

© 2023 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Chang, Chein-I, Chien-Yu Lin, and Peter Fuming Hu. "Band Sampling of Hyperspectral Anomaly Detection in Effective Anomaly Space." IEEE Transactions on Geoscience and Remote Sensing, 2023, 1–1.
<https://doi.org/10.1109/TGRS.2023.3347434>.

<https://doi.org/10.1109/TGRS.2023.3347434>

Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.

Band Sampling of Hyperspectral Anomaly Detection in Effective Anomaly Space

Chein-I Chang^{ID}, *Life Fellow, IEEE*, Chien-Yu Lin^{ID}, and Peter Fuming Hu^{ID}

Abstract—This article investigates four issues, background (BKG) suppression (BS), anomaly detectability, noise effect, and interband correlation reduction (IBCR), which have significant impacts on its performance. Despite that a recently developed effective anomaly space (EAS) was designed to use data sphering (DS) to remove the second-order data statistics characterized by BKG, enhance anomaly detectability, and reduce noise effect, it does not address the IBCR issue. To cope with this issue, this article introduces band sampling (BSam) into EAS to reduce IBCR and further suppress BKG more effectively. By implementing EAS in conjunction with BSam (EAS-BSam), these four issues can be resolved altogether for any arbitrary anomaly detector. It first modifies iterative spectral-spatial hyperspectral anomaly detection (ISSHAD) to develop a new variant of ISSHAD, called iterative spectral-spatial maximal map (ISSMax), and then generalizes ISSMax to EAS-ISSMax, which further enhances anomaly detectability and noise removal. Finally, EAS-BSam is implemented to reduce IBCR. As a result, combining EAS, BSam, and ISSMax yields four versions: EAS-BSam, EAS-ISSMax, BSam-SSMAX, and EAS-BSam-SSMAX. Such integration presents a great challenge because all these four versions are derived from different aspects, iterative spectral-spatial feedback process, compressive sensing, and low-rank and sparse matrix decomposition. Experiments demonstrate that EAS-BSam and EAS-BSam-SSMAX show their superiority to ISSHAD and many current existing hyperspectral anomaly detection (HAD) methods.

Index Terms—Anomaly detection (AD), background (BKG), band sampling (BSam), BKG suppression (BS), data sphering (DS), effective anomaly space (EAS), interband correlation reduction (IBCR), iterative spectral-spatial maximal map (ISSMax).

Manuscript received 21 September 2023; revised 19 November 2023 and 17 December 2023; accepted 23 December 2023. Date of publication 25 December 2023; date of current version 8 January 2024. The work of Chein-I Chang was supported in part by the YuShan Fellow Program sponsored by the Ministry of Education in Taiwan and in part by the National Science and Technology Council (NSTC) under Grant 111-2634-F-006-012. (*Corresponding author: Chein-I Chang.*)

Chein-I Chang is with the Center for Hyperspectral Imaging in Remote Sensing (CHIRS), Information and Technology College, Dalian Maritime University, Dalian 116026, China, also with the Remote Sensing Signal and Image Processing Laboratory, Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore, MD 21250 USA, and also with the Department of Electrical Engineering, National Cheng Kung University, Tainan 70101, Taiwan (e-mail: cchang@umbc.edu).

Chien-Yu Lin is with the Remote Sensing Signal and Image Processing Laboratory, Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore, MD 21250 USA (e-mail: chienyu2@umbc.edu).

Peter Fuming Hu is with the Department of Anesthesia, University of Maryland School of Medicine, R Adams Cowley Shock Trauma Center, Shock Trauma Anesthesia Organized Research Center, Baltimore, MD 21201 USA (e-mail: phu@som.umaryland.edu).

Digital Object Identifier 10.1109/TGRS.2023.3347434

1558-0644 © 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See <https://www.ieee.org/publications/rights/index.html> for more information.

NOMENCLATURE

AD	Anomaly detection.
ADBS	Anomaly detection BKG suppressibility.
ADP	Anomaly detection probability.
BDP	BKG detection probability.
BKG	Background.
BS	BKG suppression.
BSam	Band sampling.
BSel	Band selection.
DS	Data sphering.
EAS	Effective anomaly space.
IBCR	Interband correlation reduction.
IC	Independent component.
ISSHAD	Iterative spectral-spatial hyperspectral anomaly detection.
ISSMax	Iterative spectral-spatial maximal map.
JAD	Joint anomaly detectability.
JBS	Joint BKG suppressibility.
LSDM-MoG	Low-rank and sparse decomposition model with a mixture of Gaussians.
LRR	Low-rank representation.
OADP	Overall anomaly detection probability.
PTA	Prior-based tensor approximation.
RGAE	Robust graph autoencoder.
ROC	Receiver operating characteristic.
RXAD	Reed-Xiaoli anomaly detector.
SC	Sparsity cardinality.
SBPR	Signal-to-BKG probability ratio.
SSMax	Spectral-spatial maximal map.
TD	Target detection.

I. INTRODUCTION

AD IS a fundamental task in signal and image processing due to several unique properties [1], [2]. The first and foremost is the presence of anomalies that cannot be detected by prior knowledge or visual inspection. Accordingly, there is no training data that can be obtained to train anomalies. Otherwise, AD can be considered as TD [3]. Second, the appearance of anomalies is unexpected. That is, anomalies in AD can be only detected by surveillance as opposed to targets in TD, which can be detected by reconnaissance [1], [2]. Third, the probability of anomalies appearing in data is relatively small. Consequently, the detectability of anomalies is very low. Finally, and most importantly, the population of anomalies is rather limited and very small. As a result, anomalies cannot

be described and characterized by Gaussian distributions. Unfortunately, this Gaussian property has been widely used by deep learning-based AD methods to reconstruct or estimate BKG and then perform AD on its residual space resulting from either separating or removing BKG. All the abovementioned properties make AD very challenging.

A. Related Works

Many approaches have been investigated for hyperspectral AD (HAD) over the past years. Early approaches are based on matched filters derived from statistical detection theory [4], which can be considered statistical anomaly detectors. The most celebrated one is the anomaly detector developed by Reed and Yu [5], referred to as RXAD, which was obtained by a generalized likelihood test (GLRT). This is because the targets to be detected as anomalies are unknown and the traditional signal detector specified by LRT is not applicable. Due to such nature of anomalies with unknown properties, RXAD is derived from the Gaussian maximum likelihood estimator [6], which turns out to be the Mahalanobis distance.

However, since anomalies are not known, they can be only distinguished from their surrounding neighbor samples. To address this issue, the spectral-based RXAD was extended to ISSHAD by incorporating spatial information of anomalies into an anomaly detector [7].

Other approaches are based on an assumption that BKG and anomalies can be characterized by low-rank space, \mathbf{L} , and sparse space, \mathbf{S} , respectively, so that AD can be performed on \mathbf{S} . In other words, instead of directly designing a new anomaly detector, the idea is to separate \mathbf{L} from a hyperspectral image (HSI), \mathbf{X} , and then, the residual space will be designated as \mathbf{S} . Two ways have been investigated along this approach. One estimates or reconstructs BKG as \mathbf{L} and then separates or removes \mathbf{L} from \mathbf{X} to yield a sparse space \mathbf{S} in which an anomaly detector can operate. This gives rise to low-rank and sparse representation (LRaSR)-based approaches for AD. A typical example is to find a BKG dictionary by an LRR, which formulates AD as the following basic optimization problem [8]:

$$\min_{\mathbf{S}, \mathbf{E}} \{ \text{rank}(\mathbf{S}) + \lambda \|\mathbf{E}\|_{2,1} \} \quad \text{subject to } \mathbf{X} = \mathbf{D}\mathbf{S} + \mathbf{E} \quad (1)$$

where \mathbf{D} is the BKG dictionary with \mathbf{S} being representation coefficients, \mathbf{E} is its residual, and $\|\cdot\|_{2,1}$ is the $l_{2,1}$ -norm defined as the sum of l_2 -norm of the column of a matrix. The parameter λ is used to regularize the sparse representation \mathbf{E} . In other words, $\mathbf{D}\mathbf{S}$ and \mathbf{E} represent BKG and anomalies matrices, respectively.

As an alternative to LRaSR in (1), the other is to directly decompose \mathbf{X} into \mathbf{L} and \mathbf{S} so that anomalies can be better detected in \mathbf{S} while suppressing BKG specified by \mathbf{L} . An early approach is robust principal component analysis (RPCA) developed in [9] that can be considered as a blind source separation technique to decompose \mathbf{X} into

$$\mathbf{X} = \mathbf{L} + \mathbf{S} \quad (2)$$

in a matrix decomposition. An issue arising from the model in (2) is that it does not separate noise from the sparse

space \mathbf{S} . Thus, a more attractive model, called low-rank and sparse matrix decomposition (LRaSMD), was further developed in [9], which represents \mathbf{X} as a linear decomposition of a low-rank matrix, \mathbf{L} , and a sparse matrix, \mathbf{S} , plus a noise matrix \mathbf{N} by

$$\mathbf{X} = \mathbf{L} + \mathbf{S} + \mathbf{N} \quad (3)$$

where an approach, called go decomposition (GoDec) developed in [10], can be used to find a desired low-rank matrix \mathbf{L} and a sparse matrix \mathbf{S} . Unfortunately, a key issue in using LRaSMD is how to appropriately determine the ranks of \mathbf{L} and \mathbf{S} . To the best of our knowledge, this issue remains unsolved until [11].

Most recently, He et al. [12] developed a tensor-based low-rank approximation (TLRA) model that implements a multisubspace learning strategy to reconstruct the BKG component and also incorporates a group sparse constraint into the model to better dissociate the anomalous component from the original hyperspectral tensor. In particular, the ε -shrinkage tensor nuclear norm (ε -TNN) regularization was introduced into the model to obtain an accurate estimation of the low-rank component so that the coefficient tensor could adaptively penalize the tensor singular values with different weights to effectively address the overshrinkage problem where an effective iterative optimization algorithm based on the alternating direction method of multipliers (ADMM) was proposed to solve the proposed model. Also, Guo et al. [13] developed an antinoise hierarchical mutual-incoherence-induced discriminative learning (AHMID) method for the AD of HSI to address the issues of BKG, anomalies, and noise where the structural incoherence constraint, the first-order statistic constraint, and decentralization constraints are used to separate the BKG from the anomaly, then the noise is modeled as a mixture of Gaussian and strip noise, and, finally, a hierarchical structure is used to address the problem of inadequate single learning, so as to enhance the discrimination of BKG and anomaly with the removal of complex noise.

B. Motivation

The models (1)–(3) only offer separation of the two spaces, \mathbf{L} and \mathbf{S} , with anomalies indeed sandwiched between BKG and noise, an issue that still remains. Thus, in order to perform AD effectively, BKG, anomalies, and noise must be taken care of simultaneously. On the one end, anomalies can be embedded in BKG. On the other hand, anomalies can be mistakenly treated as noise. To address this issue, the notion of EAS was coined in [14] and developed to meet this need. It proposed a sandwich model that considers anomalies to be sandwiched between BKG and noise and then separates anomalies from both BKG and noise so that anomalies can be detected in EAS more effectively. More specifically, EAS is a space that is particularly constructed for an arbitrary anomaly detector to suppress BKG via $\mathbf{D}\mathbf{S}$, improve anomaly detectability, and remove noise jointly and simultaneously.

Similarly, ISSHAD also utilizes full-band information to enhance anomaly detectability by capturing spatial information among classified data samples via an iterative feedback

process. However, since a hyperspectral imaging sensor uses hundreds of contiguous bands to acquire data, its provided spectral information is generally more than what an application really needs. In particular, various applications generally require different wavelengths to respond to their material substances of interest. Unfortunately, ISSHAD and EAS do not address the issue of interband correlation.

To address the abovementioned interband redundancy issue, BSEL is commonly used for this purpose. Recently, Chang [15] introduced a new notion, called BSAM, which showed that BSAM performed significantly better than BSEL. This is because BSAM is a dynamic BSEL resulting from random BSAM as opposed to BSEL, which is a deterministic BSEL derived from a specific BSEL algorithm. Thus, BSAM becomes the best means to help ISSHAD and EAS deal with IBCR. To realize such a possibility, we can implement BSAM in conjunction with either EAS or ISSHAD or both to derive four new versions, EAS-BSAM, EAS-ISSMax, BSAM-SSMAX, and EAS-BSAM-SSMax, to address the four issues arising in AD, BS, anomaly detectability, noise effect, and IBCR. It is worth noting that BSAM-SSMax and EAS-BSAM-SSMax implement a new variant of ISSHAD, called SSMax, instead of ISSHAD. This is because both ISSHAD developed in [7] and BSAM developed in [15] make use of an iterative feedback process to capture and update spatial classification information. In this case, both iterative processes are merged into a single iterative process that results in SSMax. Specifically, SSMax uses the maximal map (MaxMap) obtained by maximizing the anomaly map and its spatial filtered anomaly map to produce a feedback map to be added to the current data cube to create a new expanded data cube for an iterative process. Each of ISSMax, EAS, and BSAM has its own merit, and their effectiveness was justified in [7], [14], and [15].

Among all the four issues, the most difficult and challenging task is how to deal with noise effects, specifically, non-Gaussian noise, which is generally encountered in hyperspectral imagery. EAS utilizes two steps to combat the noise. It first uses DS to remove Gaussian noise followed by its second step that uses IC analysis (ICA) coupled with SC in [10] to remove non-Gaussian noises. Unfortunately, there may have been some non-Gaussian noises still present in EAS. It is well known that hyperspectral bands are generally contaminated by noises. To further resolve this dilemma, BSEL can be used for this purpose. However, since noises are random, BSEL may not be effective. As shown in [15], BSAM is proven to effectively remove noise effects by randomly sampling bands. This leads to a motivation to combine EAS and BSAM to better remove noise effects.

Table I summarizes the effectiveness of EAS, BSAM, ISSMax, and their joint implementations, EAS-BSAM, EAS-ISSMax, BSAM-SSMAX, and EAS-BSAM-SSMax, in addressing the issues of BS, anomaly detectability, noise effect, and IBCR where a check symbol “√” indicates a “yes” in addressing a particular issue.

C. How to Evaluate Effectiveness of EAS, BSAM, and SSMax

As discussed above, we have identified four major issues encountered in AD: 1) BS; 2) anomaly detectability; 3) noise

TABLE I
EFFECTIVENESS OF EAS, BSAM, ISSMAX, EAS-BSAM,
EAS-ISSMAX, BSAM-SSMAX, AND EAS-BSAM-SSMAX
IN ADDRESSING BS, ANOMALY DETECTABILITY,
NOISE EFFECT, AND IBCR

	BKG suppression	Anomaly detectability	noise effect	IBCR
BSAM	√		√	√
EAS	√	√	√	
SSMax		√		
EAS-BSAM	√	√	√	√
EAS-ISSMax	√	√	√	
BSAM-SSMax	√		√	√
EAS-BSAM-SSMax	√	√	√	√

effect; and 4) IBCR that have significant impacts on AD performance. A commonly used evaluation criterion for AD is to plot the 2-D ROC curve of detection probability P_D versus false alarm probability P_F [4] and then calculate its area under the 2-D ROC curve (AUC), denoted by $AUC_{(D,F)}$ as a performance measure for quantitative and P_F are determined by the same threshold τ , P_D and P_F cannot stand alone to work independently. As a consequence, $AUC_{(D,F)}$ can be only used to measure the effectiveness of a detector not the effects of the four issues 1)–4) described above. Fortunately, the drawbacks of using 2-D ROC curves were remedied by a 3-D ROC developed in [16], which is a curve of three parameters, detection probability P_D , false alarm probability P_F , and threshold parameter τ . Using such a 3-D ROC curve, three 2-D ROC curves of (P_D, P_F) , (P_D, τ) , and (P_F, τ) can be generated to calculate their respective AUC values, denoted by $AUC_{(D,F)}$, $AUC_{(D,\tau)}$, and $AUC_{(F,\tau)}$, which can be further used to define eight detection measures. Most recently, these eight detection measures were further rederived for AD measures in [17], which happen to be the right performance measures to evaluate the effectiveness of the four issues identified by 1)–4) for quantitative analysis.

D. Novelities

The novelties of this article are summarized as follows.

- 1) The first novelty is to investigate and identify four issues.
 - a) BS that can be achieved by DS preprocessing used in EAS;
 - b) anomaly detectability that can be enhanced by ICs generated in EAS and the iterative process to capture spatial information among classified data samples in ISSMax;
 - c) noise effects that can be removed by SC [10] implemented in EAS;
 - d) IBCR that can be resolved by random sampling used in BSAM.
- 2) The second novelty is to introduce BSAM into EAS and ISSMax to derive four new versions: EAS-ISSMax, EAS-BSAM, BSAM-SSMax, and EAS-BSAM-SSMax. The concept of BSAM is derived from random data sampling

in CS [37], which dynamically selects band subsets to reduce interband correlation via an iterative process.

- 3) The third novelty is to implement EAS in conjunction with BS_{am} (EAS-BS_{am}) to yield better AD performance where EAS is used to enhance anomaly detectability and remove Gaussian by DS and non-Gaussian noises by SC coupling, while BS_{am} is used to suppress BKG and reduce interband correlation.

E. Contributions

There are several contributions that can be summarized as follows.

- 1) We use the eight 3-D ROC curve-derived AD measures in [17] to evaluate AD methods according to their capability in addressing four issues: BS, anomaly detectability, noise effect, and interband correlation that are closely related to AD performance.
- 2) We derive a new version of ISSHAD, ISSMax, which can be coupled with BS_{am} and EAS-BS_{am}.
- 3) With different combinations of ISSMax, EAS, and BS_{am}, four new versions are derived: EAS-ISSMax, EAS-BS_{am}, BS_{am}-SSMax, and EAS-BS_{am}-SSMax.
- 4) As will be shown in experiments, EAS-BS_{am} and EAS-BS_{am}-SSMax are the best among all test anomaly detectors and also outperform many state-of-the-art AD methods.

The remainder of this article is organized as follows. Section II identifies four issues encountered in AD and reviews their related works. Section III revisits ISSHAD, EAS, and BS_{am} that play key roles in developing EAS-BS_{am}-SSMax. Section IV pairs all possible combinations of ISSMax, EAS, and BS_{am} and integrates all three to yield EAS-BS_{am}-SSMax that turns out to be the best AD system network among all test AD methods. Section V describes images to be used for experiments. Section VI presents experimental results and discussions. Section VII further discusses the experimental results in Section VI. Section VIII provides an insightful conclusion.

II. FOUR ISSUES CLOSELY RELATED TO AD PERFORMANCE AND RELATED WORKS

In order to detect anomalies effectively, this section reviews AD methods reported in the literature that were developed to address the four issues 1)–4) noted in the introduction.

- 1) *RXAD-Based Variants*: Since RXAD was developed, it has been modified and extended in various stages [6], window-based RXAD to capture local characteristics of anomalies such as local RXAD [18], adaptive RXAD [19], dual window (DWEST) [20], collaborative representation RXAD (CRD) [21], multiple windows [22], and kernel RXAD [23]. Basically, RXAD inverts a global/local covariance matrix to suppress BKG while using a data sample as a matched signature to enhance anomaly detectability.
- 2) *BKG Estimation-/Reconstruction-Based AD Methods*: No matter how well RXAD can be improved by different modifications, these RXAD-variants still suffer from

anomalies embedded in BKG, which cannot be separated from BKG. To address this issue, many recent approaches have been directed at how BKG can be effectively suppressed to improve the detectability of an anomaly detector. The LRR-based method discussed in the introduction is one of these approaches using model (1). Another is to assume that BKG can be characterized and modeled by Gaussian distributions. For example, Qu et al. [24] developed a Gaussian mixture model (GMM)-based approach to anomaly extraction and an effective GMM-based weighting approach for fusing the extracted anomaly maps, which can be further rectified by using a guided filter to obtain the final AD map (ADMap). In addition, several deep learning-based methods have been investigated, such as BKG estimation/reconstruction by using training samples such as deep brief networks [25], an autoencoder (AE) [26], a generative adversarial network (GAN) [27], [28], and an RGAE proposed in [29]. All these methods are based on an assumption that the estimation/reconstruction error is assumed to be the main cause incurred by anomalies.

- 3) *BKG Decomposition for Separation-Based AD Methods*: However, BKG estimation/reconstruction is not only way to improve BS. Another one is to decompose BKG from the data cube as a subspace such as RPCA using model (2). Similar to RPCA, Vaswani et al. [30] also developed an approach, called robust subspace learning (RSL), which treated an outlier as an additive sparse corruption via sparse+low-rank (S+LR) matrix decomposition. Using this interpretation, RPCA can be considered as a variant of RSL via S+LR formulation. Chang et al. [31] developed a BKG-anomaly component projection and separation optimized filter (BASO) to separate anomalies from BKG. In a recent report, Li et al. [32] took advantage of tensor Tucker decomposition to develop a PTA, which decomposed a hyperspectral imagery into a BKG tensor and an anomaly tensor corresponding to BKG and anomaly spaces, respectively. Most recently, Chen et al. [33] proposed the component decomposition analysis (CDA) to decompose the data cube into principal components, ICs to represent BKG and anomaly spaces, respectively, and noise space: $\mathbf{X} = \mathbf{PCs} + \mathbf{ICs} + \mathbf{N}$.

- 4) *Enhancement of Anomaly Detectability-Based AD Methods*: Interestingly, more recent efforts have been devoted to AD enhancement but not BKG removal/suppression. For example, Wang et al. [34] proposed an autonomous HAD network (Auto-AD) to reconstruct BKG by including an adaptive-weighted loss function to assign greater weights to potential anomalous pixels, so as to increase the contrast of the anomalies against BKG. Chong et al. [35] proposed a hyperspectral deep net-based anomaly detector using a weight adjustment strategy (WAHyperDNet), which addresses the mixture of anomalous pixels. Li et al. [36] modified (2) as the LSDM-MoG where the sparse component can be described by a mixture of Gaussian noises so that anomalies can be detected by the Manhattan distance. Zhang et al. [37]

took advantage of the LRSMD model [10] using the Mahalanobis distance to develop an AD method.

- 5) *IBCR-Based AD Methods*: Interband spectral correlation is always an issue resulting from hyperspectral imaging, which uses hundreds of contiguous bands. It has been shown in [15] that compared to BSEL, BSAM can better capture spectral characteristics while removing noise and correlation among bands.

Unfortunately, to the best of our knowledge, there exists no AD method reported in the literature that can address all the aforementioned four issues effectively.

III. REVIEWS OF ISSHAD, EAS, AND BSAM

It has been shown in [7] that ISSHAD can improve any arbitrary anomaly detector via an iterative feedback system. It has been also shown in [14] that an anomaly detector can be further improved by operating in EAS rather than in its original data space or a sparse space. Most interestingly, it is understood that using full bands generally does better than using BS in applications. However, BSAM in [15] showed that using random sampling could even improve using full bands and BS in AD. This section is the first time to review these three techniques altogether in a unified manner.

A. ISSHAD

For a given anomaly detector, the key idea of ISSHAD [7] is to implement a spatial filter (SF) on an ADMap to generate spatial correlation among detected anomalies. Such an ADMap is then fed back to be added to the current data cube to create an augmented data cube that contains additional spatial information of ADMap. As the same anomaly detector is implemented on this newly expanded data cube, it is expected that its performance can be further improved. It is obvious that the same process can be continued for further improvement of AD performance indefinitely. Three issues are associated with this process. One is how to design an effective SF to better capture the spatial information of an ADMap and feed the captured SF-ed ADMaps back to the current data cube. Another is how to realize this process. A third one is how to terminate the process. The ISSHAD resolves all these three issues and makes the process iterative. The main idea of ISSHAD can be briefly described in the following steps.

Algorithm 1 ISSHAD

At the j^{th} iteration,

1. 1st step: select an anomaly detector and a desired spatial filter (SF).
 2. 2nd step: implement the selected anomaly detector on the data cube $\mathbf{\Omega}^j$ to produce an ADMap j .
 3. 3rd step: apply the selected SF to the ADMap j obtained in step 2 to produce SFADMap j .
 4. 4th step: feed SFADMap j obtained in step 3 back to the current data cube $\mathbf{\Omega}^j$ to create a new data cube $\mathbf{\Omega}^{j+1} = \mathbf{\Omega}^j \cup \text{SFADMap}^j$.
 5. Repeat steps 2-4 until it satisfies a stopping rule.
-

Two key steps are crucial in developing ISSHAD. One is step 3 that generates various SFADMap j for feedback, which leads to seven variants of ISSHAD in [7]. The other is step 5 that made use of the Taminoto index (TI) to design an automatic stopping rule [38].

B. EAS

As noted in the introduction, anomalies are generally sandwiched between BKG and noise, which is an extremely challenging issue. Interestingly, a recent work [14] constructed EAS which seemed to provide a feasible solution. It assumed that BKG was characterized by the second-order data statistics and then used DS to remove BKG and Gaussian noise. Such DS process has been shown to be more effective than many approaches that used Gaussian-based models [24], [36] or LRR characterized by second-order statistics, such as matrix decomposition. In addition, to our best knowledge, many recently developed approaches were not able to address non-Gaussian noise issues. EAS can further resolve this issue by applying the FastICA [39] to the sphered data to find an IC-formed data cube where non-Gaussian noise can be removed by the number of ICs determined by VD and SC [10].

Algorithm 2 Construction of EAS

1. 1st step: let $\mathbf{\Omega}^0$ be the original data cube and j be determined by minimax-SVD.
 2. 2nd step: apply DS to produce an $\tilde{\mathbf{\Omega}}^0$.
 3. 3rd step: apply the ATGP-FastICA to $\tilde{\mathbf{\Omega}}^0$ obtained in step 2 to produce $\{\text{IC}_l\}_{l=1}^j$.
 4. 4th step: Set $\text{SC} = k = j \times N$ to create EAS.
-

The great details of how to constrict EAS in terms of algorithms can be found in Section IV, specifically, [3, Sec. IV.C]. The effectiveness of EAS in AD has been also shown in [14] that EAS can remove second-order statistics-characterized BKG and noise while removing non-Gaussian noise by SC. To avoid duplication, we refer all details to [3] and [14].

C. BSAM

Since HSI utilizes hundreds of contiguous spectral bands to acquire data, it usually provides more spectral information than what an application needs. This is particularly true for AD where the response anomalies to bands vary; some of them may be redundant, not useful, and even conflict with one another. Thus, a general approach is DR where BS has been widely used for DR. However, in order for BS to be effective, how to determine a correct n_{BS} and how to design an appropriate BS selection algorithm are key issues to its success. Surprisingly, Chang [15] introduced a new notion, called BSAM, which was derived from random sampling, also known as compressive sensing (CS) based on two principles: incoherence and sparsity [40], [41]. It can be briefly described as follows.

BSAM has two great benefits over BSEL. One is that the BSAM rate n_{BSAM} can be smaller than the number of selected bands, n_{BSEL} . The other is that the bands sampled by BSAM

Algorithm 3 BSam

At the j^{th} iteration,

1. 1st step: let n_{BSam} be band sampling rate and Ω . Select a spatial filter and anomaly detector.
2. 2nd step: randomly select n_{BSam} bands from Ω as Ω_{BSam}^j .
3. 3rd step: implement the selected anomaly detector on Ω_{BSam}^j to produce ADMap^j .
4. 4th step: apply the selected SF to the ADMap^j obtained in step 2 to produce SFADMap^j .
5. 5th step: feed SFADMap^j obtained in step 3 back to the current data cube Ω^j to create a new data cube $\Omega^{j+1} = \Omega^j \cup \text{SFADMap}^j$.
6. 6th step: Check

$$\text{TI}^{(j)} = \frac{|\text{BMap}^{(j)} \cap \text{BMap}^{(j-1)}|}{|\text{BMap}^{(j)} \cup \text{BMap}^{(j-1)}|} > \varepsilon \quad (4)$$

where $\text{TI}^{(j)}$ is TI for the j^{th} iteration. If (4) is true, then the iterative process is terminated. Otherwise, continue.

7. 7th step: randomly select n_{BSam} bands from Ω to form $\Omega_{\text{BSam}}^{j+1}$. Let $j + 1 \leftarrow j$ and repeat steps 3-5.

are dynamic compared to bands selected by BSel, which are deterministic. As shown in [15], BSam performed significantly better than BSel even though n_{BSam} is underestimated.

IV. INTEGRATION OF ISSHAD, EAS, AND BSAM

As reviewed in Section III, ISSHAD, EAS, and BSam have shown their individual advantages in coupling with any arbitrary anomaly detector to enhance AD. Thus, as a follow-up step, it is highly desirable to combine their strengths to further improve AD performance.

A. Two Combinations of ISSMax, EAS, and BSam

There are three combinations of pairing ISSMax, EAS, and BSam.

1) *EAS-ISSMax*: EAS-ISSMax is to implement ISSMax in EAS, denoted by \mathbf{X}^{EAS} , which is a new variant of ISSHAD-ICSC in [7] where the feedback spatial filtered map (SFMap) is defined as

$$\text{MaxMap}^{(j)} = \max\{\text{SFMap}^{(j)}, \text{ADMap}^{(j)}\} \quad (5)$$

and then the new data cube is obtained by feeding $\text{MaxMap}^{(j)}$ back to $\Omega^{(j-1)}$ to create a new data cube, $\Omega^{(j)}$, by

$$\Omega^{(j)} = \Omega^{(j-1)} \cup \text{MaxMap}^{(j)}. \quad (6)$$

More specifically, it first constructs EAS and then operates ISSMax on EAS.

Fig. 1 depicts a block diagram of implementing Algorithm 4: EAS-ISSMax at the j^{th} iteration that first constructs the j^{th} EAS, $\mathbf{X}^{\text{EAS},j}$, and then implements SSMax using (2). If it satisfies the stopping rule, then the algorithm is terminated. Otherwise, the results of $\text{MaxMap}_{\text{EAS}}^j$ obtained by (8) will be fed back to the j^{th} data cube to create a new data cube $\mathbf{X}^{\text{EAS},j+1}$ by (9) for the next round $(j+1)^{\text{th}}$ iteration.

Algorithm 4 EAS-ISSMax

Let Ω^0 be the full band set and EAS be constructed from the original scene represented by Ω^0 and let it be $\mathbf{X}^{\text{EAS},0}$. At the j^{th} iteration

1. 1st step: input $\mathbf{X}^{\text{EAS},j}$.
2. 2nd step: implement a selected anomaly detector on $\mathbf{X}^{\text{EAS},j}$ to produce an $\text{ADMap}_{\text{EAS}}^j$.
3. 3th step: apply Otsu's method to $\text{ADMap}_{\text{EAS}}^j$ to produce a binary map, $\text{BMap}_{\text{EAS}}^j$.
4. 4th step: check the following stopping rule

$$\text{TI}^{(j)} = \frac{|\text{BMap}_{\text{EAS}}^j \cap \text{BMap}_{\text{EAS}}^{j-1}|}{|\text{BMap}_{\text{EAS}}^j \cup \text{BMap}_{\text{EAS}}^{j-1}|} > \varepsilon. \quad (7)$$

If (7) is true, then the iterative process is terminated. Otherwise, continue.

5. 5th step: apply a spatial filter (SF) specified by an edge preserving filter, EPF-G-g to the $\text{ADMap}_{\text{EAS}}^j$ obtained in step 2 to produce a $\text{SFMap}_{\text{EAS}}^j$, then use $\text{SFMap}_{\text{EAS}}^j$ and $\text{ADMap}_{\text{EAS}}^j$ to produce $\text{MaxMap}_{\text{EAS}}^j$ by

$$\text{MaxMap}_{\text{EAS}}^j = \max\{\text{SFMap}_{\text{EAS}}^j, \text{ADMap}_{\text{EAS}}^j\}. \quad (8)$$

6. 6th step: feed $\text{MaxMap}_{\text{EAS}}^j$ obtained in step 5 back to the current data cube $\mathbf{X}^{\text{EAS},j}$ to create a new data cube

$$\mathbf{X}^{\text{EAS},j+1} = \mathbf{X}^{\text{EAS},j} \cup \text{MaxMap}_{\text{EAS}}^j. \quad (9)$$

7. 7th step: Let $j + 1 \leftarrow j$ and repeat steps 2-6.

2) *BSam-SSMax*: It seems straightforward to implement ISSMax coupled with BSam. BSam-SSMax is by no means a trivial combination because each of ISSMax and BSam involves its own iterative feedback process. In this case, the iterative feedback process implemented by ISSMax is merged with the iterative feedback process carried out by BSam into one single iterative process. As a result, only the iterative feedback process implemented by BSam is retained where ISSMax is reduced to SSMax without involving an iterative process. Nevertheless, the used feedback SFMap implemented in BSam is the same as ISSMax defined by (2).

The detailed step-by-step implementation of BSam-SSMax at the j^{th} iteration is described in Fig. 2, which can be considered as a particular CS version of ISSMax using a random BSam strategy.

3) *EAS-BSam*: It has been shown in [15] that RXAD using BSam improved RXAD using full bands. It is also shown in [14] that RXAD can be further improved by operating EAS. Thus, as expected, combining EAS and BSam can further improve RXAD. This leads to the BSam of RXAD in EAS, denoted by EAS-BSam. However, it should be noted that EAS-BSam is not limited to RXAD, and it also works for any arbitrary anomaly detector.

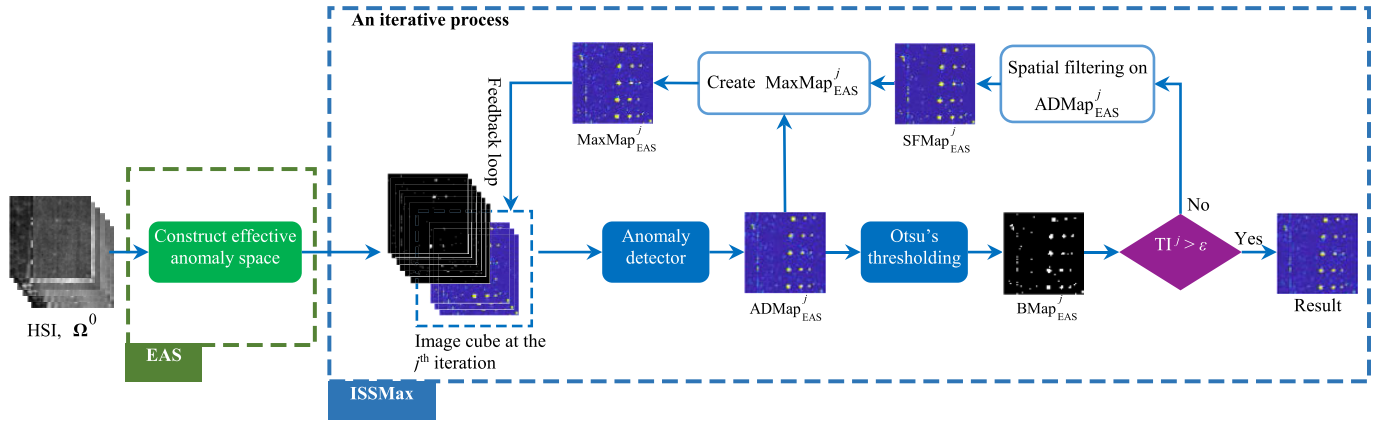
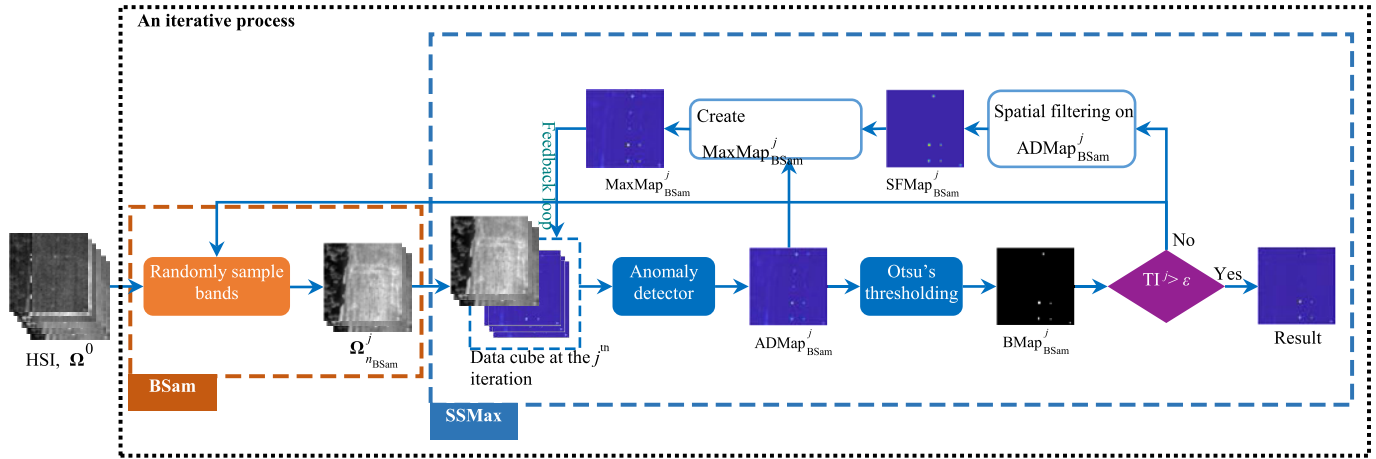
Fig. 1. Graphic block diagram of implementing EAS-ISSMax at the j th iteration.

Fig. 2. Graphic block diagram of implementing Algorithm 5.

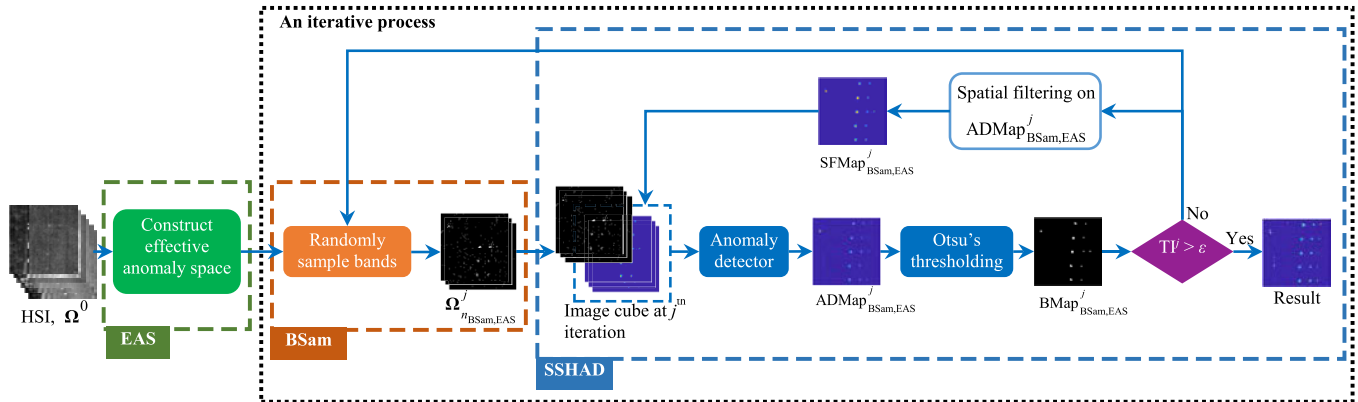


Fig. 3. Graphic block diagram of implementing Algorithm 6.

Fig. 3 shows a block diagram of implementing Algorithm 6. It should be noted that since EAS-BSam does not implement SSMax, (8) and (11) are used in step 6 for its feedback loop.

B. Three Combinations of ISSHAD, EAS, and BSAm

The EAS-BSam-SSAD presented in this section takes advantage of ISSMax, EAS, and BSAm altogether to derive

a better AD system network. In particular, like ISSHAD in [7], EAS-BSam-SSAD is also applicable to any arbitrary anomaly detector. It can be implemented in the following steps.

Fig. 4 presents a detailed step-by-step implementation block diagram of Algorithm 7. This process involves applying EAS to the HSI, followed by BSAm to randomly sampling bands. Subsequently, ISSMax is performed on the data cube formed

Algorithm 5 BSam-SSMax

Let Ω^0 be the full band set.

At the j^{th} iteration

1. 1st step: randomly select Ω_{BSam}^j from the original scene Ω^0 to form a data cube $\mathbf{X}_{\Omega_{\text{BSam}}^j}^j$.
2. 2nd step: implement a selected anomaly detector on $\mathbf{X}_{\Omega_{\text{BSam}}^j}^j$ to produce an $\text{ADMap}_{\text{BSam}}^j$.
3. 3rd step: apply Otsu's method to $\text{ADMap}_{\text{BSam}}^j$ to produce a binary map, $\text{BMap}_{\text{BSam}}^j$.
4. 4th step: check the following stopping rule.

$$\text{TI}^{(j)} = \frac{|\text{BMap}_{\text{BSam}}^j \cap \text{BMap}_{\text{BSam}}^{j-1}|}{|\text{BMap}_{\text{BSam}}^j \cup \text{BMap}_{\text{BSam}}^{j-1}|} > \varepsilon. \quad (10)$$

If (7) is true, then the iterative process is terminated. Otherwise, continue.

5. 5th step: apply a spatial filter (SF) specified by an edge preserving filter, EPF-G-g to the $\text{ADMap}_{\text{BSam}}^j$ obtained in step 2 to produce a $\text{SFMap}_{\text{BSam}}^j$ and then use $\text{SFMap}_{\text{BSam}}^j$ and $\text{ADMap}_{\text{BSam}}^j$ to produce an $\text{MaxMap}_{\text{BSam}}^j$ by

$$\text{MaxMap}_{\text{BSam}}^j = \max\{\text{SFMap}_{\text{BSam}}^j, \text{ADMap}_{\text{BSam}}^j\}. \quad (11)$$

6. 6th step: randomly select n_{BSam} bands, $\Omega_{\text{BSam}}^{j+1}$ from Ω^0 and feed $\text{MaxMap}_{\text{BSam}}^j$ obtained in step 5 to create a new data cube $\mathbf{X}_{\Omega_{\text{BSam}}^{j+1}}^{j+1}$ defined as

$$\mathbf{X}_{\Omega_{\text{BSam}}^{j+1}}^{j+1} = \Omega_{\text{BSam}}^{j+1} \cup \text{MaxMap}_{\text{BSam}}^1 \cup \dots \cup \text{MaxMap}_{\text{BSam}}^j. \quad (12)$$

7. 7th step: Let $j + 1 \leftarrow j$ and repeat steps 2-6.

Algorithm 6 EAS-BSam

Let Ω^0 be the full band set and EAS be constructed from the original scene represented by Ω^0 and let it be $\mathbf{X}^{\text{EAS},0}$. At the j^{th} iteration,

1. 1st step: input $\mathbf{X}^{\text{EAS},j}$.
2. 2nd step: randomly select $\Omega_{\text{BSam,EAS}}^j$ from $\mathbf{X}^{\text{EAS},j}$ to form $\mathbf{X}_{\Omega_{\text{BSam,EAS}}^j}^j$.
3. 3rd step: implement a selected anomaly detector on $\mathbf{X}_{\Omega_{\text{BSam,EAS}}^j}^j$ to produce an $\text{ADMap}_{\text{BSam,EAS}}^j$.
4. 4th step: apply Otsu's method to $\text{ADMap}_{\text{BSam,EAS}}^j$ to produce a binary map, $\text{BMap}_{\text{BSam,EAS}}^j$.
5. 5th step: check the following stopping rule.

$$\text{TI}^{(j)} = \frac{|\text{BMap}_{\text{BSam,EAS}}^j \cap \text{BMap}_{\text{BSam,EAS}}^{j-1}|}{|\text{BMap}_{\text{BSam,EAS}}^j \cup \text{BMap}_{\text{BSam,EAS}}^{j-1}|} > \varepsilon. \quad (13)$$

If (13) is true, then the iterative process is terminated. Otherwise, continue.

6. 6th step: apply a spatial filter (SF) specified by an edge preserving filter, EPF-G-g to the $\text{ADMap}_{\text{BSam,EAS}}^j$ obtained in step 2 to produce a $\text{SFMap}_{\text{BSam,EAS}}^j$.
7. 7th step: randomly select n_{BSam} bands, $\Omega_{\text{BSam,EAS}}^{j+1}$ from $\mathbf{X}^{\text{EAS},j}$ and feed $\text{SFMap}_{\text{BSam,EAS}}^j$ obtained in step 6 to create a new data cube $\mathbf{X}_{\Omega_{\text{BSam,EAS}}^{j+1}}^{j+1}$ defined as

$$\mathbf{X}_{\Omega_{\text{BSam,EAS}}^{j+1}}^{j+1} = \Omega_{\text{BSam,EAS}}^{j+1} \cup \text{SFMap}_{\text{BSam,EAS}}^1 \cup \dots \cup \text{SFMap}_{\text{BSam,EAS}}^j. \quad (14)$$

8. 8th step: Let $j + 1 \leftarrow j$ and repeat steps 2-7.

TABLE II
FIVE METHODS BREAK

Method	full bands	BSam	EAS	feedback map
ISSMax	√			MaxMap
EAS-ISSMax	√		√	MaxMap
BSam-SSMax		√		MaxMap
EAS-BSam		√	√	SFMap
EAS-BSam-SSMax		√	√	MaxMap

by the BSam-selected bands. The iteration continues until the stopping rule is satisfied.

Table II summarizes the configurations and specifications of ISSMax, EAS-ISSMax, BSam-SSMax, EAS-BSam, and EAS-BSam-SSMax where “√” indicates that a method uses one of three features: full bands, BSam, and EAS.

V. REAL IMAGES FOR EXPERIMENTS

In order to demonstrate the performance of ISSMax, EAS-ISSMax, EAS-BSam, BSam-SSMAX, and EAS-BSam-SSAD, three datasets were used for experiments.

A. HYDICE 15-Panel Scene

An airborne hyperspectral digital imagery collection experiment (HYDICE) scene is shown in Fig. 5. It was collected in August 1995 from a flight altitude of 10 000 ft with a ground sampling distance of approximately 1.56 m. This scene has been studied extensively by many reports such as [1], [2], and [3]. It has a total of 169 bands used for the experiments with low-signal/high-noise bands: bands 1–3 and bands 202–210 and water vapor absorption bands: bands 101–112 and bands 137–153 removed. There are 15 square panels in Fig. 5(a) with three different sizes, 3×3 , 2×2 , and 1×1 m, respectively.

Due to the ground sampling distance of approximately 1.56 m, the each of panels in the first column except the first row contains two panel pixels highlighted by red, p_{211} and p_{221} in row 2, p_{311} and p_{312} in row 3, p_{411} and p_{412} in row 4, and p_{511} and p_{521} in row 5, as shown in Fig. 5. All the remaining 11 panels in Fig. 5 contain one single panel pixel for each

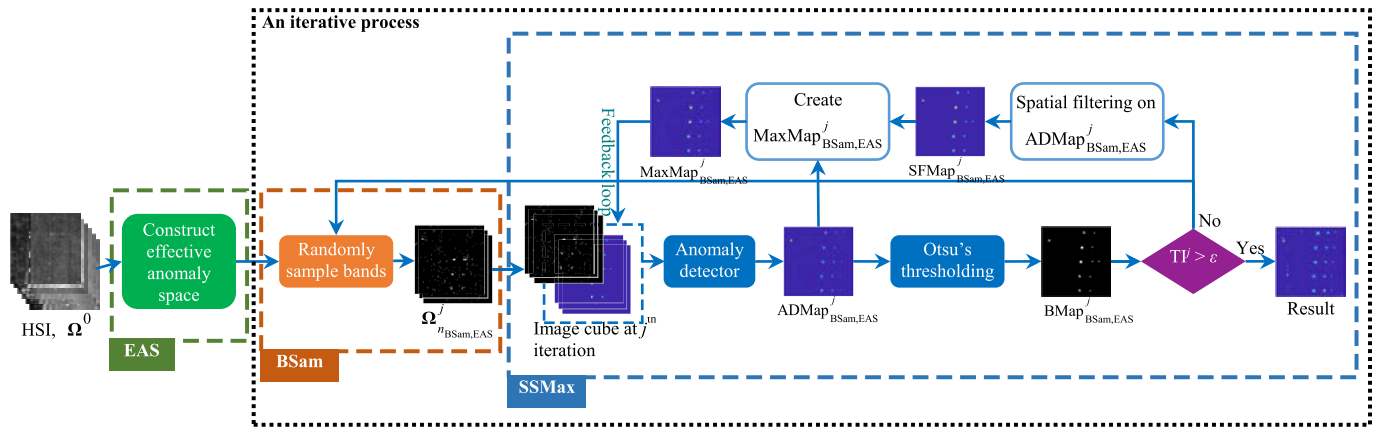


Fig. 4. Graphic block diagram of implementing Algorithm 7.

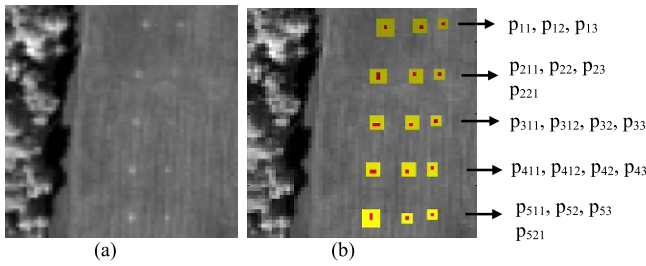


Fig. 5. (a) A HYDICE15-panel scene that contains 15 panels. (b) Ground-truth map of spatial locations of 19 red panel pixels.

panel also highlighted by red, p_{11} , p_{12} , and p_{13} in row 1, p_{22} and p_{23} in row 2, p_{32} and p_{33} in row 3, p_{42} and p_{43} in row 4, and p_{52} and p_{53} in row 5. Therefore, there are a total of 19 red panel pixels. Fig. 5(b) shows their precise spatial locations with the pixels in yellow indicating panel pixels mixed with BKG.

B. HYDICE Urban Scene

A second dataset is another HYDICE image with pseudocolor shown in Fig. 6(a), which is available and can be downloaded from the Remote Sensing Lab website.

It is an urban scene and comprises 210 spectral bands with 174 bands being used for experiments after the noise and water absorption bands had been removed. A region with a size of 80×100 pixels located at the upper right of the scene was selected as the test image shown in Fig. 6(b) along with the ground-truth map shown in Fig. 6(c), where 21 pixels were identified as anomalies that were cars and roofs because they had spectra that differ from BKG.

C. San Diego Airport Data

A third dataset used for experiments was also an airborne visible/infrared imaging spectrometer (AVIRIS) acquired scene, which is a San Diego Airport scene in California, USA, with pseudocolor shown in Fig. 7(a). It is an urban scene in which the main BKG materials are roof, shadow, and grass. It has a size of 400×400 pixels with a 3.5-m spatial resolution and 224 spectral channels in wavelengths ranging from 370 to 2510 nm. After removing the bands

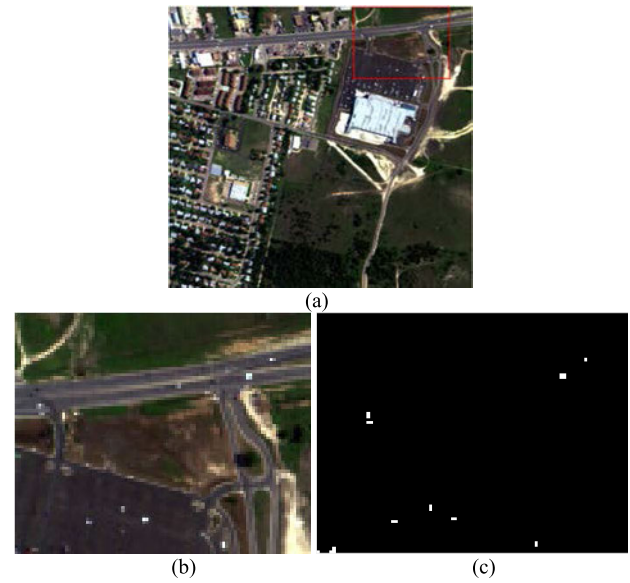


Fig. 6. HYDICE urban scene. (a) Pseudocolor image of the whole scene. (b) Pseudocolor image of the selected area. (c) Ground-truth map.

that correspond to the water absorption regions, low signal-to-noise ratio (SNR), and bad bands (1–6, 33–35, 97, 107–113, 153–166, and 221–224), 189 available bands of the data were retained in the experiments.

An area of 100×100 pixels at the upper left corner of the scene was selected as the test image shown in Fig. 7(b) with its ground-truth map shown in Fig. 7(c). This data subspace has been widely used to evaluate AD performance where the three airplanes located at the upper left corner are considered to be of major interest, consist of 85 pixels, and account for 0.33% of the test image, as shown in Fig. 7(b). This scene is less interesting than the HYDICE 15-panel data in Fig. 5 because the only ground truth available is three airplanes that are hardly considered anomalies according to the characteristics of anomalies due to their large sizes and visibility. As a result, this scene is more appropriate for TD rather than AD. As a matter of fact, the experiments to be conducted in Section V confirm this assertion, and the conclusion drawn

Algorithm 7 EAS-BSam-SSMax

Let Ω^0 be the full band set and EAS be constructed from the original scene represented by Ω^0 and let it be $\mathbf{X}^{\text{EAS},0}$. At the j^{th} iteration

1. 1st step: input $\mathbf{X}^{\text{EAS},j}$.
2. 2nd step: randomly select $\Omega_{\text{BSam,EAS}}^j$ from $\mathbf{X}^{\text{EAS},j}$ to form $\mathbf{X}_{\Omega_{\text{BSam,EAS}}^j}^j$.
3. 3rd step: implement a selected anomaly detector on $\mathbf{X}_{\Omega_{\text{BSam,EAS}}^j}^j$ to produce an $\text{ADMap}_{\text{BSam,EAS}}^j$.
4. 4th step: apply Otsu's method to $\text{ADMap}_{\text{BSam,EAS}}^j$ to produce a binary map, $\text{BMap}_{\text{BSam,EAS}}^j$.
5. 5th step: check the following stopping rule

$$\text{TI}^{(j)} = \frac{|\text{BMap}_{\text{BSam,EAS}}^j \cap \text{BMap}_{\text{BSam,EAS}}^{j-1}|}{|\text{BMap}_{\text{BSam,EAS}}^j \cup \text{BMap}_{\text{BSam,EAS}}^{j-1}|} > \varepsilon. \quad (15)$$

If (15) is true, then the iterative process is terminated. Otherwise, continue.

6. 6th step: apply a spatial filter (SF) specified by an edge preserving filter, EPF-G-g to the $\text{ADMap}_{\text{BSam,EAS}}^j$ obtained in step 2 to produce a $\text{SFMap}_{\text{BSam,EAS}}^j$, and then use $\text{SFMap}_{\text{BSam,EAS}}^j$ and $\text{ADMap}_{\text{BSam,EAS}}^j$ to produce a $\text{MaxMap}_{\text{BSam,EAS}}^j$ by

$$\begin{aligned} \text{MaxMap}_{\text{BSam,EAS}}^j &= \max\{\text{SFMap}_{\text{BSam,EAS}}^j, \text{ADMap}_{\text{BSam,EAS}}^j\}. \end{aligned} \quad (16)$$

7. 7th step: randomly select n_{BSam} bands, $\Omega_{\text{BSam,EAS}}^{j+1}$ from $\mathbf{X}^{\text{EAS},j}$ and feed $\text{MaxMap}_{\text{BSam,EAS}}^j$ obtained in step 6 to create a new data cube $\mathbf{X}_{\Omega_{\text{BSam,EAS}}^{j+1}}^{j+1}$ defined as

$$\begin{aligned} \mathbf{X}_{\Omega_{\text{BSam,EAS}}^{j+1}}^{j+1} &= \Omega_{\text{BSam,EAS}}^{j+1} \cup \text{MaxMap}_{\text{BSam,EAS}}^1 \cup \dots \cup \text{MaxMap}_{\text{BSam,EAS}}^j. \end{aligned} \quad (17)$$

8. 8th step: Let $j + 1 \leftarrow j$ and repeat steps 2-7.

from this scene is indeed quite different from that drawn from Figs. 5 and 6.

D. AVIRIS Bay Champagne Beach Scene

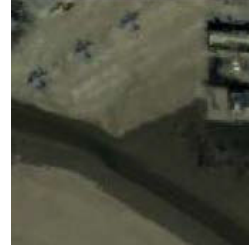
A fourth dataset is another AVIRIS scene, which is the Bay Champagne Beach. The data was acquired on July 4, 2010, and has a size of 100×100 pixels. After removing the noisy bands from the original dataset, 188 bands remained. The spatial resolution is approximately 4.4 m. Fig. 8(a) and (b) shows the false-color map and ground-truth map of the Bay Champagne Beach scene, respectively.

VI. EXPERIMENTAL RESULTS AND DISCUSSIONS

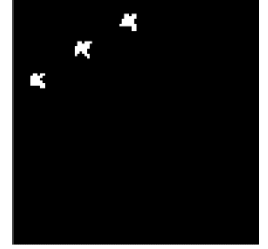
This section presents a very detailed comprehensive study of experiments and discussions where RXAD, EAS-RXAD,



(a)

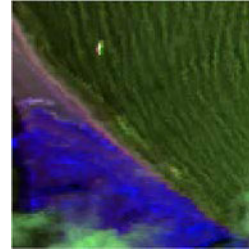


(b)



(c)

Fig. 7. San Diego Airport scene. (a) Pseudocolor image of the whole scene. (b) Pseudocolor image of the selected area. (c) Ground-truth map.



(a)



(b)

Fig. 8. AVIRIS Bay Champagne Beach scene. (a) False-color image. (b) Ground-truth map.

BSam-RXAD, ISSMax, EAS-ISSMax, BSam-SSMax, EAS-BSam, and EAS-BSam-SSMax presented in Figs. 1–4 were used for experiments. It should be noted that all the used anomaly detectors were RXAD, i.e., the Mahalanobis distance. To evaluate their performance, a comparative analysis was conducted against RXAD and the four most recent state-of-the-art AD methods, LRR, LSDM-MoG, PTA, and RGAE, on the four real hyperspectral data scenes described in Section V. These four AD methods were selected and compared to evaluate performance because their source codes are available on website, and each of them represents one particular category. For example, an LRR represents a category of LRA-SR methods. It separates BKG and anomalies by approximating and representing BKG via a dictionary with its representation coefficients of all pixels to form a low-rank matrix subject to a sparsity-inducing regularization term to constrain the representation coefficients where anomaly tensor was used for AD by the Euclidean distance (or l_2 -norm). An LSDM-MoG proposed in [36] represents a category of using Gaussian distributions for BKG estimation where AD was performed by the Manhattan distance in a BKG-separated space. A PTA proposed in [32] represents a category of tensor decomposition to decompose an HSI data cube into BKG and anomaly tensors where the anomaly tensor was used for AD

TABLE III
PARAMETERS SETTING FOR TEST AD METHODS

n_{BSam}	HYDICE panel	HYDICE urban	San Diego airport	Bay Champagne Beach
	13	13	11	15
ISSMax	EPF-G-g [46]			
LRR [8]	$\beta = 0.1$			
LDSM-MoG [36]	$(l_0, K) = (20, 2)$ $(l_0, \mu_0, \sigma_0, b_0, c_0, d_0, \alpha_0, \beta_0) = (10^{-3}, 0, 10^{-6}, 10^{-6}, 10^{-6}, 10^{-6}, 10^{-6}, 10^{-6})$			
PTA [32]	$(\gamma, \alpha, \mu, \beta, \tau, \maxIter) = (1, 1.7, 0.069, 0.1, 400)$			
RGAE [29]	$(\lambda, S, d) = (10^{-2}, 150, 100)$			
Max number of iterations	20			

TABLE IV
CAPABILITY OF TEST AD METHODS IN ADDRESSING FOUR ISSUES:
BS, NOISE EFFECT, ANOMALY DETECTABILITY, AND
INTERBAND CORRELATION

	BS	anomaly detectability	noise effect	inter-band correlation
RXAD [5]		✓		
LRR [8]	✓	✓		
LSDM-MoG [36]	✓		✓	
PTA [32]		✓		
RGAE [29]	✓			
ISSHAD [7]	✓	✓		
EAS [14]	✓	✓	✓	
BSam [15]			✓	✓
ISSMax	✓	✓		
EAS-ISSMax	✓	✓	✓	
EAS-BSam	✓	✓		✓
BSam-SSMax	✓		✓	✓
EAS-BSam-SSMax	✓	✓	✓	✓

by the Euclidean distance (or l_2 -norm). An RGAE proposed in [29] represents a category of graph-based AD methods. It is a graph-based AD where the BKG is modeled with Gaussian multivariate distribution assumption, and anomalies are detected based on the Euclidean distance or the Mahalanobis distance. Since these AD methods used various measures, such as Euclidean distance and the Manhattan distance to perform AD, we replaced these measures with RXAD for consistency when we implemented all their corresponding SSMax, EAS, and BSam. Also, RPCA, LRSMD, and CDA have been shown in [14] not to perform, as well as EAS, and they were not included in the comparison.

The experiments were performed under MATLAB R2022b with macOS Ventura version 13.4, CPU chip Apple M1 Pro, eight cores, and 16-GB memory. Table III shows the parameter setting for the test AD methods including the number of bands to be sampled, n_{BSam} , which is determined by the Harsanyi-Farrend-Chang (HFC) method [42] for VD with the false alarm probability set to 10^{-4} [43], [44], [45]. Table IV tabulates various most recent particularly developed AD methods to deal with the four issues addressed in the introduction where check symbols “✓” in red and black indicate a strong “yes” and a weak “yes” in addressing a particular issue, respectively.

A. Quantitative Detection Measures

It should be noted that the detection measures [15] were derived from the signal detection noise model for TD where

the hypothesis H_0 was assumed to be noise. However, in AD, BKG is a crucial factor in determining detection performance. To address this issue, Chang [17] reformulated the binary hypothesis testing problem as an AD in BKG by replacing noise with BKG in hypothesis H_0 . Thus, to quantitatively evaluate the relative performance among all the AD methods in Table IV, the eight detection measures proposed in [17] were used for this purpose.

1) *Effectiveness of a Detector*: This is the standard 2-D ROC curve of (P_D, P_F) generated by the same threshold τ , $AUC_{(D,F)}$, given by

$$0 \leq AUC_{(D,F)} \leq 1. \quad (18)$$

As a result, P_D and P_F cannot work alone such as anomaly detectability and BKG suppressibility and can be only used to evaluate the effectiveness of a detector.

2) *Anomaly Detection Probability*: The value of $AUC_{(D,\tau)}$ calculated from the 2-D ROC curve of (P_D, τ) , denoted by AUC_{ADP} , is given by

$$0 \leq AUC_{ADP} \leq 1 \quad (19)$$

which can be used to evaluate anomaly detectability. By factoring (18) into (19), we can define a new measure, called JAD, AUC_{JAD} , given by

$$0 \leq AUC_{JAD} = AUC_{(D,F)} + AUC_{ADP} \leq 2. \quad (20)$$

3) *Background Detection Probability*: In signal detection of noise, P_F is the false alarm probability and $1 - P_F$ is the detection of noise. Thus, using $1 - P_F$ to evaluate noise detection performance for signal detection does not make sense. However, this is not true from AD where $1 - P_F$ is actually the detection probability of BKG. Thus, we can define BDP, AUC_{BDP} , as a counterpart of ADP in (19) by

$$0 \leq AUC_{BDP} = 1 - AUC_{(F,\tau)} \leq 1 \quad (21)$$

and JBS, AUC_{JBS} , given by

$$0 \leq AUC_{JBS} = AUC_{(D,F)} + AUC_{BDP} \leq 2. \quad (22)$$

4) *JAD With BKG Suppressibility*: As pointed out in $AUC_{(D,F)}$, P_D and P_F are calculated by the same threshold τ for their joint performance without accounting for ADP and BDP separately. There are three ways to factor $AUC_{(F,\tau)}$ into $AUC_{(D,\tau)}$ to evaluate JAD with the BKG suppressibility of a detector.

a) *AD in BKG (ADBS)*: Due to the fact that P_F is caused by the probability of misdetecting BKG as anomalies, this error probability should be subtracted from ADP. To take care of this effect, a measure, factoring P_F into ADP, to be called ADBS, can be defined as AUC_{ADBS} by

$$-1 \leq AUC_{ADBS} = AUC_{(D,\tau)} - AUC_{(F,\tau)} \leq 1. \quad (23)$$

b) *Overall AD power*: Another important measure is derived from a classification point of view. It is called OADP defined by

$$0 \leq AUC_{OADP} = AUC_{ADP} + AUC_{BDP} \leq 2 \quad (24)$$

where AUC_{ADP} is given in (19) and AUC_{BDP} is given in (18).

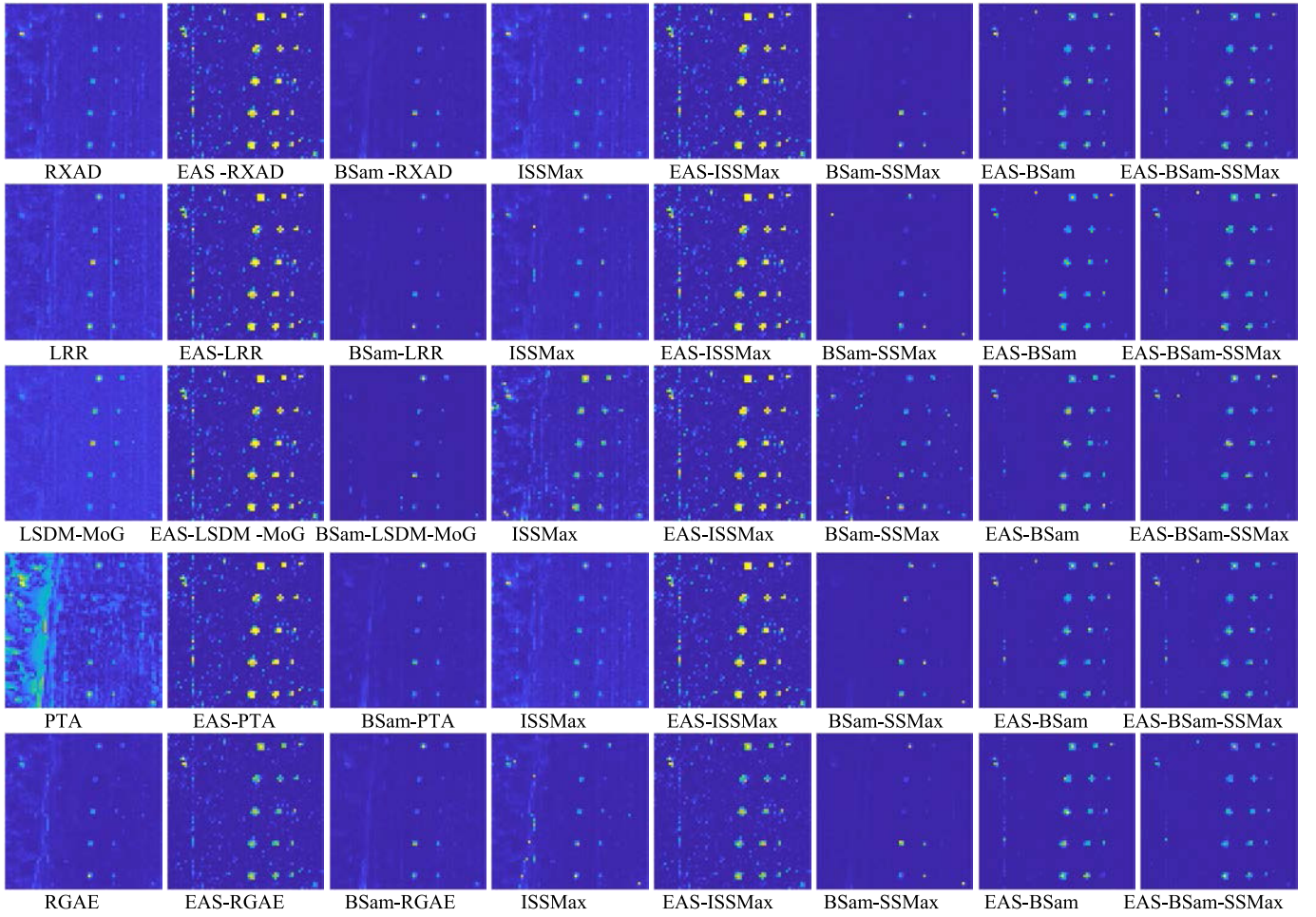


Fig. 9. ADMaps produced by RXAD, LRR, LSDM-MoG, PTA, and RGAE along with their corresponding EAS, BSam, ISSMax, EAS-BSam, and EAS-BSam-SSMax counterparts using RXAD as the anomaly detector for HYDICE panel scene.

TABLE V

EIGHT DETECTION MEASURES IN ADDRESSING FOUR ISSUES:
BS, ANOMALY DETECTABILITY, NOISE EFFECT, AND IBCR

detection measure	BS	anomaly detectability	noise effect	inter-band correlation
$AUC_{(D,F)}$	✓	✓		
AUC_{ADP}		✓		
$=AUC_{(D,F)}$				
AUC_{BDP}			✓	
$=1-AUC_{(F,F)}$				
AUC_{JAD}		✓		✓
AUC_{JBS}	✓			✓
AUC_{ADBS}	✓	✓		
AUC_{OADP}	✓	✓	✓	
AUC_{SBPR}	✓	✓	✓	✓

c) *Signal-to-BKG probability ratio*: A third one is derived from an idea similar to SNR. It is called SBPR that was also derived in [17] by

$$0 \leq AUC_{SBPR} = \frac{AUC_{ADP}}{AUC_{BDP}} \quad (25)$$

where AUC_{ADP} and AUC_{BDP} defined in (19) and (21) can be assumed to be anomaly detectability and BKG detectability.

Table V tabulates the eight detection measures to respond to the four major AD issues noted in the introduction where check symbols “✓” in red and black indicate a strong “yes” and a weak “yes” in addressing a particular issue, respectively.

B. HYDICE 15-Panel Scene

Fig. 9 shows the ADMaps of RXAD, LRR, LSDM-MoG, PTA, and RGAE along with their corresponding counterparts, EAS, BSam, ISSMax, EAS-ISSMax, BSam-SSMax, EAS-BSam, and EAS-BSam-SSMax for the HYDICE panel scene. By visual inspection, it was obvious that EAS did improve RXAD, LRR, LSDM-MoG, and PTA as expected by detecting all 25 panels as anomalies. Interestingly, BSam cleaned up most BKG at the expense of weakening the ability of RXAD, LRR, LSDM-MoG, PTA, RGAE, and ISSMax in detecting anomalies. Most surprisingly, if BSam coupled with EAS alone, the AD performances of RXAD, LRR, LSDM-MoG, and PTA were improved significantly without appealing for SSMax. Finally, if all three, EAS, BSam, and SSMax, were implemented jointly, it clearly showed that EAS-BSam-SSMax produced the best results. In other words, EAS-ISSMax has the best AD results, while EAS-BSam and EAS-BSam-SSMax give better BS results and retain the ability to detect the subpixel panels. Interestingly, as demonstrated in Fig. 9, BSam

might weaken the ability of anomaly detectors. However, by integrating EAS into BSAM-SSMax, anomaly detectors regained their detectability and better BS.

Fig. 10 plots 3-D ROC curves of (P_D, P_F, τ) in Fig. 9 and their corresponding three 2-D ROC curves of (P_D, P_F) , (P_D, τ) , and (P_F, τ) for RXAD, LRR, LSFDM-MoG, PTA, and RGAE detectors along with their EAS, BSAM, ISSMax, EAS-BSAM, and EAS-BSAM-SSMax counterparts.

The ROC curve plots in Fig. 10 can be assessed for comparison by visual inspection. In particular, the profiles of 3-D ROC curves of 12 test anomaly detectors exhibited different behaviors. Unfortunately, their AUCs are real-valued and cannot be used for quantitative analysis. In this case, the eight AUC values, $AUC_{(D,F)}$, AUC_{ADP} , AUC_{BDP} , AUC_{JAD} , AUC_{JBS} , AUC_{ADBS} , AUC_{SBPR} , and AUC_{OADP} , can be calculated and tabulated in Table VI for quantitative analysis where the best results are boldfaced by red. In Table VI, the computing times and the number of iterations, j , along with $TI_{j,j-1}$ values are also documented. According to Table IV, EAS-ISSMax produced the best anomaly detectability regardless of which anomaly detector was used. Furthermore, EAS-BSAM had the best BS performance while maintaining decent AD performance; specifically, the values of AUC_{SBPR} by EAS-BSAM and EAS-BSAM-SSMax were higher than that by ISSMax and EAS-ISSMax. All these results were also confirmed by visual inspection in Fig. 9.

C. HYDICE Urban Scene

Fig. 11 shows the ADMaps of RXAD, LRR, LSDM-MoG, PTA, and RGAE along with their corresponding counterparts, EAS, BSAM, ISSMax, EAS-ISSMax, BSAM-SSMax, EAS-BSAM, and EAS-BSAM-SSMax for the HYDICE urban scene. From Fig. 6, there are only 21 anomalous pixels in the urban scene, which makes it difficult to evaluate the performance based on visual inspection of Fig. 11.

Fig. 12 plots 3-D ROC curves of (P_D, P_F, τ) in Fig. 11 and their corresponding three 2-D ROC curves of (P_D, P_F) , (P_D, τ) , and (P_F, τ) for RXAD, LRR, LSFDM-MoG, PTA, and RGAE detectors along with their EAS, BSAM, ISSMax, EAS-BSAM, and EAS-BSAM-SSMax counterparts. Despite that the 3-D ROC curve plots in Fig. 11 provided visual qualitative analysis, their AUC values are real-valued and cannot be used for quantitative comparison. In this case, the eight AUC values, $AUC_{(D,F)}$, AUC_{ADP} , AUC_{BDP} , AUC_{JAD} , AUC_{JBS} , AUC_{ADBS} , AUC_{SBPR} , and AUC_{OADP} , can be calculated and tabulated in Table VII where the best results are boldfaced by red. According to Table VII, EAS-ISSMax yielded the best result regardless of which anomaly detector was used. In addition, EAS-BSAM and EAS-BSAM-SSMax had the best BS. Specifically, EAS-BSAM and EAS-BSAM-SSMax produced much higher values of AUC_{SBPR} than that produced by ISSMax and EAS-ISSMax. This indicated that by compressing BKG and noise from the image, anomalies could be brought to light from the BKG and noise.

D. AVIRIS San Diego Airport Scene

As noted earlier and also shown in [14], the three airplanes in the scene are so visible and hardly considered anomalies due

to their large sizes. As a result, this scene is more suitable for TD than for AD. Consequently, the conclusions drawn from the scene are expected to be different from the previous two scenes.

Fig. 13 shows the ADMaps of RXAD, LRR, LSDM-MoG, PTA, and RGAE along with their corresponding counterparts, EAS, BSAM, ISSMax, EAS-ISSMax, BSAM-SSMax, EAS-BSAM, and EAS-BSAM-SSMax for the AVIRIS San Diego Airport scene. By visual inspection, EAS-ISSMax produced the best results in detecting the three aircraft while improving BS compared to their original AD without it. Interestingly, the three aircrafts were not visible in BSAM-SSMax, EAS-BSAM, and EAS-BSAM-SSMax. This indicated that these methods did not treat the airplanes as anomalies.

To evaluate ADMaps in Fig. 13 quantitatively, Fig. 14 plots 3-D ROC curves of (P_D, P_F, τ) in Fig. 13 and their corresponding three 2-D ROC curves of (P_D, P_F) , (P_D, τ) , and (P_F, τ) for RXAD, LRR, LSFDM-MoG, PTA, and RGAE detectors along with their EAS, BSAM, ISSMax, EAS-BSAM, and EAS-BSAM-SSMax counterparts. Interestingly, compared with the 3-D ROC curves in Figs. 10 and 12, the 3-D ROC curves in Fig. 14 exhibited quite different behaviors. To further analyze 3-D/2-D ROC curves quantitatively, the eight AUC values, $AUC_{(D,F)}$, AUC_{ADP} , AUC_{BDP} , AUC_{JAD} , AUC_{JBS} , AUC_{ADBS} , AUC_{SBPR} , and AUC_{OADP} , can be calculated and tabulated in Table VIII where the best results are boldfaced by red. Also included in Table VIII are their computing times and the number of iterations, j , along with $TI_{j,j-1}$ values also documented. Interestingly, based on the results in Table VIII, the airplanes were not considered anomalies since their detection rates, AUC_{ADP} , and AUC_{ADBS} were very low and BS was very poor.

E. AVIRIS Bay Champagne Beach Scene

Fig. 15 shows the ADMaps of RXAD, LRR, LSDM-MoG, PTA, and RGAE along with their corresponding counterparts, EAS, BSAM, ISSMax, EAS-ISSMax, BSAM-SSMax, EAS-BSAM, and EAS-BSAM-SSMax for the AVIRIS Bay Champagne Beach scene. By visual examination, it is evident that EAS, EAS-ISSMax, EAS-BSAM, and EAS-BSAM-SSMax enhanced the performance of RXAD, LRR, LSDM-MoG, PTA, and RGAE. Moreover, EAS-BSAM and EAS-BSAM-SSMax exhibited superior BS compared to other methods while still retaining their ability to detect anomalous pixels.

To evaluate ADMaps in Fig. 15 quantitatively, Fig. 16 plots 3-D ROC curves of (P_D, P_F, τ) in Fig. 15 and their corresponding three 2-D ROC curves of (P_D, P_F) , (P_D, τ) , and (P_F, τ) for RXAD, LRR, LSFDM-MoG, PTA, and RGAE detectors along with their EAS, BSAM, ISSMax, EAS-BSAM, and EAS-BSAM-SSMax counterparts.

By examining the 3-D/2-D ROC curves Figs. 10, 12, and 14, the 3-D/2-D ROC curves in Fig. 16 behaved more like those in Figs. 10 and 12 but quite different from that in Fig. 14. These facts further confirmed that the targets in the HYDICE panel scene, the HYDICE urban scene, and the Bay Champaign beach scene are indeed anomalies compared to three airplanes in the San Diego Airport scene, which are, in fact, visible large targets and are not supposed to be anomalies. For quantitative analysis, the eight AUC values, $AUC_{(D,F)}$, AUC_{ADP} , AUC_{BDP} ,

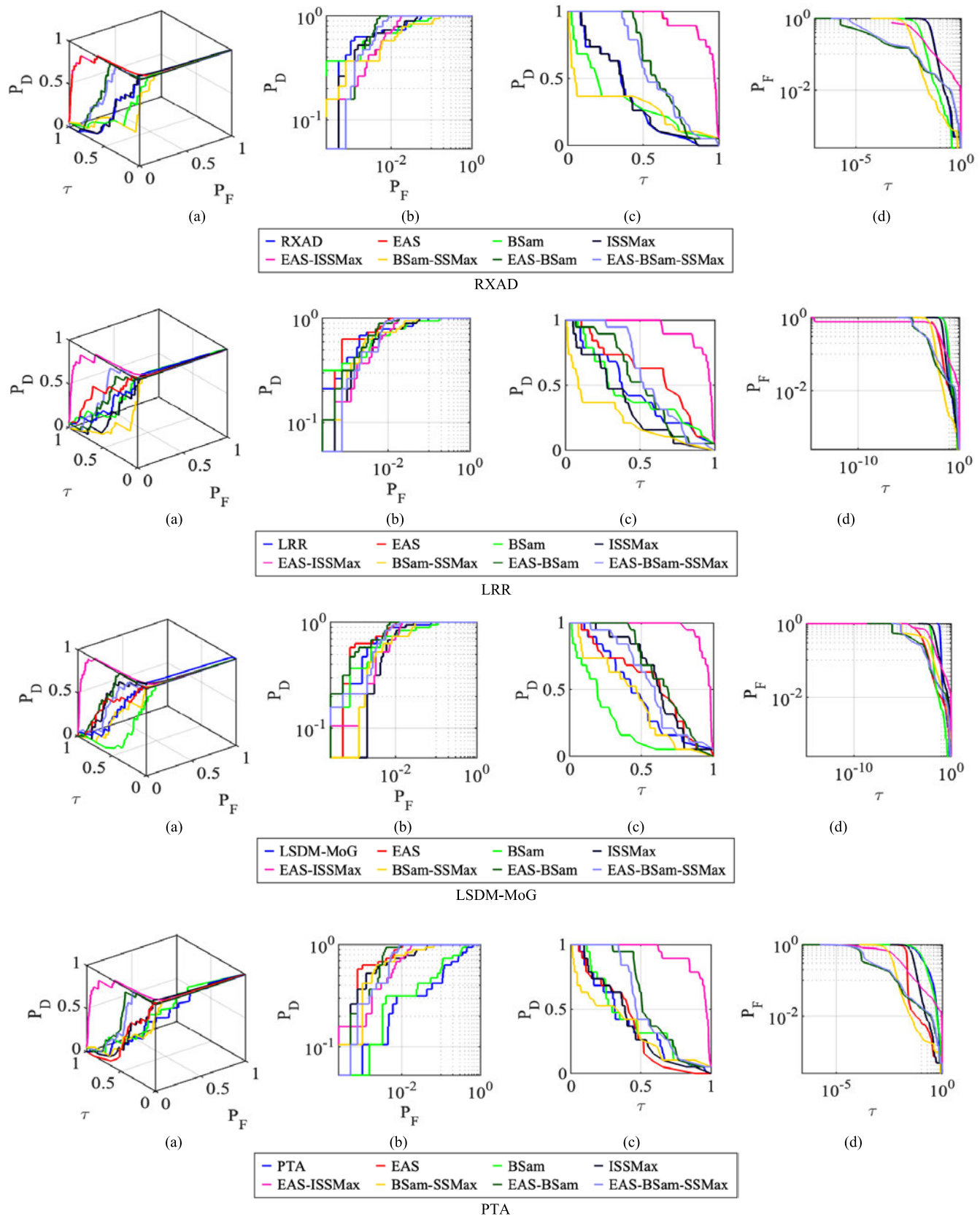


Fig. 10. 3-D ROC curves along with their corresponding 2-D ROC curves generated from Fig. 9 for the HYDICE panel scene. (a) 3-D ROC curves. (b) (P_D, P_F) curves. (c) (P_D, τ) curves. (d) (P_F, τ) curves.

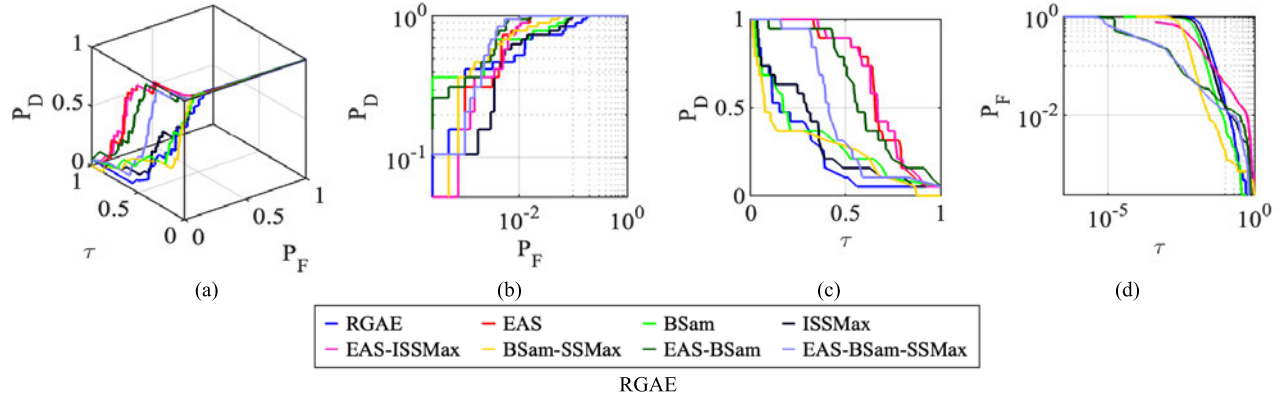


Fig. 10. (Continued.) 3-D ROC curves along with their corresponding 2-D ROC curves generated from Fig. 9 for the HYDICE panel scene. (a) 3-D ROC curves. (b) (P_D, P_F) curves. (c) (P_D, τ) curves. (d) (P_F, τ) curves.

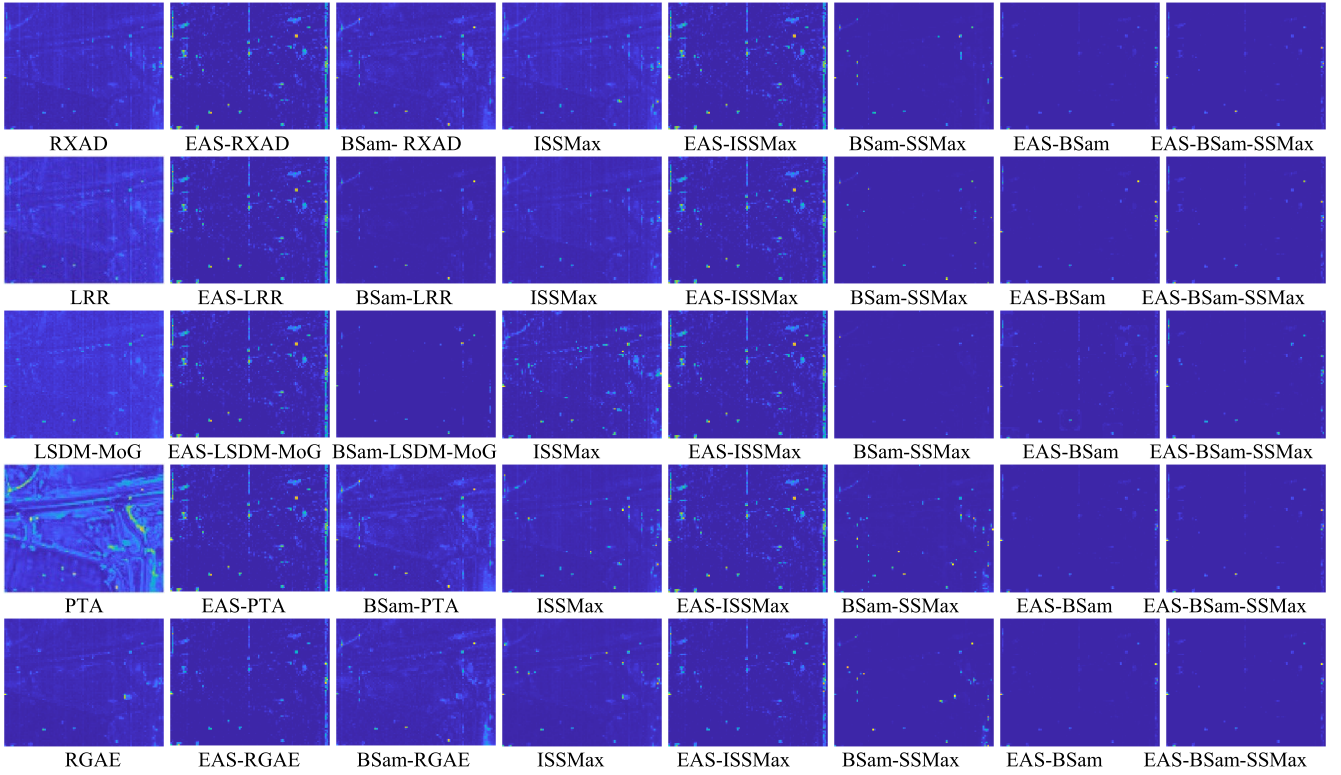


Fig. 11. ADMaps produced by RXAD, LRR, LSDM-MoG, PTA, and RGAE along with their corresponding EAS, BSAm, ISSMax, EAS-BSAm, and EAS-BSAm-SSMax counterparts using RXAD as the anomaly detector for the HYDICE urban scene.

AUC_{JAD} , AUC_{JBS} , AUC_{ADBS} , AUC_{SBPR} , and AUC_{OADP} , can be calculated and tabulated in Table VII where the best results are boldfaced by red.

Also, in Table VII, the computation times and the number of iterations, j , along with $TI_{j,j-1}$ values, are also documented. According to Table IX, (EAS-RXAD, EAS-BSAm-RXAD), (EAS-ISSMax-LRR, EAS-BSAm-SSMax-LRR), (EAS-ISSMax-LSDM-MoG, EAS-BSAm-LSDM-MoG), (EAS-PTA, EAS-BSAm-PTA), (EAS-RGAE, EAS-BSAm-SSMax-RGAE) were among the best for a given anomaly detector. Furthermore, EAS-BSAm and EAS-BSAm-SSMax exhibited the best BS while maintaining decent AD performance. Specifically, the values of AUC_{SBPR} for EAS-BSAm and

EAS-BSAm-SSMax were higher than other methods, which were confirmed by Fig. 15.

F. Time Complexity Analysis

From Tables VI to IX, the computing time is completely determined by the AD method and ISSMax, EAS, and BSAm. Since ISSMax and BSAm implemented iterative feedback processes, their computational complexities are nearly the same with their most computing times determined by how many iterations are required to terminate the iterative processes. As also shown in Tables VI–VIII, LRR required the most computing time followed by PTA and RGAE whose computing times

TABLE VI

AUC VALUES PRODUCED BY RXAD, LRR, LSDM-MOG, PTA, AND RGAE ALONG WITH THEIR CORRESPONDING EAS, BSAM, ISSMAX, EAS-BSAM, AND EAS-BSAM-SSMAX COUNTERPARTS USING RXAD AS THE ANOMALY DETECTOR FOR THE HYDICE PANEL SCENE

Anomaly Detector	j	$Tl_{j,j-1}$	AUC _(D,F)	AUC _{ADP}	AUC _{BDP}	AUC _{JAD}	AUC _{JBS}	AUC _{ADBS}	AUC _{SBPR}	AUC _{OADP}	time(s)
RXAD	--	--	0.99	0.3574	0.9573	1.3474	1.9473	0.3147	0.3733	1.3147	0.09
EAS-RXAD	--	--	0.9932	0.9227	0.9627	1.9159	1.9559	0.8854	0.9585	1.8854	0.07
BSam-RXAD	--	--	0.9839	0.2816	0.9862	1.2655	1.9701	0.2678	0.2855	1.2678	0.01
ISSMax-RXAD	1	0.976	0.9918	0.3618	0.9584	1.3537	1.9502	0.3202	0.3775	1.3202	2.82
EAS-ISSMax-RXAD	1	0.992	0.9932	0.9233	0.9625	1.9165	1.9557	0.8858	0.9593	1.8858	2.65
BSam-SSMax-RXAD	6	1	0.9727	0.2383	0.9949	1.2109	1.9675	0.2331	0.2395	1.2331	7.88
EAS-BSam-RXAD	12	0.954	0.9974	0.5932	0.9919	1.5906	1.9893	0.5851	0.5980	1.5851	14.02
EAS-BSam-SSMax-RXAD	12	0.969	0.997	0.5911	0.9916	1.5881	1.9886	0.5827	0.5961	1.5827	14.43
LRR	--	--	0.9877	0.429	0.9481	1.4167	1.9358	0.377	0.4525	1.377	73.11
EAS-LRR	--	--	0.9933	0.9183	0.9628	1.9116	1.9561	0.8811	0.9538	1.8811	1.45
BSam-LRR	--	--	0.9763	0.1947	0.9939	1.171	1.9702	0.1886	0.1959	1.1886	20.09
ISSMax-LRR	6	0.951	0.9897	0.3463	0.9732	1.3359	1.9629	0.3195	0.3558	1.3195	164.20
EAS-ISSMax-LRR	1	0.976	0.9935	0.9212	0.9627	1.9147	1.9562	0.8839	0.9569	1.8839	8.12
BSam-SSMax-LRR	5	1	0.9918	0.2308	0.9965	1.2226	1.9883	0.2273	0.2316	1.2273	131.14
EAS-BSam-LRR	14	0.957	0.996	0.4857	0.9937	1.4818	1.9897	0.4794	0.4888	1.4794	43.51
EAS-BSam-SSMax-LRR	19	0.961	0.9957	0.5809	0.9912	1.5767	1.987	0.5722	0.5861	1.5722	63.87
LSDM-MoG	--	--	0.9928	0.4351	0.9213	1.4279	1.9141	0.3565	0.4723	1.3565	9.47
EAS-LSDM-MoG	--	--	0.9932	0.928	0.9627	1.9213	1.9559	0.8907	0.9640	1.8907	9.42
BSam-LSDM-MoG	--	--	0.975	0.2414	0.9959	1.2164	1.9709	0.2373	0.2424	1.2373	0.47
ISSMax-LSDM-MoG	6	0.965	0.9926	0.5909	0.9618	1.5835	1.9544	0.5527	0.6144	1.5527	39.85
EAS-ISSMax-LSDM-MoG	2	0.97	0.9936	0.9496	0.9624	1.9431	1.9559	0.9119	0.9867	1.9119	28.30
BSam-SSMax-LSDM-MoG	16	0.981	0.9911	0.4054	0.9899	1.3966	1.9811	0.3953	0.4095	1.3953	52.45
EAS-BSam-LSDM-MoG	17	0.957	0.9978	0.6607	0.9921	1.6586	1.99	0.6528	0.6660	1.6528	59.99
EAS-BSam-SSMax-LSDM-MoG	16	0.962	0.9961	0.5485	0.9929	1.5446	1.989	0.5414	0.5524	1.5414	71.00
PTA	--	--	0.8315	0.3894	0.8632	1.221	1.6948	0.2527	0.4511	1.2527	48.93
EAS-PTA	--	--	0.9933	0.9177	0.9627	1.911	1.956	0.8804	0.9533	1.8804	35.04
BSam-PTA	--	--	0.9834	0.2386	0.9861	1.222	1.9695	0.2247	0.2420	1.2247	3.09
ISSMax-PTA	1	1	0.9915	0.3862	0.9548	1.3778	1.9464	0.3411	0.4045	1.3411	71.41
EAS-ISSMax-PTA	1	0.969	0.9933	0.9182	0.9625	1.9115	1.9558	0.8806	0.9540	1.8806	71.09
BSam-SSMax-PTA	8	1	0.9907	0.3012	0.9941	1.2919	1.9849	0.2953	0.3030	1.2953	64.57
EAS-BSam-PTA	12	0.952	0.9973	0.5605	0.9922	1.5578	1.9895	0.5526	0.5649	1.5526	102.62
EAS-BSam-SSMax-PTA	12	0.985	0.996	0.526	0.9913	1.522	1.9873	0.5173	0.5306	1.5173	101.63
RGAE	--	--	0.9643	0.2078	0.9748	1.1721	1.939	0.1826	0.2132	1.1826	43.89
EAS-RGAE	--	--	0.9945	0.654	0.9764	1.6486	1.971	0.6305	0.6698	1.6305	40.95
BSam-RGAE	--	--	0.9835	0.2695	0.9861	1.2531	1.9697	0.2557	0.2733	1.2557	14.02
ISSMax-RGAE	16	0.976	0.9775	0.2778	0.9787	1.2553	1.9562	0.2566	0.2838	1.2566	784.36
EAS-ISSMax-RGAE	1	0.966	0.9944	0.6619	0.9765	1.6563	1.9709	0.6384	0.6778	1.6384	83.55
BSam-SSMax-RGAE	4	1	0.9916	0.2641	0.9945	1.2557	1.9861	0.2586	0.2656	1.2586	76.81
EAS-BSam-RGAE	20	0.937	0.9966	0.5894	0.9925	1.586	1.9891	0.582	0.5939	1.582	360.09
EAS-BSam-SSMax-RGAE	16	0.966	0.997	0.4359	0.9946	1.4329	1.9917	0.4305	0.4383	1.4305	284.06

were nearly at the same order with RXAD requiring the least computing time.

VII. FURTHER DISCUSSIONS ON EXPERIMENTAL RESULTS

A. HYDICE Panel and Urban Scenes

As for RXAD, it clearly showed that all the four EAS-ISSMax, BSam-SSMax, EAS-BSam, and EAS-BSam-SSMax

performed better than RXAD and ISSMax alone where EAS-ISSMax produced the best anomaly detectability at the expense of P_F . On the contrary, BSam-SSMax produced the best BS at the expense of anomaly detectability. However, if we compared EAS-ISSMax to ISSMax, EAS-SSMax increased anomaly detectability significantly but also increased P_F . Thus, including BSam could alleviate both issues by reducing P_F while retaining anomaly detectability at a

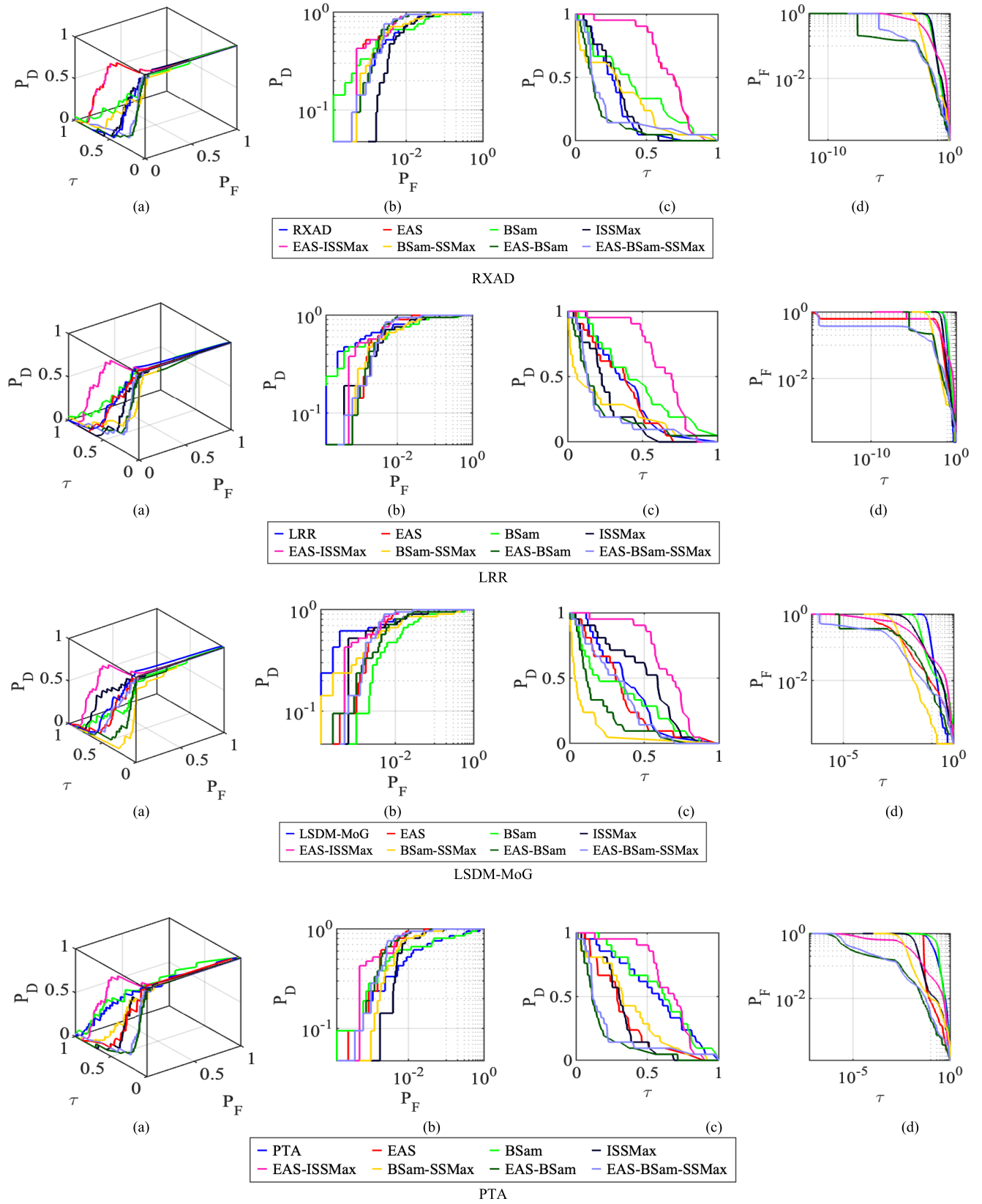


Fig. 12. 3-D ROC curves along with their corresponding 2-D ROC curves generated from Fig. 11 for the HYDICE urban scene. (a) 3-D ROC curves. (b) (P_D, P_F) curves. (c) (P_D, τ) curves. (d) (P_F, τ) curves.

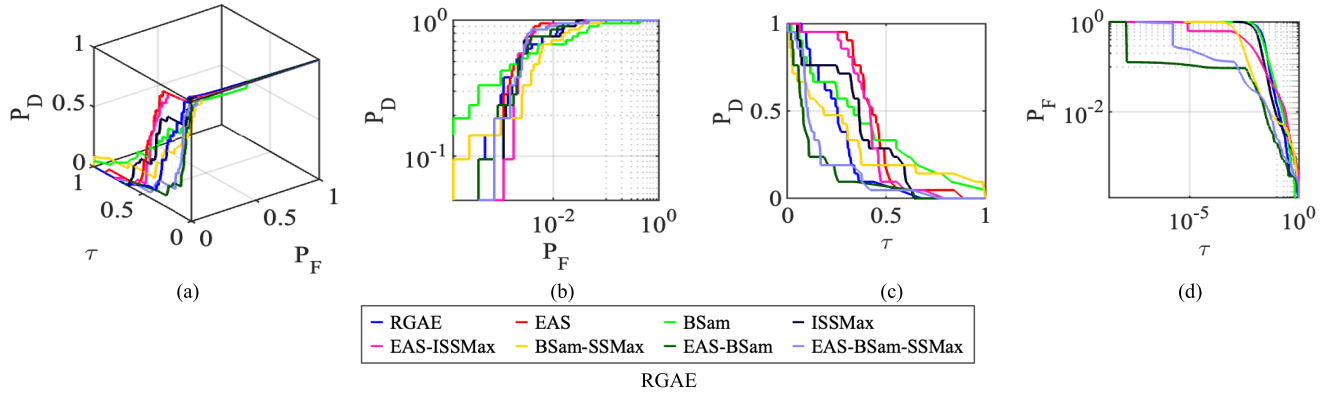


Fig. 12. (Continued.) 3-D ROC curves along with their corresponding 2-D ROC curves generated from Fig. 11 for the HYDICE urban scene. (a) 3-D ROC curves. (b) (P_D, P_F) curves. (c) (P_D, τ) curves. (d) (P_F, τ) curves.

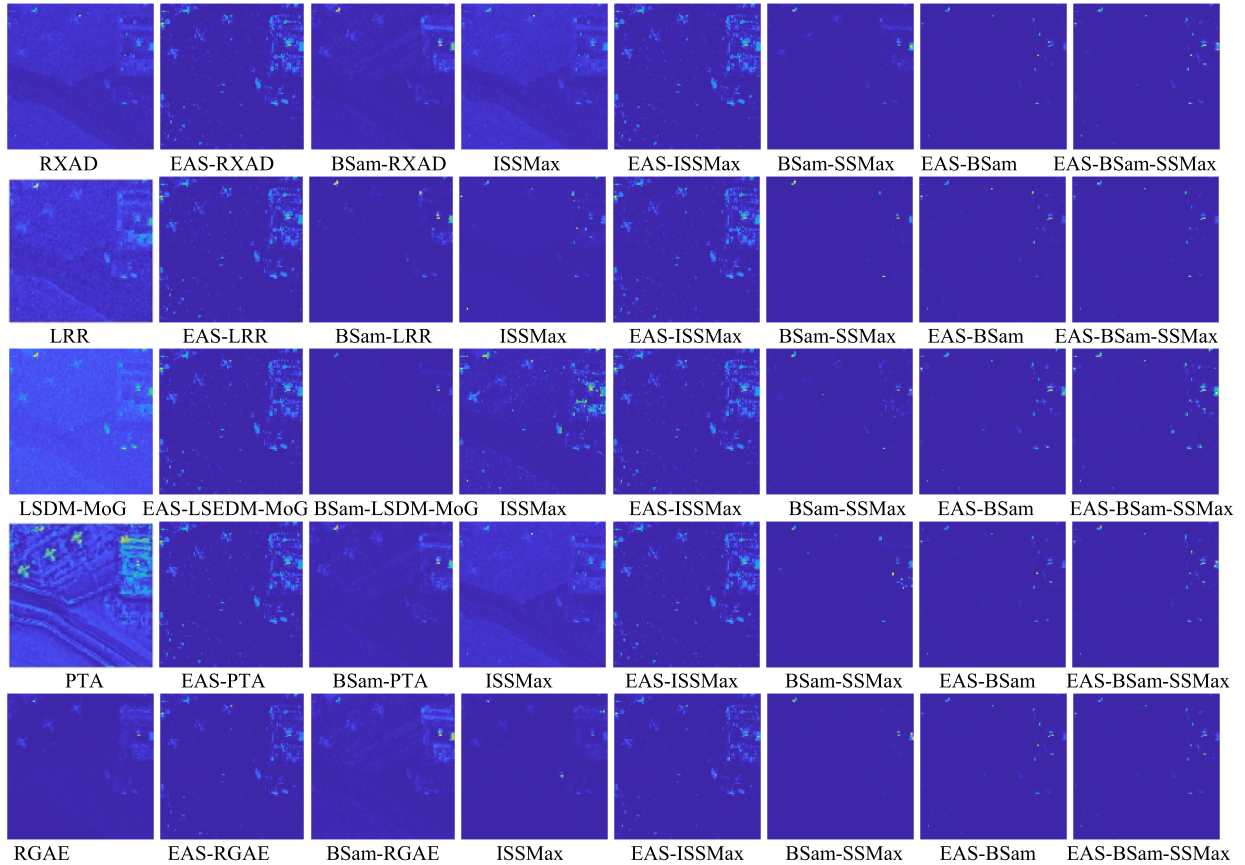


Fig. 13. ADMaps produced by RXAD, LRR, LSDM-MoG, PTA, and RGAE along with their corresponding EAS, BSAm, ISSMax, EAS-BSAm, and EAS-BSAm-SSMax counterparts using RXAD as the anomaly detector for the AVIRIS San Diego Airport scene.

satisfactory level. This was illustrated and validated in Fig. 8 by visual inspection and quantitative analysis in Table VI. Furthermore, according to Fig. 8 and Table VI, EAS-ISSMax and EAS-BSAm complemented each other according to the eight 3-D ROC-derived detection measures. It suggested that if we combined all three, EAS, BSAm, and ISSMax together, the resulting performance might be the best. Interestingly, this was not true as expected. This phenomenon can be explained as follows.

The experiments conducted on ISSMax and BSAm-SSMax in Fig. 8 showed that BSAm oversuppressed the BKG of

ISSMax. If we further examine EAS-ISSMax, the target pixels were double-detected because both EAS and ISSMax are designed to enhance anomaly detectability. This implied that the overdetection by EAS-ISSMax can be corrected by coupling with BSAm. However, the iterative process used by BSAm to capture spatial information among classified data samples also serves as the same functionality of the iterative process used in ISSMax. In this case, ISSMax must sacrifice its iterative process and be reduced to SSMax without feedback. As a consequence, including SSMax in EAS-BSAm did not necessarily improve BSAm-SSMax since SSMax does

TABLE VII
AUC VALUES PRODUCED BY RXAD, LRR, LSDM-MOG, PTA, AND RGAE ALONG WITH THEIR CORRESPONDING EAS, BSAM, ISSMAX, EAS-BSAM, AND EAS-BSAM-SSMAX COUNTERPARTS USING RXAD AS THE ANOMALY DETECTOR FOR THE HYDICE URBAN SCENE

Anomaly Detector	j	$TI_{j,j-1}$	$AUC_{(D,F)}$	AUC_{ADP}	AUC_{BDP}	AUC_{JAD}	AUC_{JBS}	AUC_{ADBS}	AUC_{SBPR}	AUC_{OADP}	time(s)
RXAD	--	--	0.9874	0.2575	0.9641	1.2449	1.9514	0.2216	0.2671	1.2216	0.08
EAS-RXAD	--	--	0.9957	0.6439	0.9767	1.6395	1.9724	0.6206	0.6593	1.6206	0.06
BSam-RXAD	--	--	0.9636	0.3627	0.9708	1.3263	1.9344	0.3335	0.3736	1.3335	0.00
ISSMax-RXAD	5	0.965	0.9865	0.2758	0.9647	1.2623	1.9511	0.2405	0.2859	1.2405	13.35
EAS-ISSMax-RXAD	1	0.996	0.9956	0.6442	0.9766	1.6398	1.9722	0.6208	0.6596	1.6208	4.91
BSam-SSMax-RXAD	10	0.97	0.9816	0.2977	0.9951	1.2793	1.9767	0.2928	0.2992	1.2928	35.75
EAS-BSam-RXAD	1	1	0.9958	0.1582	0.9981	1.154	1.9939	0.1564	0.1585	1.1564	6.46
EAS-BSam-SSMax-RXAD	3	1	0.9956	0.2116	0.9978	1.2072	1.9934	0.2094	0.2121	1.2094	13.16
LRR	--	--	0.9924	0.3712	0.941	1.3636	1.9333	0.3122	0.3945	1.3122	114.45
EAS-LRR	--	--	0.9955	0.6064	0.9784	1.6019	1.9739	0.5847	0.6198	1.5847	6.29
BSam-LRR	--	--	0.9703	0.2081	0.9928	1.1784	1.9631	0.2009	0.2096	1.2009	42.08
ISSMax-LRR	3	0.978	0.9862	0.2446	0.9741	1.2308	1.9603	0.2187	0.2511	1.2187	171.52
EAS-ISSMax-LRR	1	0.962	0.9955	0.6338	0.9773	1.6293	1.9728	0.6111	0.6485	1.6111	18.40
BSam-SSMax-LRR	8	1	0.9812	0.1849	0.9975	1.1662	1.9787	0.1824	0.1854	1.1824	364.31
EAS-BSam-LRR	4	1	0.9736	0.1745	0.9976	1.1481	1.9712	0.1721	0.1749	1.1721	39.32
EAS-BSam-SSMax-LRR	3	1	0.995	0.2077	0.9977	1.2027	1.9927	0.2054	0.2082	1.2054	30.30
LSDM-MoG	--	--	0.9915	0.3741	0.9281	1.3656	1.9196	0.3022	0.4031	1.3022	15.04
EAS-LSDM-MoG	--	--	0.9956	0.6459	0.9765	1.6416	1.9722	0.6225	0.6614	1.6225	7.75
BSam-LSDM-MoG	--	--	0.952	0.1837	0.9976	1.1357	1.9496	0.1813	0.1841	1.1813	0.69
ISSMax-LSDM-MoG	20	0.772	0.9867	0.4821	0.9754	1.4688	1.9621	0.4575	0.4943	1.4575	241.42
EAS-ISSMax-LSDM-MoG	1	0.979	0.9956	0.6468	0.9764	1.6424	1.9721	0.6232	0.6624	1.6232	28.38
BSam-SSMax-LSDM-MoG	2	1	0.9534	0.0893	0.9971	1.0427	1.9505	0.0863	0.0896	1.0863	19.63
EAS-BSam-LSDM-MoG	5	0.952	0.9774	0.1836	0.9951	1.1609	1.9725	0.1787	0.1845	1.1787	31.51
EAS-BSam-SSMax-LSDM-MoG	10	0.952	0.9961	0.3055	0.997	1.3016	1.9931	0.3025	0.3064	1.3025	83.96
PTA	--	--	0.916	0.5504	0.8617	1.4664	1.7777	0.4121	0.6387	1.4121	79.18
EAS-PTA	--	--	0.9956	0.6453	0.9767	1.6409	1.9724	0.622	0.6607	1.622	58.38
BSam-PTA	--	--	0.961	0.4173	0.9646	1.3783	1.9257	0.3819	0.4326	1.3819	5.11
ISSMax-PTA	20	0.978	0.9887	0.2959	0.9777	1.2846	1.9664	0.2736	0.3026	1.2736	1370.25
EAS-ISSMax-PTA	1	1	0.9956	0.6456	0.9766	1.6412	1.9722	0.6223	0.6611	1.6223	121.89
BSam-SSMax-PTA	16	0.975	0.9904	0.3652	0.9907	1.3555	1.981	0.3558	0.3686	1.3558	230.33
EAS-BSam-PTA	1	1	0.9956	0.1636	0.998	1.1592	1.9936	0.1617	0.1639	1.1617	17.27
EAS-BSam-SSMax-PTA	3	1	0.9955	0.2088	0.9977	1.2043	1.9932	0.2065	0.2093	1.2065	39.16
RGAE	--	--	0.9946	0.252	0.9737	1.2466	1.9683	0.2257	0.2588	1.2257	67.23
EAS-RGAE	--	--	0.9958	0.4263	0.9869	1.4221	1.9826	0.4132	0.4320	1.4132	67.45
BSam-RGAE	--	--	0.965	0.3567	0.9728	1.3217	1.9378	0.3295	0.3667	1.3295	20.19
ISSMax-RGAE	20	0.925	0.9939	0.3482	0.9825	1.3422	1.9765	0.3307	0.3544	1.3307	1492.96
EAS-ISSMax-RGAE	4	0.964	0.9955	0.3989	0.987	1.3944	1.9825	0.3859	0.4042	1.3859	361.50
BSam-SSMax-RGAE	17	0.958	0.988	0.2956	0.9947	1.2837	1.9827	0.2903	0.2972	1.2903	468.87
EAS-BSam-RGAE	2	1	0.9941	0.1332	0.9984	1.1273	1.9925	0.1315	0.1334	1.1315	70.95
EAS-BSam-SSMax-RGAE	3	1	0.9953	0.1664	0.9981	1.1617	1.9934	0.1645	0.1667	1.1645	94.88

not include an iterative feedback process, which is taken over by BSAM. This leads to the fact that BSAM-SSMax oversuppressed BKG, as shown in Fig. 8. Nevertheless, EAS-BSam-SSMax still managed to perform comparably to EAS-BSam with its performance nearly identical to that of EAS-BSam with no visual difference. Also, quantitatively

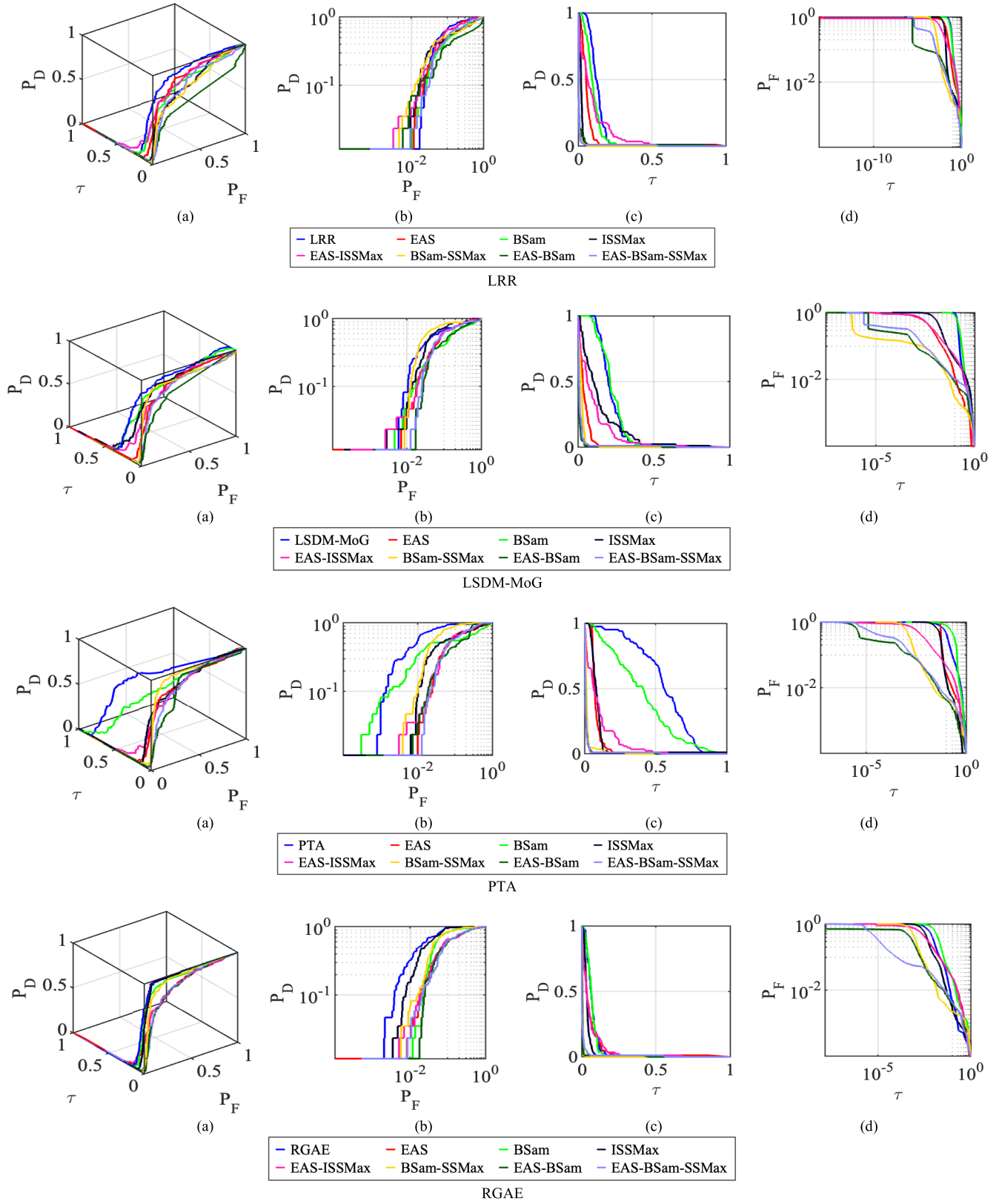


Fig. 14. 3-D ROC curves along with their corresponding 2-D ROC curves generated from Fig. 13 for the San Diego Airport scene. (a) 3-D ROC curves. (b) (P_D , P_F) curves. (c) (P_D , τ) curves. (d) (P_F , τ) curves.

speaking, EAS-BSam and EAS-BSam-ISSMax only differed by a threshold of less than 10^{-2} with EAS-BSam slightly better than EAS-BSam-SSMax.

In summary, all these experimental results demonstrated that IBCR by BSam allowed AD to better reduce P_F and noise effects. Specifically, all three EAS-ISSMax, EAS-BSam, and

TABLE VIII

AUC VALUES PRODUCED BY RXAD, LRR, LSDM-MOG, PTA, AND RGAE ALONG WITH THEIR CORRESPONDING EAS, BSAM, ISSMAX, EAS-BSAM, AND EAS-BSAM-SSMAX COUNTERPARTS USING RXAD AS THE ANOMALY DETECTOR FOR THE AVIRIS SAN DIEGO AIRPORT SCENE

Anomaly Detector	j	$TI_{j,j-1}$	$AUC_{(D,F)}$	AUC_{ADP}	AUC_{BDP}	AUC_{JAD}	AUC_{JBS}	AUC_{ADBS}	AUC_{SBPR}	AUC_{OADP}	time(s)
RXAD	--	--	0.8314	0.0802	0.9547	0.9116	1.7861	0.0349	0.0840	1.0349	0.09
EAS-RXAD	--	--	0.8512	0.0966	0.98	0.9477	1.8312	0.0766	0.0986	1.0766	0.09
BSam-RXAD	--	--	0.9147	0.0641	0.9786	0.9788	1.8933	0.0428	0.0655	1.0428	0.00
ISSMax-RXAD	1	0.964	0.8317	0.0765	0.9566	0.9082	1.7883	0.033	0.0800	1.033	6.38
EAS-ISSMax-RXAD	1	0.983	0.8512	0.0972	0.98	0.9484	1.8311	0.0772	0.0992	1.0772	7.30
BSam-SSMax-RXAD	3	0.96	0.9332	0.0345	0.9933	0.9677	1.9265	0.0278	0.0347	1.0278	20.99
EAS-BSam-RXAD	8	0.958	0.8495	0.0079	0.998	0.8573	1.8474	0.0058	0.0079	1.0058	45.50
EAS-BSam-SSMax-RXAD	7	0.97	0.8408	0.0162	0.9961	0.857	1.8369	0.0124	0.0163	1.0124	41.90
LRR	--	--	0.8789	0.1251	0.9314	1.004	1.8103	0.0565	0.1343	1.0565	129.18
EAS-LRR	--	--	0.8512	0.094	0.9806	0.9452	1.8318	0.0746	0.0959	1.0746	11.39
BSam-LRR	--	--	0.7906	0.0167	0.9915	0.8072	1.7821	0.0082	0.0168	1.0082	48.29
ISSMax-LRR	12	0.969	0.7911	0.0233	0.9872	0.8144	1.7783	0.0104	0.0236	1.0104	640.40
EAS-ISSMax-LRR	1	0.971	0.8511	0.0948	0.9806	0.9459	1.8317	0.0753	0.0967	1.0753	29.98
BSam-SSMax-LRR	8	1	0.7785	0.005	0.9975	0.7835	1.776	0.0025	0.0050	1.0025	422.12
EAS-BSam-LRR	2	1	0.6155	0.0154	0.9979	0.6308	1.6133	0.0133	0.0154	1.0133	61.07
EAS-BSam-SSMax-LRR	5	0.955	0.7956	0.009	0.9979	0.8046	1.7935	0.0069	0.0090	1.0069	114.03
LSDM-MoG	--	--	0.8403	0.2111	0.8571	1.0513	1.6973	0.0681	0.2463	1.0681	9.73
EAS-LSDM-MoG	--	--	0.8508	0.1003	0.9802	0.951	1.831	0.0805	0.1023	1.0805	42.64
BSam-LSDM-MoG	--	--	0.9068	0.0106	0.9963	0.9174	1.903	0.0069	0.0106	1.0069	1.02
ISSMax-LSDM-MoG	20	0.805	0.8976	0.1394	0.9721	1.037	1.8697	0.1115	0.1434	1.1115	308.31
EAS-ISSMax-LSDM-MoG	1	0.959	0.8504	0.0988	0.9801	0.9492	1.8305	0.0788	0.1008	1.0788	69.05
BSam-SSMax-LSDM-MoG	12	1	0.8756	0.0211	0.9974	0.8967	1.873	0.0185	0.0212	1.0185	110.99
EAS-BSam-LSDM-MoG	11	0.956	0.769	0.012	0.996	0.781	1.765	0.008	0.0120	1.008	125.30
EAS-BSam-SSMax-LSDM-MoG	14	0.958	0.8223	0.0171	0.9943	0.8394	1.8166	0.0114	0.0172	1.0114	244.93
PTA	--	--	0.9613	0.5501	0.8883	1.5113	1.8495	0.4383	0.6193	1.4383	104.12
EAS-PTA	--	--	0.8522	0.0965	0.9799	0.9487	1.8321	0.0764	0.0985	1.0764	74.27
BSam-PTA	--	--	0.9377	0.0571	0.9807	0.9948	1.9185	0.0378	0.0582	1.0378	5.15
ISSMax-PTA	1	0.964	0.849	0.08	0.9551	0.929	1.8041	0.0351	0.0838	1.0351	158.41
EAS-ISSMax-PTA	1	0.989	0.8514	0.0968	0.9798	0.9482	1.8313	0.0766	0.0988	1.0766	155.30
BSam-SSMax-PTA	15	1	0.9534	0.018	0.9966	0.9714	1.95	0.0146	0.0181	1.0146	260.98
EAS-BSam-PTA	4	0.966	0.8316	0.0116	0.9977	0.8432	1.8294	0.0093	0.0116	1.0093	42.57
EAS-BSam-SSMax-PTA	14	0.971	0.8637	0.0128	0.9967	0.8765	1.8605	0.0095	0.0128	1.0095	203.11
RGAE	--	--	0.9655	0.0667	0.9887	1.0322	1.9542	0.0553	0.0675	1.0553	85.43
EAS-RGAE	--	--	0.8511	0.0555	0.9887	0.9066	1.8398	0.0442	0.0561	1.0442	79.78
BSam-RGAE	--	--	0.9202	0.0649	0.9787	0.985	1.8989	0.0436	0.0663	1.0436	21.37
ISSMax-RGAE	12	1	0.9607	0.0344	0.9934	0.9951	1.9541	0.0278	0.0346	1.0278	1090.52
EAS-ISSMax-RGAE	20	0.937	0.8503	0.0553	0.9888	0.9055	1.839	0.044	0.0559	1.044	1786.91
BSam-SSMax-RGAE	13	1	0.9311	0.006	0.9975	0.9371	1.9287	0.0035	0.0060	1.0035	404.28
EAS-BSam-RGAE	8	0.96	0.849	0.0081	0.9977	0.857	1.8467	0.0058	0.0081	1.0058	252.46
EAS-BSam-SSMax-RGAE	9	1	0.8536	0.0103	0.998	0.8639	1.8516	0.0083	0.0103	1.0083	284.31

EAS-BSam-SSMax detected the subpixel panels as anomalies and BSam could not work alone to achieve the best AD in the third column. This further showed that EAS, ISSMax, performance.

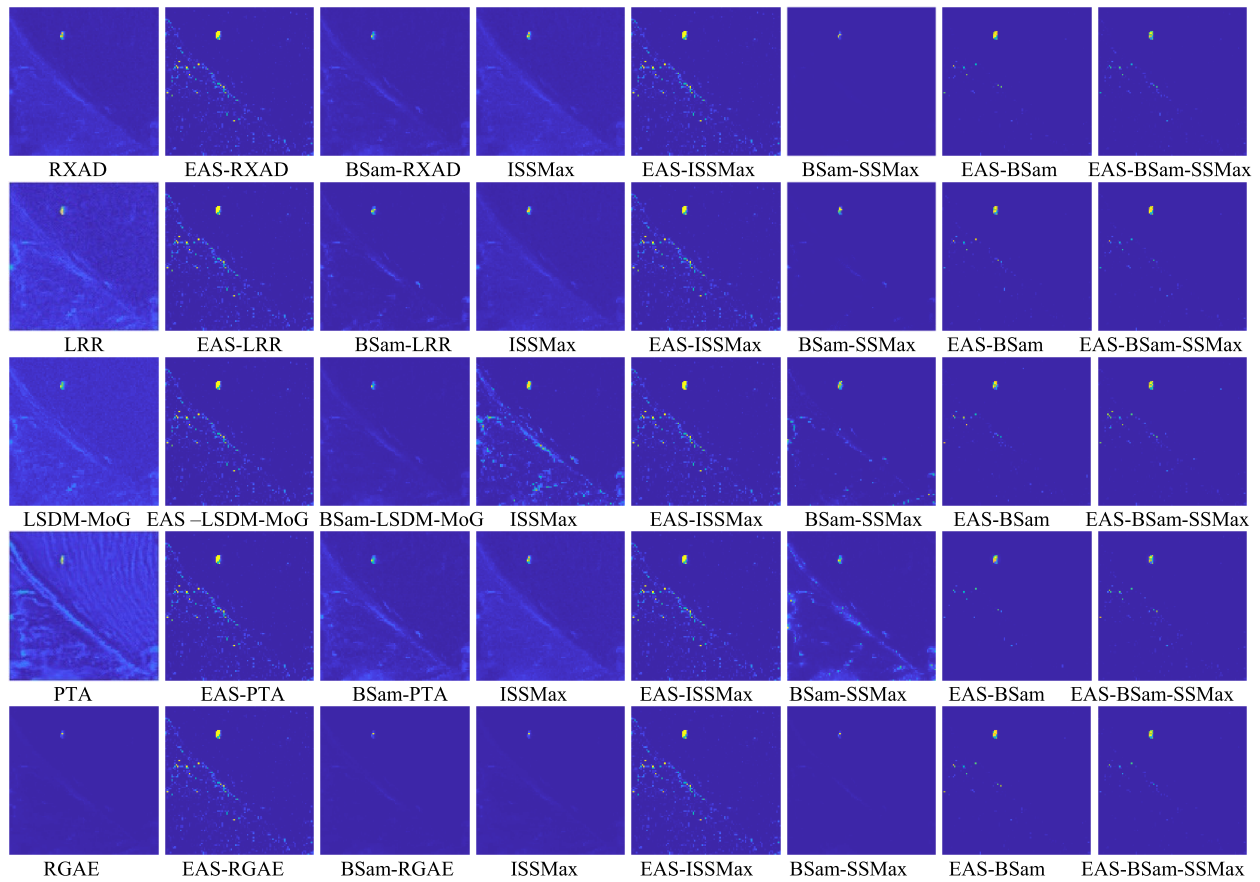


Fig. 15. ADMaps produced by RXAD, LRR, LSDM-MoG, PTA, and RGAE along with their corresponding EAS, BSAM, ISSMax, EAS-BSAM, and EAS-BSAM-SSMax counterparts using RXAD as the anomaly detector for the AVIRIS Bay Champagne Beach scene.

As for LRR, LSDM-MoG, and PTA, similar conclusions were also applied except in the case of RGAE where EAS-ISSMax was the best in anomaly detectability, while EAS-BSAM-SSMax was best and performed slightly better than EAS-BSAM in BS and noise effect reduction.

With regard to the HYDICE urban scene, it is less interesting than the HYDICE panel scene since it is also a HYDICE dataset but does not have subpixel targets to be detected as anomalies. Thus, all the conclusions drawn for the panel dataset were also applied to this scene.

B. AVIRIS San Diego Airport and Bay Champagne Beach Scenes

The San Diego Airport scene in Fig. 7 was discussed in detail in Section VIII of the EAS [14]. The main issue that arises in this scene is the three airplanes to be considered anomalies. Since these three airplanes occupy a total of 85 pixels, they are so large and very visible. They are hardly considered anomalous targets. Because of that, RXAD and CRD did not detect these airplanes as anomalies, but other methods such as tensor-decomposition-based PTA did pick up airplanes as anomalies. This fact was confirmed by Fig. 11 where RXAD did not detect these three airplanes, while LRR, LSDM-MoG, and RGAE did detect these airplanes weakly with PTA detecting these targets very visibly. It is obvious that the results of RXAD and CRD contradict those produced

by PTA. Their results were indeed against each other. Thus, if RXAD is an anomaly detector, then PTA is not. On the other hand, if PTA is an anomaly detector, then RXAD is not. Thus, based on our experience, if the targets are visible and can be inspected visually, these targets should not be considered as anomalies, as detailed in [17, Sec. IX.C]. For example, in the San Diego airport dataset, airplanes should be considered anomalies. Instead, anomalies should be those detected in the right upper corner. Unfortunately, since there is no ground truth, we cannot verify these anomalies.

Furthermore, if we consider these three airplanes as anomalous targets, then, according to Fig. 10 and Table VIII, PTA yielded the best performance among all the test methods. Are these conclusions true? As also shown in [14], since the three airplanes can be identified by visual inspection, we can simply select a few pixels located in the centers of these airplanes as the desired target signatures and then apply a subpixel target detector, constrained energy minimization (CEM) [47], [48], [49], [50] and its two variants, hierarchical CEM [51], and ensemble CEM (ECM) [52] to this scene. The detection rates were nearly 100%, which was much better than PTA's detection map. These experiments demonstrated that the San Diego airport scene is better to be used for TD but not appropriate to be used for AD. Moreover, due to the fact that SSMax can improve anomaly detectability and the three airplanes can be considered as visible targets, EAS-BSAM-SSMax did perform better than EAS-BSAM according

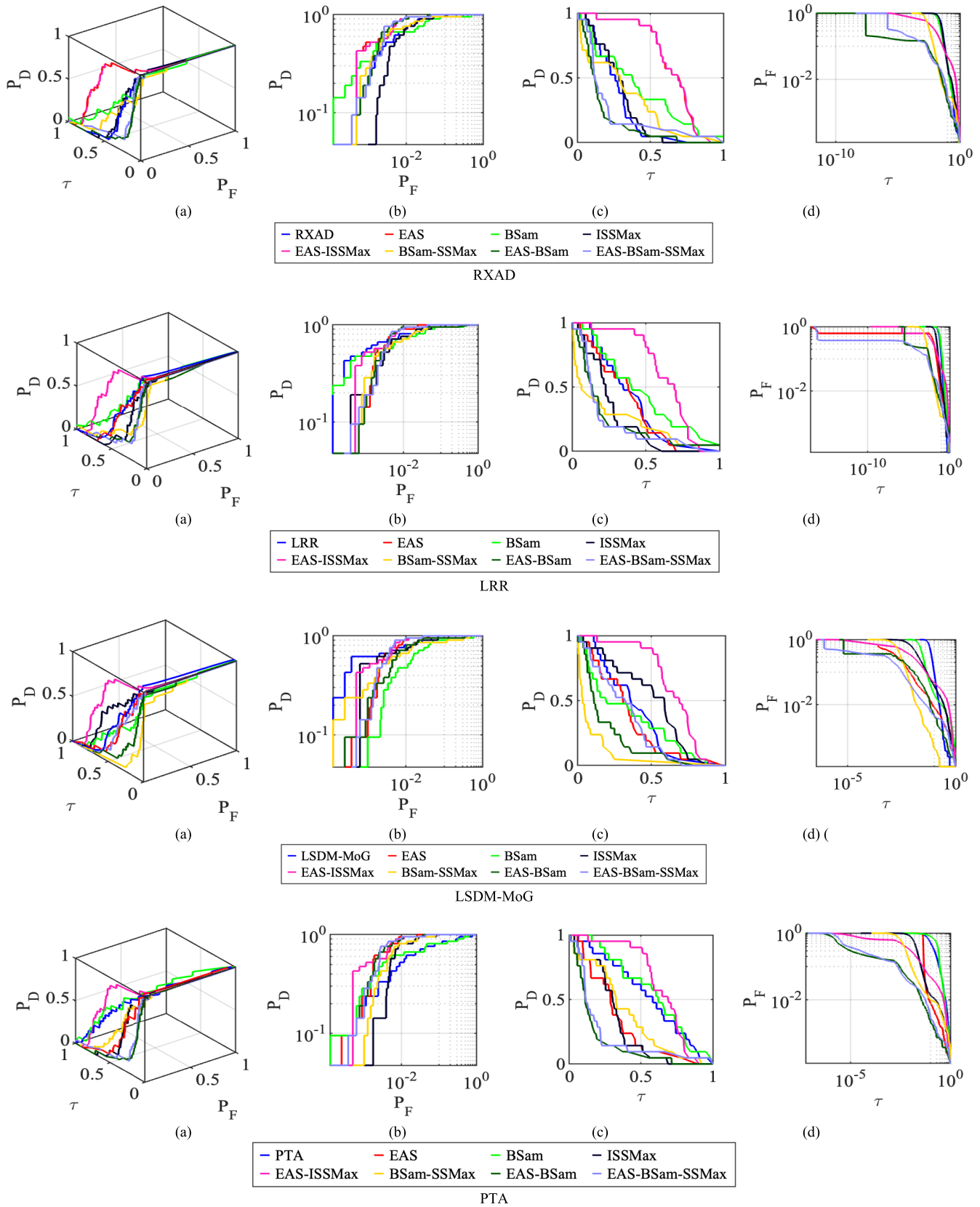


Fig. 16. 3-D ROC curves along with their corresponding 2-D ROC curves generated from Fig. 15 for the Bay Champagne Beach scene. (a) 3-D ROC curves. (b) (P_D , P_F) curves. (c) (P_D , τ) curves. (d) (P_F , τ) curves.

to Table VIII. This indicated that including SSMax was an advantage to improve EAS-BSAm in terms of detection targets.

Unlike the San Diego Airport scene in Fig. 7, which has three large airplanes, the Bay Champagne Beach scene

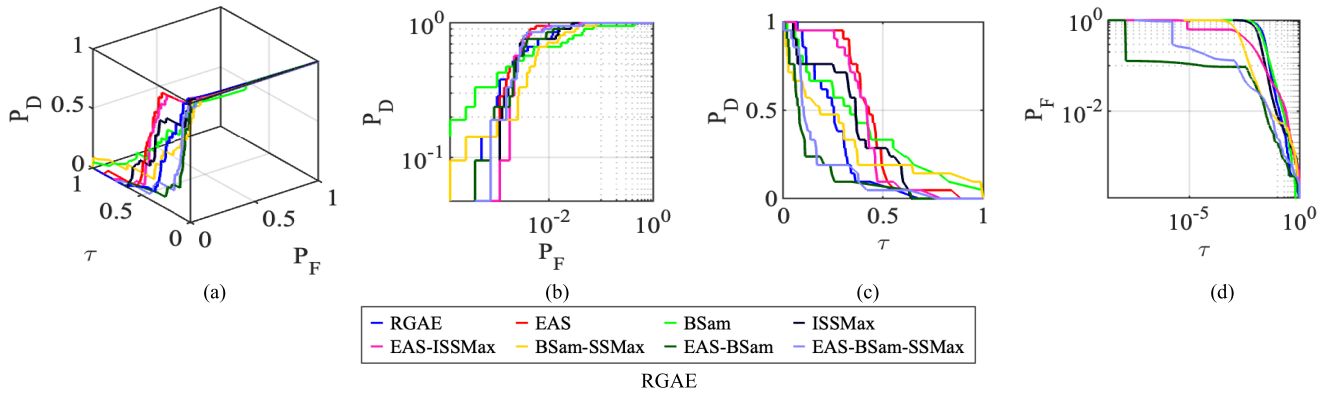


Fig. 16. (Continued.) 3-D ROC curves along with their corresponding 2-D ROC curves generated from Fig. 15 for the Bay Champagne Beach scene. (a) 3-D ROC curves. (b) (P_D , P_F) curves. (c) (P_D , τ) curves. (d) (P_F , τ) curves.

in Fig. 8 contains a relatively small target, which can be considered an anomaly. This scene is a very typical scene used for AD since only one anomaly is located upper left quadrant. Compared to the anomalies in HYDICE data scenes in Figs. 5 and 6, which either are subpixels or occupy one or two pixels, the anomaly in Fig. 8 is a single anomaly, which is made up of 11 pixels, but occupies only 0.1% of data size. As a result, the conclusions drawn from this scene were similar to those drawn from HYDICE datasets but quite different from the conclusions made for the San Diego Airport scene. For example, the PTA that was the best detector for the San Diego Airport scene became the worst detector for the Bay Champagne Beach scene. On the contrary, EAS-BSam-RXAD and EAS-BSam-SSMax that did not detect the airplanes well in the San Diego Airport scene turned out to be among the best detectors, which detected the anomaly effectively in the Bay Champagne Beach scene. These two scenes provided completely contrasting examples to show how crucial to select an appropriate dataset for anomaly detectors. This evidence was also shown in 3-D/2-D ROC curves in Figs. 14 and 16.

In summary, there are some detectors that claim to be anomaly detectors using the San Diego Airport scene, such as PTA, LRR, and LSDM-MoG, and are actually target detectors not really designed for anomaly detectors. Similar scenarios to the San Diego Airport data are also observed in [12] and [13] where the used datasets, SD1, SD2, Gainesville, and Gulfport, all contain airplanes that are very visible. We can very easily identify these targets by visual inspection and perform TD, which is more appropriate AD. It is generally understood that a target is called an anomaly mainly because it cannot be identified by visual inspection or prior knowledge. If this is the case, why should we consider these airplanes as anomalies? In this case, we can use target knowledge obtained by visual inspection to simply perform TD rather than AD. Unfortunately, such dilemmas have never been investigated for AD in the past. An effort devoted to exploring these issues is currently being undertaken [53].

C. Separability Map Versus 3-D ROC Curve

There is a commonly used separability map that utilizes box plots as a means of separating anomalies from BKG.

According to [54, p. 77], a box plot is constructed to “visualize” anomaly-BKG separability to evaluate the performance of an AD method. More specifically, a separability map is composed of two boxes that represent the distribution of anomalous pixel values and the distribution of BKG pixel values with a line in the middle of each box to indicate either the average value or median of the pixels value. The vertical length of each box in a separability map can be considered as a variation from the average value used to specify a tolerance range. The larger the gap between the anomaly box and the BKG box, the better the separation between anomalies and the BKG. Also, the shorter the length of an anomaly box, the greater the anomaly detectability and so is the BKG box. Although the box length was commonly adopted as [10%, 90%], such as [34] and [37], this was an empirical choice. The key issue arising in using a separability map is how to calculate the box length, which is actually determined by an appropriate value τ to threshold real-valued detector statistics produced by an AD method. A separability map only shows the detection performance of a detector, which uses a particular value, τ , to threshold its detector statistics. In other words, various values of τ used to threshold the detector statistics result in different separability maps produced by the same detector. Correspondingly, this can be explained by a Neyman–Pearson detector. For a given false alarm probability, P_F , we can find a threshold value τ , which determines P_F . This τ is then used to calculate its corresponding detection probability, P_D . Thus, in terms of the 2-D ROC terminology, a separability map produced by a detector using a specific threshold τ to threshold its detector statistics actually corresponds to a pair of (P_D , P_F) as a point on a 2-D ROC curve generated by the Neyman–Pearson detector using a particular threshold τ where the lengths of anomaly and BKG boxes in a separability map are calculated by P_D and $1 - P_F$, respectively. Thus, varying threshold values τ for a separability map correspond to varying the value of P_F for a 2-D ROC curve. More specifically, each separability map only corresponds to one point on a 2-D ROC curve. Thus, technically speaking, in order to see a full picture of a detector using a separability map, we need to show separability maps produced by all possible threshold values of τ . It is practically impossible. However, this is

TABLE IX
AUC VALUES PRODUCED BY RXAD, LRR, LSDM-MOG, PTA, AND RGAE ALONG WITH THEIR CORRESPONDING EAS, BSAM, ISSMAX, EAS-BSAM, AND EAS-BSAM-SSMAX COUNTERPARTS USING RXAD AS THE ANOMALY DETECTOR FOR THE AVIRIS BAY CHAMPAGNE BEACH SCENE

Anomaly Detector	j	TI_{jj-1}	$AUC_{(D,F)}$	AUC_{ADP}	AUC_{BDP}	AUC_{JAD}	AUC_{JBS}	AUC_{ADBS}	AUC_{SBPR}	AUC_{OADP}	time(s)
RXAD	--	--	0.9998	0.4845	0.9741	1.4844	1.9739	0.4586	0.4974	1.4586	0.27
EAS-RXAD	--	--	0.9991	0.9015	0.9901	1.9006	1.9891	0.8916	0.9105	1.8916	0.13
BSam-RXAD	--	--	0.9998	0.3319	0.9883	1.3317	1.9881	0.3202	0.3358	1.3202	0.01
ISSMax-RXAD	1	1	0.9998	0.4537	0.9771	1.4536	1.9769	0.4308	0.4643	1.4308	6.99
EAS-ISSMax-RXAD	1	0.981	0.9988	0.9	0.99	1.8988	1.9888	0.89	0.9091	1.89	6.75
BSam-SSMax-RXAD	4	1	0.9997	0.2316	0.9985	1.2314	1.9982	0.2301	0.2319	1.2301	16.27
EAS-BSam-RXAD	6	1	0.9997	0.8284	0.9988	1.8281	1.9985	0.8271	0.8294	1.8271	22.30
EAS-BSam-SSMax-RXAD	11	0.96	0.9995	0.6096	0.998	1.6091	1.9975	0.6076	0.6108	1.6076	38.34
LRR	--	--	0.9998	0.6329	0.9401	1.6327	1.9398	0.573	0.6732	1.573	58.80
EAS-LRR	--	--	0.9988	0.9001	0.9903	1.8989	1.989	0.8904	0.9089	1.8904	22.11
BSam-LRR	--	--	0.9993	0.3786	0.9934	1.3779	1.9927	0.372	0.3811	1.372	53.06
ISSMax-LRR	1	1	0.9997	0.4852	0.9827	1.4849	1.9824	0.468	0.4937	1.468	119.53
EAS-ISSMax-LRR	1	1	0.9992	0.9009	0.9902	1.9002	1.9894	0.8911	0.9098	1.8911	53.22
BSam-SSMax-LRR	7	1	0.9996	0.4121	0.9981	1.4117	1.9977	0.4102	0.4129	1.4102	463.03
EAS-BSam-LRR	5	1	0.9996	0.7059	0.9987	1.7055	1.9983	0.7046	0.7068	1.7046	80.57
EAS-BSam-SSMax-LRR	6	1	0.9997	0.7892	0.9987	1.7889	1.9984	0.7879	0.7902	1.7879	81.13
LSDM-MoG	--	--	0.9996	0.517	0.9387	1.5166	1.9384	0.4558	0.5508	1.4558	10.89
EAS-LSDM-MoG	--	--	0.999	0.9091	0.9902	1.9081	1.9892	0.8993	0.9181	1.8993	7.28
BSam-LSDM-MoG	--	--	0.9998	0.2861	0.9912	1.2858	1.9909	0.2772	0.2886	1.2772	1.40
ISSMax-LSDM-MoG	20	0.801	0.9993	0.7798	0.9799	1.7791	1.9793	0.7597	0.7958	1.7597	365.13
EAS-ISSMax-LSDM-MoG	2	1	0.9989	1	0.9901	1.9989	1.989	0.9901	1.0100	1.9901	36.43
BSam-SSMax-LSDM-MoG	20	0.922	0.9994	0.6575	0.9949	1.6569	1.9943	0.6524	0.6609	1.6524	188.23
EAS-BSam-LSDM-MoG	7	0.955	0.9994	0.651	0.9984	1.6504	1.9979	0.6494	0.6520	1.6494	42.29
EAS-BSam-SSMax-LSDM-MoG	13	1	0.9993	0.6909	0.9979	1.6902	1.9972	0.6888	0.6924	1.6888	80.26
PTA	--	--	0.9982	0.5371	0.9203	1.5353	1.9185	0.4574	0.5836	1.4574	90.29
EAS-PTA	--	--	0.9995	0.9012	0.9902	1.9007	1.9898	0.8914	0.9101	1.8914	87.89
BSam-PTA	--	--	0.9995	0.4022	0.9789	1.4016	1.9784	0.3811	0.4109	1.3811	6.31
ISSMax-PTA	1	0.981	0.9999	0.4794	0.9739	1.4793	1.9737	0.4533	0.4922	1.4533	186.44
EAS-ISSMax-PTA	1	0.981	0.9991	0.9009	0.9901	1.9001	1.9893	0.8911	0.9099	1.8911	193.24
BSam-SSMax-PTA	20	0.84	0.9949	0.4397	0.9833	1.4346	1.9782	0.423	0.4472	1.423	364.65
EAS-BSam-PTA	4	1	0.9997	0.6725	0.9988	1.6721	1.9984	0.6713	0.6733	1.6713	66.85
EAS-BSam-SSMax-PTA	9	0.955	0.9992	0.5865	0.9981	1.5858	1.9974	0.5847	0.5876	1.5847	148.19
RGAE	--	--	0.9997	0.1376	0.9961	1.1373	1.9958	0.1337	0.1381	1.1337	74.57
EAS-RGAE	--	--	0.9996	0.902	0.9905	1.9016	1.9901	0.8924	0.9107	1.8924	74.45
BSam-RGAE	--	--	0.9994	0.1226	0.9975	1.122	1.9969	0.12	0.1229	1.12	21.00
ISSMax-RGAE	1	1	0.9998	0.1337	0.9962	1.1335	1.9959	0.1299	0.1342	1.1299	155.55
EAS-ISSMax-RGAE	2	1	0.9995	0.8992	0.9904	1.8987	1.9899	0.8896	0.9079	1.8896	235.37
BSam-SSMax-RGAE	1	1	0.9995	0.1103	0.9985	1.1098	1.998	0.1088	0.1105	1.1088	47.88
EAS-BSam-RGAE	6	1	0.9996	0.7552	0.9987	1.7548	1.9983	0.7539	0.7562	1.7539	178.74
EAS-BSam-SSMax-RGAE	3	0.952	0.9998	0.6952	0.999	1.695	1.9988	0.6942	0.6959	1.6942	99.69

not a problem for a 2-D ROC curve because a 2-D ROC curve can be generated by running through the value of P_F over $[0, 1]$, while a separability map cannot. If we further look at a 2-D ROC curve, it only provides qualitative results by visual inspection as a separability map does. In order to quantify a 2-D ROC curve, we need to calculate the AUC

TABLE X
EIGHT DETECTION MEASURES CALCULATED BY RXAD ALONG WITH ITS CORRESPONDING EAS, BSAM, ISSMAX, EAS-BSAM,
AND EAS-BSAM-SSMAX COUNTERPARTS FOR AND HYDICE PANEL SCENE

Anomaly Detector	AUC _(D,F)	AUC _{ADP}	AUC _{BDP}	AUC _{JAD}	AUC _{JBS}	AUC _{ADBS}	AUC _{SBPR}	AUC _{OADP}
RXAD	0.99	0.3574	0.9573	1.3474	1.9473	0.3147	0.3733	1.3147
EAS-RXAD	0.9932	0.9227	0.9627	1.9159	1.9559	0.8854	0.9585	1.8854
BSam-RXAD	0.9839	0.2816	0.9862	1.2655	1.9701	0.2678	0.2855	1.2678
ISSMax-RXAD	0.9918	0.3618	0.9584	1.3537	1.9502	0.3202	0.3775	1.3202
EAS-ISSMax-RXAD	0.9932	0.9233	0.9625	1.9165	1.9557	0.8858	0.9593	1.8858
BSam-SSMax-RXAD	0.9727	0.2383	0.9949	1.2109	1.9675	0.2331	0.2395	1.2331
EAS-BSam-RXAD	0.9974	0.5932	0.9919	1.5906	1.9893	0.5851	0.5980	1.5851
EAS-BSam-SSMax-RXAD	0.997	0.5911	0.9916	1.5881	1.9886	0.5827	0.5961	1.5827

TABLE XI
EIGHT DETECTION MEASURES CALCULATED BY PTA ALONG WITH ITS CORRESPONDING EAS, BSAM, ISSMAX, EAS-BSAM,
AND EAS-BSAM-SSMAX COUNTERPARTS FOR AND HYDICE PANEL SCENE

Anomaly Detector	AUC _(D,F)	AUC _{ADP}	AUC _{BDP}	AUC _{JAD}	AUC _{JBS}	AUC _{ADBS}	AUC _{SBPR}	AUC _{OADP}
PTA	0.8315	0.3894	0.8632	1.221	1.6948	0.2527	0.4511	1.2527
EAS-PTA	0.9933	0.9177	0.9627	1.911	1.956	0.8804	0.9533	1.8804
BSam-PTA	0.9834	0.2386	0.9861	1.222	1.9695	0.2247	0.2420	1.2247
ISSMax-PTA	0.9915	0.3862	0.9548	1.3778	1.9464	0.3411	0.4045	1.3411
EAS-ISSMax-PTA	0.9933	0.9182	0.9625	1.9115	1.9558	0.8806	0.9540	1.8806
BSam-SSMax-PTA	0.9907	0.3012	0.9941	1.2919	1.9849	0.2953	0.3030	1.2953
EAS-BSam-PTA	0.9973	0.5605	0.9922	1.5578	1.9895	0.5526	0.5649	1.5526
EAS-BSam-SSMax-PTA	0.996	0.526	0.9913	1.522	1.9873	0.5173	0.5306	1.5173

to represent its quantitative value. Since a separability map is a result of one single unspecified threshold, there is no such counterpart similar to a 2-D ROC curve. Accordingly, no corresponding AUC values can be calculated to quantify a separability map. This is the reason that when a separability map is used, it is generally coupled with the AUC of a 2-D ROC curve to be used for quantitative analysis. Because of that, a separability map can be viewed as a point on a 2-D ROC curve with P_D and $1 - P_F$ interpreted as AD and BS probabilities calculated by a specific threshold value, respectively [17]. Unfortunately, both (anomaly-box and BKG-box) are used for a separability map and (P_D , P_F) used for a 2-D ROC curve are calculated by the same threshold τ . In other words, as the length of the anomaly box is changed, so is the length of the BKG box. This is the same issue that arises in using P_D and P_F for a 2-D ROC curve. To address this threshold issue, a 3-D ROC curve of (P_D , P_F , τ) is particularly designed by including the threshold value τ as an independent parameter to take care of the dependence between P_D and P_F . Therefore, once a 3-D ROC curve is used, there is no need for a separability map that does not provide additional values.

Using the HYDICE panel scene in Fig. 5 as an illustrative example, Fig. 17(a) plots its separability maps of RXAD and the corresponding EAS, BSam, ISSMax, EAS-BSam, and EAS-BSam-SSMax counterparts using RXAD as the anomaly detector where red and black boxes represent

anomaly and BKG with the range of [10%, 90%], which was used in [34] and [37] corresponding to a certain value of the threshold τ . According to Fig. 17(a), EAS yielded the best anomaly detectability (EAS and EAS-ISSMax), while BSam produced the best BS (BSam-SSMax, EAS-BSam, and EAS-BSam-SSMax). As for the separation of anomalies from BKG, the best anomaly detectors were EAS and EAS-ISSMax. However, these observations are all based on visual inspection. Quantitatively speaking, there is no way to distinguish one from another in each of these three groups.

To further demonstrate the advantages of using 3-D ROC curve-derived eight detection measures over separability maps, we reproduced Table X from Table VI in comparison with separability maps in Fig. 17(a). In the first group of best anomaly detectability, EAS-ISSMax produced the best values of AUC_{ADP} and AUC_{JAD}. In the second group of best BS, the black boxes produced by BSam-SSMax, EAS-BSam, and EAS-BSam-SSMax were too thin to tell them apart visually. However, from Table X, EAS-SSMax-RXAD and EAS-BSam-RXAD yielded the best values of AUC_{BDP} and AUC_{JBS}, respectively. In the third group of separating anomalies from BK, EAS-ISSMax produced the best values of AUC_{ADBS}, AUC_{SBPR}, and AUC_{OADP}. Most interestingly, according to all the eight AUC values, both EAS-BSam-RXAD and EAS-BSam-SSMax-RXAD produced very close values within an error range of 10^{-2} . Apparently, the separability maps in Fig. 17 cannot offer all these quantitative comparisons.

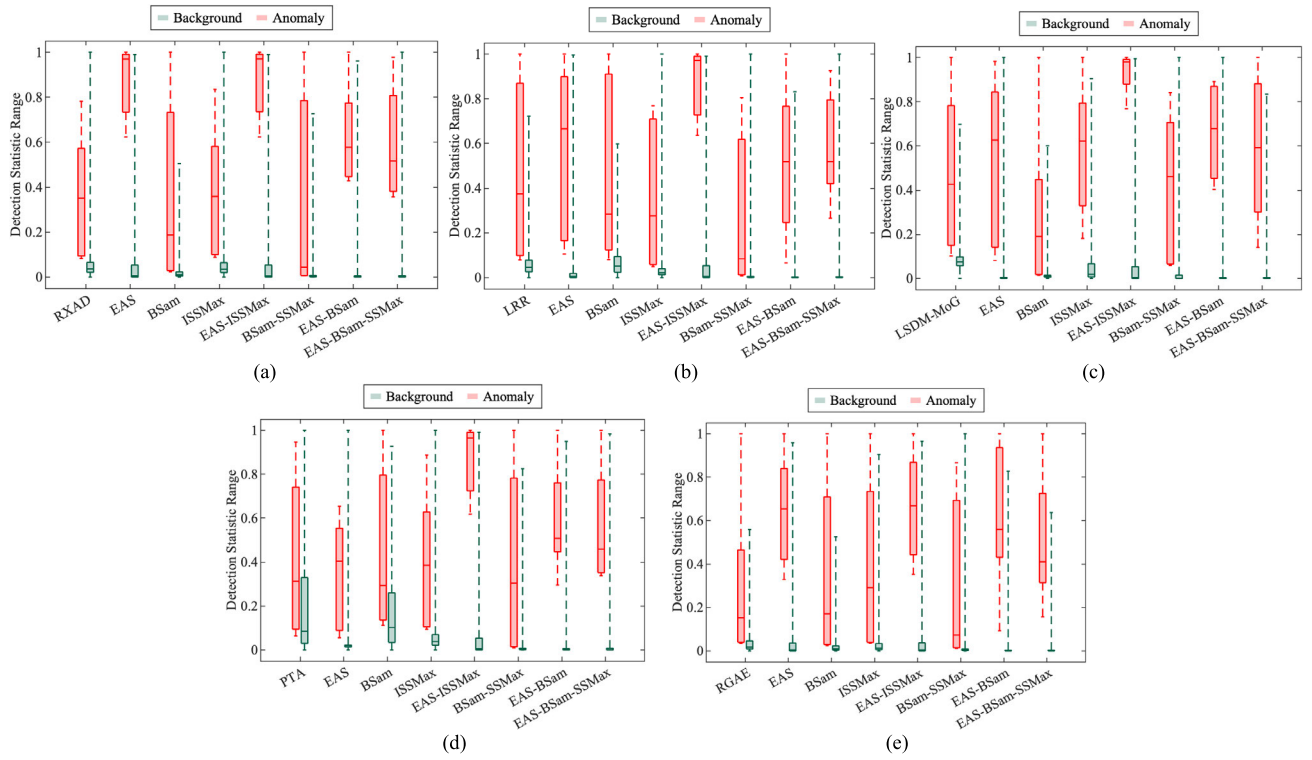


Fig. 17. BKG-anomaly separability maps of (a) RXAD, (b) LRR, (c) LSDM-MoG, (d) PTA, and (e) RGAE and their corresponding EAS, BSAM, ISSMax, EAS-BSAM, and EAS-BSAM-SSMax counterparts using RXAD as the anomaly detector for the HYDICE panel scene.

Similarly, Fig. 17(b)–(e) plots separability maps of the other four anomaly detectors: LRR, LSDM-MoG, PTA, and RGAE for the HYDICE panel scene. The discussions on the separability maps in Fig. 17(a) are also applicable to Fig. 17(b)–(e) and Tables VII–IX. Of particular interest is Fig. 17(d) for PTA where PTA and BSAM-PTA had large overlaps between anomaly and BKG boxes. However, the separability map did not quantitatively measure how much the overlap is unless we used Table X reproduced from Table VI for quantitative analysis. As for the other three data scenes, the above same argument can be readily applied. Due to limited space, their experimental results are not included.

As a concluding remark, in order to see the whole picture of comparative analysis among the five test anomaly detectors, RXAD, LSDM-MoG, PTA, and RGAE for five datasets, it requires 20 separability maps to complete the task only just for “qualitative” analysis. On the other hand, using 3-D ROC curve-generated Tables VI–IX, we can complete a “quantitative” analysis, which cannot be done by separability maps. The evidence clearly shows the superiority of the 3-D ROC analysis to the separability map.

VIII. CONCLUSION

Whether or not an AD method is effective is determined by four factors: 1) how well BKG can be suppressed; 2) how much anomaly detectability can be enhanced; 3) how effective noise can be removed; and 4) how much interband correlation can be reduced, each of which corresponds to one of four major issues 1)–4) related to AD described in the introduction. Over the past years, BKG estimation and

reconstruction have been one of the major trends. This is because it is based on assumptions that either BKG can be characterized by Gaussian models or learned by training samples. Unfortunately, these assumptions are generally not justified because BKG may exhibit high-order statistics and cannot be described by Gaussian distributions. Specifically, when BKG is very complicated, how to find or determine effective training samples to faithfully represent BKG is a challenging issue. Another mainstream is to decompose a data cube \mathbf{X} into a low-rank space \mathbf{L} that corresponds to BKG and a sparse space \mathbf{S} that contains anomalies. The associated key issues are how to determine the ranks of \mathbf{L} and \mathbf{S} and how to cope with noise, which may be non-Gaussian and cannot be separated by a matrix decomposition. This article provides feasible solutions to cope with the issues. It combines EAS and BSAM to reduce noise effects, while it combines EAS and SSMax to enhance anomaly detectability.

Moreover, since datasets are real, there is no way to identify what type of noise is really involved in the real data. In particular, the noise in real data is generally characterized by several factors. If the noise is Gaussian, it can be easily removed by noise whitening. Unfortunately, the noise in hyperspectral imagery is generally non-Gaussian. Another is that the noise is not additive. In other words, noise is nonlinearly mixed with anomalies and BKG. This means that noise is not uncorrelated with target signal sources. A third one is that noise can be of any type. For example, in communications, noise can be the Raleigh noise. However, in magnetic resonance imaging, the noise can be Rician. In an HSI, there is no specific type of noise that can be used to simulate synthetic

images. A fourth one is that anomalies to be implanted in synthetic data cannot be simply by superimposition as additive targets. By factoring in all these issues, it is impractical and unrealistic to use synthetic experiments for validation because synthetic experiments cannot really reflect what a real situation is. Specifically, the noise simulated for synthetic images is generally not additive. With all things considered, coming up with reliable synthetic experiments may be very difficult if not impossible. Besides, even though an anomaly detector performs well on synthetic experiments, in many cases, it may be very likely to fail on real data. Accordingly, it is believed that synthetic experiments may sometimes lead to unnecessary confusion and inconclusive claims.

REFERENCES

- [1] C.-I. Chang, *Real-Time Progressive Hyperspectral Image Processing: Endmember Finding and Anomaly Detection*. New York, NY, USA: Springer, 2016.
- [2] C.-I. Chang, "Hyperspectral anomaly detection: A dual theory of hyperspectral target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5511720.
- [3] C.-I. Chang, "Target-to-anomaly conversion for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5540428.
- [4] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York, NY, USA: Springer-Verlag, 1991.
- [5] I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 10, pp. 1760–1770, Oct. 1990.
- [6] D. Manolakis, E. Truslow, M. Pieper, T. Cooley, and M. Brueggeman, "Detection algorithms in hyperspectral imaging systems," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 24–33, Jan. 2014.
- [7] C.-I. Chang, C.-Y. Lin, P.-C. Chung, and P. F. Hu, "Iterative spectral-spatial hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5511705.
- [8] Y. Xu, Z. Wu, J. Li, A. Plaza, and Z. Wei, "Anomaly detection in hyperspectral images based on low-rank and sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, pp. 1990–2000, Apr. 2016.
- [9] E. J. Candes, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *J. ACM*, vol. 58, no. 3, pp. 1027–1063, 2009.
- [10] T. Zhou and D. Tao, "GoDec: Randomized low-rank and sparsity matrix decomposition in noisy case," in *Proc. 28th Int. Conf. Mach. Learn.*, Bellevue, WA, USA, 2011.
- [11] C.-I. Chang, H. Cao, S. Chen, X. Shang, M. Song, and C. Yu, "Orthogonal subspace projection-based GoDec for low rank and sparsity matrix decomposition for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 3, pp. 2403–2429, Mar. 2021.
- [12] X. He, J. Wu, Q. Ling, Z. Li, Z. Lin, and S. Zhou, "Anomaly detection for hyperspectral imagery via tensor low-rank approximation with multiple subspace learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5509917.
- [13] T. Guo, L. He, F. Luo, X. Gong, Y. Li, and L. Zhang, "Anomaly detection of hyperspectral image with hierarchical antinoise mutual-incoherence-induced low-rank representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5510203.
- [14] C.-I. Chang, "Effective anomaly space for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5526624.
- [15] C.-I. Chang, "Band sampling for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5514024.
- [16] C.-I. Chang, "An effective evaluation tool for hyperspectral target detection: 3D receiver operating characteristic curve analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 6, pp. 5131–5153, Jun. 2021.
- [17] C.-I. Chang, "Comprehensive analysis of receiver operating characteristic (ROC) curves for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5541124.
- [18] J. M. Molero, E. M. Garzón, I. García, and A. Plaza, "Analysis and optimizations of global and local versions of the RX algorithm for anomaly detection in hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 801–814, Apr. 2013.
- [19] S. Matteoli, T. Veracini, M. Diani, and G. Corsini, "A locally adaptive background density estimator: An evolution for RX-based anomaly detectors," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 323–327, Jan. 2014.
- [20] H. Kwon, S. Z. Der, and N. M. Nasrabadi, "Adaptive anomaly detection using subspace separation for hyperspectral imagery," *Opt. Eng.*, vol. 42, no. 11, pp. 3342–3351, Nov. 2003.
- [21] W. Li and Q. Du, "Collaborative representation for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1463–1474, Mar. 2015.
- [22] W.-M. Liu and C.-I. Chang, "Multiple-window anomaly detection for hyperspectral imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 644–658, Apr. 2013.
- [23] H. Kwon and N. M. Nasrabadi, "Kernel RX-algorithm: A nonlinear anomaly detector for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 2, pp. 388–397, Feb. 2005.
- [24] J. Qu, Q. Du, Y. Li, L. Tian, and H. Xia, "Anomaly detection in hyperspectral imagery based on Gaussian mixture model," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 11, pp. 9504–9517, Nov. 2021.
- [25] N. Ma, Y. Peng, S. Wang, and P. Leong, "An unsupervised deep hyperspectral anomaly detector," *Sensors*, vol. 18, no. 3, p. 693, Feb. 2018, doi: 10.3390/s18030693.
- [26] W. Xie, B. Liu, Y. Li, J. Lei, and Q. Du, "Autoencoder and adversarial-learning-based semisupervised background estimation for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 8, pp. 5416–5427, Aug. 2020.
- [27] T. Jiang, Y. Li, W. Xie, and Q. Du, "Discriminative reconstruction constrained generative adversarial network for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 7, pp. 4666–4679, Jul. 2020.
- [28] J. Zhong, W. Xie, Y. Li, J. Lei, and Q. Du, "Characterization of background-anomaly separability with generative adversarial network for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 7, pp. 6017–6028, Jul. 2021.
- [29] G. Fan, Y. Ma, X. Mei, F. Fan, J. Huang, and J. Ma, "Hyperspectral anomaly detection with robust graph autoencoders," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5511314.
- [30] N. Vaswani, T. Bouwmans, S. Javed, and P. Narayanamurthy, "Robust subspace learning: Robust PCA, robust subspace tracking, and robust subspace recovery," *IEEE Signal Process. Mag.*, vol. 35, no. 4, pp. 32–55, Jul. 2018.
- [31] S. Chang, B. Du, and L. Zhang, "BASO: A background-anomaly component projection and separation optimized filter for anomaly detection in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 7, pp. 3747–3761, Jul. 2018.
- [32] L. Li, W. Li, Y. Qu, C. Zhao, R. Tao, and Q. Du, "Prior-based tensor approximation for anomaly detection in hyperspectral imagery," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 3, pp. 1037–1050, Mar. 2022.
- [33] S. Chen, C.-I. Chang, and X. Li, "Component decomposition analysis for hyperspectral target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5516222.
- [34] S. Wang, X. Wang, L. Zhang, and Y. Zhong, "Auto-AD: Autonomous hyperspectral anomaly detection network based on fully convolutional autoencoder," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5503314.
- [35] D. Chong, B. Hu, X. Gao, H. Gao, P. Xia, and Y. Wu, "Hyperspectral deep convolution anomaly detection based on weight adjustment strategy," *Appl. Opt.*, vol. 59, no. 31, pp. 9633–9642, 2020.
- [36] L. Li, W. Li, Q. Du, and R. Tao, "Low-rank and sparse decomposition with mixture of Gaussian for hyperspectral anomaly detection," *IEEE Trans. Cybern.*, vol. 51, no. 9, pp. 4363–4372, Sep. 2021.
- [37] Y. Zhang, B. Du, L. Zhang, and S. Wang, "A low-rank and sparse matrix decomposition-based Mahalanobis distance method for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1376–1389, Mar. 2016.
- [38] S. Theodoridis and K. Koutroubas, *Pattern Recognition*. New York, NY, USA: Academic, 1999, p. 366.
- [39] A. Hyvärinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," *Neural Comput.*, vol. 9, no. 7, pp. 1483–1492, Jul. 1997.
- [40] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
- [41] R. G. Baraniuk, "Compressive sensing," *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 116–124, Jul. 2007.

- [42] J. C. Harsanyi, W. Farrand, and C.-I. Chang, "Detection of subpixel spectral signatures in hyperspectral image sequences," in *Proc. Amer. Congr. Surveying Mapping (ACSM)/Amer. Soc. Photogramm. Remote Sens. (ASPRS) Annu. Conv. Expo.*, vol. 1, Baltimore, MD, USA, 1994, pp. 236–247.
- [43] C.-I. Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*. Norwell, MA, USA: Kluwer Academic, 2003.
- [44] C.-I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, pp. 608–619, Mar. 2004.
- [45] C.-I. Chang, "A review of virtual dimensionality for hyperspectral imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 4, pp. 1285–1305, Apr. 2018.
- [46] X. Kang, S. Li, and J. A. Benediktsson, "Spectral-spatial hyperspectral image classification with edge-preserving filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2666–2677, May 2014.
- [47] C.-I. Chang, "Constrained energy minimization anomaly detection for hyperspectral imagery via dummy variable trick," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5517119.
- [48] J. C. Harsanyi, "Detection and classification of subpixel spectral signatures in hyperspectral image sequences," Ph.D. dissertation, Dept. Elect. Eng., Univ. Maryland Baltimore County, Baltimore, MD, USA, 1993.
- [49] W. Farrand, "Mapping the distribution of mine tailings in the Coeur d'Alene River Valley, Idaho, through the use of a constrained energy minimization technique," *Remote Sens. Environ.*, vol. 59, no. 1, pp. 64–76, Jan. 1997.
- [50] C.-I. Chang, "Target signature-constrained mixed pixel classification for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 5, pp. 1065–1081, May 2002.
- [51] Z. Zou and Z. Shi, "Hierarchical suppression method for hyperspectral target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 330–342, Jan. 2016.
- [52] R. Zhao, Z. Shi, Z. Zou, and Z. Zhang, "Ensemble-based cascaded constrained energy minimization for hyperspectral target detection," *Remote Sens.*, vol. 11, no. 11, p. 1310, Jun. 2019, doi: [10.3390/rs11111310](https://doi.org/10.3390/rs11111310).
- [53] C.-I. Chang, S. Chen, S. Zhong, and Y. Shi, "Exploration of data scene characterization and 3D ROC evaluation for hyperspectral anomaly detection," *Remote Sens.*, vol. 16, no. 1, p. 135, 2024, doi: [10.3390/rs16010135](https://doi.org/10.3390/rs16010135).
- [54] H. Su, Z. Wu, H. Zhang, and Q. Du, "Hyperspectral anomaly detection: A survey," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 1, pp. 64–90, Mar. 2022.



Chein-I Chang (Life Fellow, IEEE) received the Ph.D. degree in electrical engineering from the University of Maryland, College Park, MD, USA, in 1987.

He is currently a Professor with the University of Maryland, Baltimore County (UMBC), Baltimore, MD, USA. He has been a Chang Jiang Scholar Chair Professor and the Director of the Center for Hyperspectral Imaging in Remote Sensing (CHIRS), Dalian Maritime University, Dalian, China, since 2016. He is also a Yushan Scholar Chair Professor

with National Cheng Kung University, Tainan, Taiwan. He has been a Chair Professor with National Yang Ming Chiao Tung University, Hsinchu, Taiwan,

since 2019. He has authored four books, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification* (Kluwer Academic Publishers, 2003), *Hyperspectral Data Processing: Algorithm Design and Analysis* (John Wiley & Sons, 2013), *Real-Time Progressive Hyperspectral Image Processing: Endmember Finding and Anomaly Detection* (Springer, 2016), and *Real-Time Recursive Hyperspectral Sample and Band Processing: Algorithm Architecture and Implementation* (Springer, 2017), and also edited three books, *Recent Advances in Hyperspectral Signal and Image Processing* (2006), *Hyperspectral Data Exploitation: Theory and Applications* (John Wiley & Sons, 2007), and *Advances in Hyperspectral Image Processing Techniques* (Wiley, 2023). In addition, he also coedited with A. Plaza the book titled *High Performance Computing in Remote Sensing* (CRC Press, 2007). His research interests include multispectral/hyperspectral image processing, automatic target recognition, and medical imaging.

Dr. Chang is a fellow of SPIE. He is serving as an Associate Editor for *Remote Sensing* and IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.



Chien-Yu Lin received the B.S. degree in electrical engineering from National Central University, Taoyuan, Taiwan, in 2013, the M.S. degree in electrical engineering from National Cheng Kung University, Tainan, Taiwan, in 2015, and the M.S. degree in electrical engineering from the University of Maryland, Baltimore County (UMBC), Baltimore, MD, USA, in May 2023, where she is currently pursuing the Ph.D. degree in electrical engineering with the Remote Sensing Signal and Image Processing Laboratory.

Her research interests include hyperspectral imaging, remote sensing, medical data analysis, machine learning, and digital signal processing.



Peter Fuming Hu received the B.S. degree in electrical engineering from Shanghai University, Shanghai, China, in 1984, and the M.S. and Ph.D. degrees in computer science from the University of Maryland, Baltimore County (UMBC), Baltimore, MD, USA, in 1992 and 2013, respectively.

He has been with the University of Maryland School of Medicine, Baltimore, since 1992, where he is currently a Professor with the Department of Anesthesiology, Surgery and Epidemiology. He is also an Adjunct Professor with UMBC. He is also

the Chief Technologist with the Shock Trauma Center, University of Maryland, Baltimore. He has more than 25 years of experience in clinical research at the University of Maryland School of Medicine and the Shock Trauma Center. During that time, he worked as a principal investigator (PI), a Co-PI, or a co-investigator on over 50 successfully completed peer-reviewed grants funded by various U.S. government agencies. His research interests include applying big data machine learning (AI) methods for the prediction of trauma patient outcomes.

Dr. Hu is also a member of the Editorial Board of *Critical Care Medicine* that is the leading clinical journal dedicated to intensive care unit (ICU) patient care.