques for the detection of a spectral anomaly in multispectral imagery—see references in report. It was employed by the DARPA MUSIC program to detect military vehicles in an intense clutter background, and it became known in the community as the RX anomaly detector. Eventually this detector became the industry standard for utility and comparison. Using the context of our discussion, a popular version of the RX anomaly detector is shown below:

$$Z_{RX} = \frac{n_0 n_1}{n_0 + n_1}\left(\bar{\mathbf{x}}_{\mathbf{in}} - \bar{\mathbf{x}}_{\mathbf{out}}\right)^{\mathrm{t}}_{(1\mathrm{x}b)}\left(\mathbf{C}^{-\mathbf{1}}_{\mathbf{out}}\right)_{(b\mathrm{x}b)}\left(\bar{\mathbf{x}}_{\mathbf{in}} - \bar{\mathbf{x}}_{\mathbf{out}}\right)_{(b\mathrm{x}1)}, \tag{A-26}$$

where $\bar{\mathbf{x}}_{in}$ is the sample mean vector using the columns of $\mathbf{X}_{\mathrm{in}}$ in (7A), $\bar{\mathbf{x}}_{out}$ is the sample mean vector using the columns of $\mathbf{X}_{\mathrm{out}}$ in (8A), and $\mathbf{C}_{out}^{-1}$ is the inverse of a sample covariance matrix using the rows of $\mathbf{X}_{\mathrm{out}}$.

Analysis of Variance (ANOVA) Anomaly Detector

Our interest to have a well known method operating in the same feature space of our detectors' feature space motivated us to adapt the ANOVA method into anomaly detection. In the context of our discussion, the ANOVA detector is defined as

$$Z_{ANOVA} = \frac{\sum_{i=0}^{1} n_0\left(\bar{y}_i - \bar{\bar{y}}\right)^2}{S^2} \tag{A-19}$$

where, S^2 is the square root of equation A-12, $\bar{y}_i$ ($i = 0, 1$) is defined in equations A-13 and A-14, and using equations A-5 and A-6 $\bar{\bar{y}}$ is

$$\bar{\bar{y}} = \sum_{i=0}^{1} \sum_{j=1}^{n_0} \frac{y_{ij}}{n_0}. \tag{A-20}$$

Fisher's Linear Discriminant (FLD) Anomaly Detector

Fisher's linear discriminant analysis is a standard technique in pattern recognition. It projects the original high dimensional data onto a low dimensional space, where all the classes are well separated by maximizing the Raleigh quotient, i.e., the ratio of between-class scatter matrix determinant to within-class scatter matrix determinant. The application of the FLD detector to hyperspectral imagery has been investigated for anomaly detection and for object classification. (References are given in the report.)

A version of FLD for the two-class (anomaly or not an anomaly) problem is show below:

$$Z_{FLD} = \left| \left(\mathbf{E}^{\mathbf{t}}_{\mathbf{S}_B/\mathbf{S}_w} \right)_{(1\text{x}b)} \left(\bar{\mathbf{x}}_{\mathbf{in}} - \bar{\mathbf{x}}_{\mathbf{out}} \right)_{(b\text{x}1)} \right|, \tag{A-21}$$

where $\bar{\mathbf{x}}_{in}$ is the sample mean vector using the columns of $\mathbf{X}_{\text{in}}$ in (7A), $\bar{\mathbf{x}}_{out}$ is the sample mean vector using the columns of $\mathbf{X}_{\text{out}}$ in (8A), $\left| \bullet \right|$ denotes the absolute value operator, and $\mathbf{E}^{\mathbf{t}}_{\mathbf{S}_B/\mathbf{S}_w}$ is the transposed highest energy eigenvector (*1* x *b*) from the principal component decomposition using as input the scatter matrices ratio $\mathbf{S}_B\mathbf{S}_W^{-1}$, where

$$\begin{aligned} \mathbf{S_W} = & \sum_{i=1}^{n_{in}} (\mathbf{x}^{(i)}_{\mathbf{in}} - \bar{\mathbf{x}}_{in})(\mathbf{x}^{(i)}_{\mathbf{in}} - \bar{\mathbf{x}}_{in})^t + \\ & \sum_{i=1}^{n_{out}} (\mathbf{x}^{(i)}_{\mathbf{out}} - \bar{\mathbf{x}}_{out})(\mathbf{x}^{(i)}_{\mathbf{out}} - \bar{\mathbf{x}}_{out})^t, \end{aligned} \tag{A-22}$$

$$\begin{aligned} \mathbf{S_B} = & \sum_{i=1}^{n_{in}} (\mathbf{x}^{(i)}_{\mathbf{in}} - \bar{\mathbf{x}}_{\mathbf{total}})(\mathbf{x}^{(i)}_{\mathbf{in}} - \bar{\mathbf{x}}_{\mathbf{total}})^t + \\ & \sum_{i=1}^{n_{out}} (\mathbf{x}^{(i)}_{\mathbf{out}} - \bar{\mathbf{x}}_{\mathbf{total}})(\mathbf{x}^{(i)}_{\mathbf{out}} - \bar{\mathbf{x}}_{\mathbf{total}})^t, \end{aligned} \tag{A-23}$$

Where $\bar{\mathbf{x}}_{total}$ $\bar{\mathbf{x}}_{total}$ is the total sample mean using the columns of $\mathbf{X} = [\mathbf{X}_{\text{in}}\ \mathbf{X}_{\text{out}}]$, $\mathbf{x}^{(i)}_{in}$ and $\mathbf{x}^{(i)}_{out}$ are the *i*-th columns of $\mathbf{X}_{\text{in}}$ and $\mathbf{X}_{\text{out}}$, respectively, and $n_{in} = n_1$ and $n_{out} = n_0$.

Dominant Principle Component (DPC) and Eigen Separation Transform (EST)

The DPC and EST techniques are both based on the same general principle, i.e., data are projected from their original high dimensional space onto a significantly lower dimensional

space (in our case, only one dimension) using a criterion that promotes highest sample variability within each domain in this lower dimensional space. (References on these approaches and their applications to the problem in context are given in the report.)

Differences between DPC and EST anomaly detectors are better appreciated through their mathematical representations:

$$Z_{DPC} = \left| \left(\mathbf{E}_{out}^{\mathbf{t}} \right)_{(1\text{x}b)} \left(\overline{\mathbf{x}}_{\mathbf{in}} - \overline{\mathbf{x}}_{\mathbf{out}} \right)_{(b\text{x}1)} \right|, \tag{A-24}$$

$$Z_{EST} = \left| \left(\mathbf{E}_{\Delta\mathbf{C}}^{\mathbf{t}} \right)_{(1\text{x}b)} \left(\overline{\mathbf{x}}_{\mathbf{in}} - \overline{\mathbf{x}}_{\mathbf{out}} \right)_{(b\text{x}1)} \right|, \tag{A-25}$$

where $\overline{\mathbf{x}}_{in}$ is the sample mean vector using the columns of $\mathbf{X}_{\text{in}}$ in (A-7) $\overline{\mathbf{x}}_{out}$ is the sample mean vector using the columns of $\mathbf{X}_{\text{out}}$ in (A-8), $\mathbf{E}_{out}^{\mathbf{t}}$ is the transposed highest energy eigenvector (*1* x *b*) from the principal component decomposition using as input the covariance matrix C_{out} estimated from the rows of $\mathbf{X}_{\text{out}}$, and $\mathbf{E}_{\Delta C}^{\mathbf{t}}$ is the transposed highest positive energy eigenvector (*1* x *b*) from the principal component decomposition using as input the difference ΔC between the estimated covariance matrix C_{in} from the rows of $\mathbf{X}_{\text{in}}$ and the estimated covariance matrix C_{out} from the columns of $\mathbf{X}_{\text{out}}$, i.e, $\Delta C = \left(C_{in} - C_{out} \right)_{(b\text{x}b)}$.

Estimated Parameters (Full Table) Used to Specify Data Models in the Simulation Experiments Described in Section 2.

Table A-1. (Complete table for table 3) Estimated parameter values using real HS data from three arbitrarily chosen material classes, such that, two of these classes are relatively similar and the remaining one is distinct. Estimated means $\hat{\boldsymbol{\mu}}_0$ (class 1) and $\hat{\boldsymbol{\mu}}_1$ (class 2) were attained from two similar classes, and estimated mean $\hat{\boldsymbol{\mu}}_2$ (class 3) was obtained from a distinct class. Estimates $diag\left(\hat{\boldsymbol{\Sigma}}_0\right)$, $diag\left(\hat{\boldsymbol{\Sigma}}_1\right)$, and $diag\left(\hat{\boldsymbol{\Sigma}}_2\right)$ are the diagonal terms of $\hat{\boldsymbol{\Sigma}}_0$, $\hat{\boldsymbol{\Sigma}}_1$, and $\hat{\boldsymbol{\Sigma}}_2$, respectively. These vectors are 1 by 120 representing these statistical results from 120 frequency bands.

$\hat{\boldsymbol{\mu}}_0$ (1.0e3)	$\hat{\boldsymbol{\mu}}_1$ (1.0e3)	$\hat{\boldsymbol{\mu}}_2$ (1.0e3)	$diag\left(\hat{\boldsymbol{\Sigma}}_0\right)$ (1.0e3)	$diag\left(\hat{\boldsymbol{\Sigma}}_1\right)$ (1.0e5)	$diag\left(\hat{\boldsymbol{\Sigma}}_2\right)$ (1.0e4)
0.204325	0.175925	0.087980	0.160109	0.004161	0.011063
0.220555	0.190895	0.094050	0.121826	0.005398	0.010110
0.244355	0.216750	0.102410	0.138863	0.007945	0.012112
0.260325	0.236395	0.106035	0.158180	0.011452	0.014576
0.282800	0.260630	0.111530	0.172251	0.016150	0.015648
0.303000	0.285110	0.114955	0.192190	0.022174	0.018835
0.306015	0.295890	0.116005	0.155532	0.028275	0.018890
0.343840	0.329070	0.128410	0.195692	0.039375	0.025303
0.400310	0.385030	0.145650	0.229260	0.063204	0.033793

0.426200	0.414335	0.152510	0.253849	0.080857	0.039194
0.479340	0.469640	0.168075	0.300336	0.115043	0.046178
0.525745	0.513335	0.182130	0.440964	0.149612	0.052138
0.584655	0.570855	0.193020	0.629634	0.204707	0.060838
0.610340	0.604405	0.200715	0.622818	0.239670	0.070759
0.649755	0.640735	0.209895	0.907532	0.287662	0.084692
0.690790	0.686715	0.215935	1.049553	0.342080	0.094601
0.702835	0.697190	0.216030	1.004279	0.370186	0.104717
0.726870	0.722825	0.220090	1.348907	0.416116	0.115709
0.762175	0.757420	0.224515	1.644939	0.476546	0.121292
0.780015	0.774675	0.227240	1.767170	0.519702	0.134708
0.787180	0.786200	0.222435	1.849846	0.550012	0.129150
0.732555	0.739030	0.207110	1.669775	0.487791	0.123226
0.775930	0.773770	0.214565	1.867784	0.553223	0.137767
0.796820	0.801540	0.218215	2.201474	0.614233	0.147447
0.809625	0.808330	0.216085	2.242838	0.633637	0.146689
0.794150	0.797925	0.213265	2.272309	0.614483	0.144601
0.828195	0.827975	0.220475	2.615393	0.683276	0.162850
0.840075	0.849915	0.226515	2.615617	0.715535	0.161526
0.831380	0.849795	0.228955	2.512106	0.721765	0.160688
0.808975	0.827500	0.232415	2.415130	0.659957	0.155005
0.870040	0.890830	0.256610	2.966651	0.791883	0.172193
0.880055	0.910140	0.270255	2.880826	0.808746	0.197910
0.921765	0.951300	0.291850	3.118864	0.884936	0.228888
0.929725	0.965590	0.302075	3.239084	0.909067	0.232111
0.932220	0.969280	0.305220	3.216936	0.899421	0.241986
0.921085	0.960860	0.305565	3.111193	0.880365	0.235402
0.931650	0.971825	0.308275	3.091625	0.899690	0.246211
0.925530	0.965370	0.307330	3.177989	0.894036	0.247417
0.923355	0.963485	0.304205	3.032310	0.870426	0.258352
0.895975	0.936730	0.290005	2.919803	0.830141	0.227518
0.904015	0.940870	0.283295	2.916648	0.833355	0.237817
0.897305	0.933275	0.272225	2.827288	0.819754	0.234874
0.898590	0.929095	0.260715	2.944866	0.798283	0.225671
0.910205	0.936845	0.253370	2.799882	0.823004	0.223510
0.919870	0.941215	0.247705	2.836294	0.821598	0.234855
0.950490	0.972275	0.245425	3.206763	0.886147	0.237883
0.908910	0.939200	0.227990	2.952202	0.821648	0.224186
0.900230	0.918710	0.223610	2.725032	0.786142	0.221190
0.920960	0.943310	0.221570	2.962862	0.832184	0.229621
0.925045	0.950590	0.220935	2.952274	0.852092	0.225073
0.947905	0.959410	0.221330	3.135744	0.869849	0.243380
0.958455	0.976075	0.218455	3.165847	0.896249	0.242417
0.970680	0.984340	0.215080	3.068278	0.911621	0.260409
0.985745	0.990600	0.211135	3.181859	0.914875	0.258862
1.025030	1.034935	0.215405	3.440069	1.001250	0.275730
1.043480	1.055125	0.216075	3.393607	1.042665	0.297600
1.069190	1.065800	0.217685	3.647220	1.069506	0.305052
1.096320	1.106635	0.221330	4.174480	1.143990	0.331439
1.142055	1.140765	0.225910	4.202403	1.224413	0.350069
1.161080	1.151510	0.225105	4.282978	1.246320	0.364165
1.177265	1.172715	0.222430	4.379311	1.296082	0.400537
1.187560	1.172275	0.218775	4.562036	1.294309	0.387834
1.177625	1.171750	0.213845	4.180486	1.293686	0.384884
1.151170	1.139275	0.205155	4.049277	1.224283	0.373273
1.237895	1.212965	0.214355	4.649119	1.397175	0.414268

1.241160	1.215345	0.211450	4.898225	1.403088	0.421946
1.251420	1.225765	0.208305	4.627360	1.429392	0.422963
1.270000	1.237640	0.212275	5.315728	1.459915	0.457859
1.277710	1.249570	0.212710	4.811875	1.478746	0.436082
1.282930	1.268280	0.217945	5.131763	1.517451	0.455665
1.196590	1.187770	0.210745	4.359529	1.296109	0.385094
1.234565	1.234435	0.232650	4.502458	1.362996	0.412412
1.327255	1.342730	0.276340	5.158020	1.589815	0.490936
1.383260	1.411650	0.336030	5.029861	1.701567	0.538960
1.422875	1.463445	0.411905	5.262270	1.772021	0.560185
1.462250	1.515500	0.505030	5.300158	1.836494	0.606341
1.493065	1.561440	0.602705	5.364221	1.915909	0.715015
1.423135	1.508760	0.670420	5.226288	1.708367	0.660536
1.334520	1.426945	0.749470	4.466210	1.482968	0.728704
1.381710	1.469105	0.877970	4.985041	1.553158	0.855156
1.393405	1.491240	1.001965	4.207418	1.562387	1.112205
1.466285	1.560730	1.176320	5.637853	1.671780	1.643040
1.483555	1.583840	1.289370	5.622489	1.722372	2.056169
1.486735	1.579900	1.381130	4.934738	1.684039	2.382719
1.493840	1.587180	1.460400	5.310939	1.683869	2.732464
1.460615	1.551640	1.475775	4.581474	1.571794	2.864592
1.426705	1.520430	1.474890	4.644379	1.491099	2.829500
1.182315	1.265305	1.207680	3.184297	0.962515	1.728644
0.766795	0.843315	0.803555	1.544234	0.393423	0.751571
1.110490	1.182340	1.193975	2.928974	0.841235	1.783286
1.302920	1.384200	1.381090	3.327139	1.185519	2.190944
1.314435	1.386145	1.396215	3.478528	1.176905	2.181804
1.295460	1.371110	1.377460	3.346571	1.118177	2.190823
1.260880	1.337615	1.336440	3.119583	1.035903	1.909793
1.222720	1.289615	1.295285	3.141509	0.949031	1.779276
1.177835	1.244455	1.241470	3.044982	0.861052	1.622846
1.153900	1.214110	1.214440	2.866532	0.835313	1.572928
1.125975	1.187700	1.177085	3.012285	0.790342	1.328350
1.086095	1.150385	1.128610	2.467674	0.727008	1.279527
1.051070	1.107460	1.087065	2.789874	0.652285	1.139547
0.984130	1.042080	1.013115	2.311932	0.569328	0.868234
0.912005	0.965170	0.933495	2.117954	0.478695	0.760942
0.901550	0.955095	0.920625	2.167595	0.451016	0.734419
0.911205	0.962870	0.934750	2.021148	0.469622	0.751153
0.892230	0.931560	0.902545	2.366037	0.436912	0.735990
0.906215	0.944995	0.913460	2.347496	0.435759	0.678744
0.918200	0.962805	0.920620	2.494070	0.463938	0.736099
0.908605	0.951180	0.908170	2.525084	0.426147	0.626397
0.877635	0.920710	0.875335	2.141479	0.405547	0.550242
0.845505	0.878730	0.826515	2.072753	0.368693	0.459082
0.830430	0.867685	0.810290	2.318025	0.365784	0.476302
0.858150	0.890915	0.824305	2.375163	0.370618	0.504363
0.825355	0.851690	0.777845	2.343978	0.335777	0.465364
0.802280	0.826430	0.755920	2.285790	0.318725	0.482355
0.802610	0.819845	0.752745	2.325012	0.322827	0.454256
0.797050	0.813390	0.735130	2.245886	0.322522	0.459890
0.810120	0.816685	0.727345	2.337965	0.306014	0.401832
0.793160	0.809950	0.705925	2.490024	0.302738	0.373889
0.779695	0.792315	0.693840	2.174403	0.287481	0.321542
0.751900	0.760385	0.639680	2.374914	0.277350	0.316555

INTENTIONALLY LEFT BLANK

Distribution List

ADMNSTR
DEFNS TECHL INFO CTR
ATTN DTIC-OCP (ELECTRONIC COPY)
8725 JOHN J KINGMAN RD STE 0944
FT BELVOIR VA 22060-6218

DARPA
ATTN IXO S WELBY
ATTN J RICKLIN
3701 N FAIRFAX DR
ARLINGTON VA 22203-1714

NGA
ATTN R S RAND
12310 SURSISE VALLEY DR
MAIL STOP DN 11
RESTON VA 20191-3449

OFC OF THE SECY OF DEFNS
ATTN ODDRE (R&AT)
THE PENTAGON
WASHINGTON DC 20301-3080

US ARMY TRADOC
BATTLE LAB INTEGRATION & TECHL DIRCTRT
ATTN ATCH-B
10 WHISTLER LANE
FT MONROE VA 23651-5850

CECOM NVESD
ATTN AMSRD-CER-NV-OD J RATCHES
10221 BURBECK RD STE 430
FT BELVOIR VA 22060-5806

US MILITARY ACDMY
MATHEMATICAL SCI CTR OF EXCELLENCE
ATTN MAJ J HARTKE
PHOTONICS CENTER
WEST POINT NY 10996-1786

US ARMY ABERDEEN TEST CENTER
ATTN CSTE-DT-AT-WC-A F CARLEN
ATTN CSTE-DTC-AT-TC-N
D L JENNINGS
400 COLLERAN ROAD
ABERDEEN PROVING GROUND MD 21005-5059

US ARMY CERDEC, NVESD
ATTN AMSRD-CER-NV-MS
R DRIGGERS
ATTN AMSRD-CER-NV-ST J HILGER
ATTN AMSRD-CER-NV-ST
P PERCONTI
10221 BURBECK RD STE 430
FT BELVOIR VA 22060-5806

US ARMY ERDC
ATTN CEERD-TR-S
7701 TELEGRAPH RD BLDG 2592
ALEXANDRIA VA 22315

US ARMY MATERIEL SYS ANAL ACTVTY
ATTN AMSRD-AMS-SC G KISTNER
ATTN AMSRD-AMS-SC J MAZZ
ATTN AMSRD-AMS-SC R WHEELER
392 HOPKINS RD
ABERDEEN PROVING GROUND MD 21005-5071

US ARMY NATICK RDEC ACTING TECHL DIR
ATTN SBCN-TP P BRANDLER
KANSAS STREET BLDG 78
NATICK MA 01760-5056

US ARMY PM NV/RSTA
ATTN SFAE-IEW&S-NV
10221 BURBECK RD
FT BELVOIR VA 22060-5806

COMMANDER
US ARMY RDECOM
ATTN AMSRD-AMR W C MCCORKLE
5400 FOWLER RD
REDSTONE ARSENAL AL 35898-5000

US ARMY RDECOM AMRDEC
ATTN AMSRD-AMR-SG W PITTMAN
ATTN AMSRD-AMR-SG-IP
H F ANDERSON
BLDG 5400
REDSTONE ARSENAL AL 35809

US ARMY RDECOM AMRDEC
ATTN AMSRD-AMR-SG-IR R SIMS
BLDG 5400
REDSTONE ARSENAL AL 35898

US ARMY RDECOM AMRDEC
ATTN AMSRD-AMR-WS-PL
W DAVENPORT
BLDG 7804
REDSTONE ARSENAL AL 35898

US ARMY RDECOM ARDEC
ATTN AMSRD-AAR-AEP-S P WILLSON
ATTN AMSRD-AAR-AER-S J ROMANO
ATTN AMSRD-AAR-AEP-S
P GRANGER
PICATINNEY ARSENAL NJ 07806-5000

US ARMY RDECOM TARDEC
ATTN AMSRD-TAR-R G R GERHART
6501 E ELEVEN RD MS 263
WARREN MI 48397-5000

US ARMY RSRCH LAB
ATTN AMSRD-ARL-CI-OK-TP TECHL
LIB T LANDFRIED
BLDG 4600
ABERDEEN PROVING GROUND MD
21005-5066

DIRECTOR
US ARMY RSRCH LAB
ATTN AMSRD-ARL-RO-EL L DAI
ATTN AMSRD-ARL-RO-M D ARNEY
ATTN AMSRD-ARL-RO-MI
R ZACHERY
ATTN AMSRD-ARL-RO-MM
M-H CHANG
PO BOX 12211
RESEARCH TRIANGLE PARK NC 27709-
2211

US ARMY SOLDIER & BIOLOGICAL
CHEM CTR
ATTN AMSRD-ECB-RT-DE W LOEROP
ATTN AMSRD-ECB-RT-DE
A BEN-DAVID
ATTN AMSRD-ECB-RT-DE E ROESE
EDGEWOOD CHEM & BIOLOGICAL CTR
BLDG E-5554
ABERDEEN PROVING GROUND MD
21010-5424

COMMANDER
USAISEC
ATTN AMSEL-TD BLAU
BUILDING 61801
FT HUACHUCA AZ 85613-5300

AFRL/SNAA
ATTN M JARRATT
2241 AVIONICS CIRCLE, AREA B,
BLDG 620
WRIGHT PATTERSON AFB OH 45433-
7321

CMTCO
ATTN MAJ A SUZUKI
1030 S HIGHWAY A1A
PATRICK AFB FL 23925-3002

US GOVERNMENT PRINT OFF
DEPOSITORY RECEIVING SECTION
ATTN MAIL STOP IDAD J TATE
732 NORTH CAPITOL ST., NW
WASHINGTON DC 20402

SITAC
ATTN H STILES
ATTN K WHITE
ATTN R DOWNIE
11981 LEE JACKSON MEMORIAL HWY
SUITE 500
FAIRFAX VA 22033-3309

US ARMY RSRCH LAB
ATTN AMSRD-ARL-WM-BF G HAAS
ABERDEEN PROVING GROUND MD
21005-5067

US ARMY RSRCH LAB
ATTN AMSRD-ARL-WM-BF
W OBERLE
BLDG 390
ABERDEEN PROVING GROUND MD
21005-5067

DIRECTOR
US ARMY RSRCH LAB
ATTN AMSRD-ARL-RO-EV W D BACH
PO BOX 12211
RESEARCH TRIANGLE PARK NC 27709

US ARMY RSRCH LAB
ATTN AMSRD-ARL-CI-OK-T TECHL
PUB (2 COPIES)
ATTN AMSRD-ARL-CI-OK-TL TECHL
LIB (2 COPIES)
ATTN AMSRD-ARL-D J M MILLER
ATTN AMSRD-ARL-SE J PELLEGRINO
ATTN AMSRD-ARL-SE-S J EICKE
ATTN AMSRD-ARL-SE-SA N SROUR
ATTN AMSRD-ARL-SE-SE D ROSARIO
(6 COPIES)
ATTN AMSRD-ARL-SE-SE L KAPLAN
ATTN AMSRD-ARL-SE-SE
N NASRABADI
ATTN AMSRD-ARL-SE-SE
P GILLESPIE
ATTN IMNE-ALC-IMS MAIL &
RECORDS MGMT
ADELPHI MD 20783-1197

INTENTIONALLY LEFT BLANK