Parametric Adaptive Signal Detection for Hyperspectral Imaging

Hongbin Li, *Member, IEEE,* and James H. Michels, *Fellow, IEEE*

***Abstract*—Hyperspectral imaging (HSI) sensors can provide very fine spectral resolution that allows remote identification of ground objects smaller than a full pixel in an HSI image. Traditional approaches to the so-called subpixel target signal detection problem are *training inefficient* due to the need for an estimate of a large-size covariance matrix of the background from target-free training pixels. This imposes a training requirement that is often difficult to meet in a heterogeneous environment. In this paper, a class of training-efficient adaptive signal detectors is presented by exploiting a parametric model that takes into account the nonstationarity of HSI data in the spectral dimension. A maximum-likelihood (ML) estimator is developed to estimate the parameters associated with the proposed parametric model. Several important issues are discussed, including model order selection, training screening, and time-series-based whitening and detection, which are intrinsic parts of the proposed parametric adaptive detectors. Experimental results using real HSI data reveal that the proposed parametric detectors are more training efficient and outperform conventional covariance-matrix-based detectors when the training size is limited.**

***Index Terms*—Adaptive signal detection, hyperspectral imaging (HSI), nonstationarity, parameter estimation.**

I. Introduction

HYPERSPECTRAL sensors are a new class of imaging spectroscopy sensors that divide the waveband of interest into hundreds of contiguous narrow bands. Their fine spectral resolution enables remote identification of ground objects based on their spectral signatures. Hyperspectral imaging (HSI) has a wide range of applications, including terrain classification, environmental and agricultural monitoring, geological exploration, ordinance remediation, tactical surveillance, and others [1].

A challenging problem in HSI applications is the so-called *subpixel target detection*, which involves detecting objects occupying only a portion of a full pixel in an HSI image [2]. In such a case, the signal produced by the HSI sensors consists of both the *object* and *background*, with the latter behaving effectively as interference that has to be suppressed for effective detection. The problem is reminiscent of that of detecting a known signal with unknown amplitude in colored noise with unknown correlation[1] (e.g., [5]). A multitude of solutions have been developed, including the Kelly's generalized-likelihood ratio (GLR) test [6], the adaptive matched filter (AMF) [7], and the adaptive coherence estimator (ACE) test [8], [9], among others. While these detectors can be used to solve the HSI subpixel target detection problem, there is a major difficulty with them in training-limited scenarios. In particular, the above detectors are *covariance-matrix-based techniques* in that they all rely on an estimate of the background covariance matrix, which is obtained from *target-free* training pixels. The size of the background covariance matrix is identical to the number of spectral bands that is typically in the order of hundreds. A good estimate of the covariance matrix would require several hundred or more target-free training pixels, which may not be available in heterogeneous or dense-target environments. Another problem with the above covariance-matrix-based detectors is complexity, since the large-size covariance matrix has to be estimated and inverted frequently.

There is a significant interest in developing *training-efficient* detection techniques for training-limited applications, such as the above HSI target detection approaches applied in heterogeneous environments. Another example is target detection based on space–time adaptive processing (STAP) for airborne radars [10], where range-dependent clutter characteristics, along with other issues, prevent inclusion of a large number of range cells far away from the test cell in the training set. One effective way to reduce training requirement in STAP detection is to utilize a suitable parametric model for the radar clutter and exploit the model for target detection. In particular, multichannel autoregressive (AR) models have been found to be very effective in representing the temporal correlation among pulse returns [11]–[16]. A parametric detector based on such a multichannel AR clutter model is developed in [11]–[14], which is referred to as the parametric adaptive matched filter (PAMF). The PAMF detector has been shown to significantly outperform the covariance-matrix-based detectors for small training size.

For HSI applications, however, the data is nonstationary in the spectral domain (see Section IV-A for details of such nonstationarity),[2] whereas AR models are by definition stationary. To account for such nonstationarity, we introduce in this paper a sliding-window-based nonstationary AR (NS-AR) model to capture the spectral correlation of HSI data. We propose a class of parametric adaptive signal detectors for HSI subpixel

Manuscript received March 13, 2005; revised July 11, 2005. This work was supported by the Air Force Research Laboratory (AFRL) Visiting Faculty Research Program (VFRP). Part of this work was presented at the 2004 International Conference on Acoustic, Speech, and Signal Processing (ICASSP 2004), Montreal, Canada, May 2004. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Fulvio Gini.

H. Li is with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030 USA (e-mail: hli@stevens.edu).

J. H. Michels is with JHM Technologies, Ithaca, NY 14852-4142 USA (e-mail: jmichels@americu.net).

Digital Object Identifier 10.1109/TSP.2006.873589

[1]We take a stochastic approach herein by modeling the background as a correlated random vector with an unknown covariance matrix. There are other detectors based on modeling the background as a deterministic quantity. (See [2]–[4] and references therein for details).

[2]Such spectral nonstationarity shall not be confused with the spatial stationarity that is often assumed for HSI data [2].

target detection, and develop a maximum-likelihood (ML) estimation algorithm to estimate the parameters associated with the NS-AR model. In addition, we develop model order selection, training screening, and time-series-based whitening and detection techniques, which are intrinsic parts of the proposed parametric adaptive detectors. We show via experimental results with real HSI data that our proposed parametric detectors are more efficient in training usage and outperform the conventional covariance-matrix-based detectors when the training size is limited.

The rest of the paper is organized as follows. Section II contains the data model and problem statement. The covariance-matrix-based detectors are briefly reviewed and discussed in Section III. The proposed techniques, including an NS-AR model, a class of parametric adaptive detectors, an ML parameter estimation algorithm, a model order selection method, and a training screening approach, are detailed in Section IV. Experimental results illustrating the performance of the proposed detectors under homogeneous, heterogeneous, and dense-target environments are presented in Section V. Finally, Section VI contains our concluding remarks.

II. Data Model and Problem Statement

Obtained through both spatial and spectral sampling, HSI data is usually described as a data cube, whose face is a function of the spatial coordinates and depth is a function of spectral bands or wavelengths. Each pixel can be represented as an $L \times 1$ *real-valued* vector: $\boldsymbol{x} = [x(0), x(1), \ldots, x(L-1)]^T$, where L denotes the total number of spectral bands, $x(l)$ denotes the spectral response at the lth spectral band, and $(\cdot)^T$ denotes transpose. Since HSI data has nonzero mean [2], [17], a preprocessing stage is usually invoked to remove the sample mean estimated using the neighbor pixels.

In vector notation, the *subpixel signal detection* problem is described by the following composite hypothesis test [2]:

$$\begin{aligned} H_0 &: \boldsymbol{x} = \boldsymbol{b}, \quad \text{target absent} \\ H_1 &: \boldsymbol{x} = a\boldsymbol{s} + \boldsymbol{b}, \quad \text{target present} \end{aligned} \tag{1}$$

where $\boldsymbol{x} \in \mathbb{R}^{L\times 1}$ is the demeaned test pixel, $\boldsymbol{s} \in \mathbb{R}^{L\times 1}$ is the *signature vector* of the target object with amplitude a, and $\boldsymbol{b} \in \mathbb{R}^{L\times 1}$ denotes the background plus system noise. We adopt the standard assumption that the signature vector $\boldsymbol{s}$ is deterministic and known to the detector[3]; the amplitude a, however, is assumed unknown. For the background, we follow a *statistical approach* that models the background interference $\boldsymbol{b}$ as a multivariate Gaussian random vector with zero mean and an unknown covariance matrix $\boldsymbol{R}_b \triangleq E\{\boldsymbol{b}\boldsymbol{b}^T\}$. The Gaussian assumption has been widely used for multispectral (e.g., [17]) and HSI data [2]. It leads to mathematical tractability and reasonably good performance. Nevertheless, it should be noted that a Gaussian model is not fully appropriate to characterize the statistical behavior of HSI data in many realistic cases, and alternative modeling approaches have been considered in [18]–[20].

Equation (1) implies that the background interference covariance matrix is the same under both hypotheses. Since for a subpixel target the area covered by background is different under the two hypotheses, it is more appropriate to consider the following modified hypothesis [2], [21]:

$$\begin{aligned} H_0 &: \boldsymbol{x} = \boldsymbol{b}, \quad \text{target absent} \\ H_1 &: \boldsymbol{x} = a\boldsymbol{s} + \sigma\boldsymbol{b}, \quad \text{target present} \end{aligned} \tag{2}$$

where σ is unknown and, along with the signature amplitude a, determined by the target *fill factor*, i.e., the percentage of the pixel area occupied by the target [2].

Similar to [2], we assume that in addition to the test pixel $\boldsymbol{x}$, we have N training pixels $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$. In surveillance applications when the target class is rare or sparsely populated, the training pixels are usually taken as those surrounding the test pixel and assumed target free [2]. Again similar to [2], we assume that $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ are independent and identically distributed (i.i.d.) Gaussian random vectors with zero mean and covariance matrix $\boldsymbol{R}_b$, and independent of the text pixel $\boldsymbol{x}$.

The problem in question is to find an efficient decision rule for the composite hypothesis testing problem (1) or (2), given knowledge of the test pixel $\boldsymbol{x}$, target signal signature $\boldsymbol{s}$, and training pixels $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$. Our goal is to achieve good detection performance for small N.

Before closing this section, we remark that our parametric detection schemes, as well as many others (e.g., in [2]), rely on the perfect knowledge of the target spectral signature. Generally, the target signature is available in its *reflectance spectrum*, whereas the HSI sensors measure the *radiance spectrum* of the observed materials. In order to apply these detection techniques, the HSI data must be preprocessed to obtain reflectance data from the radiance ones (e.g., through atmospheric correction) or, alternatively, target spectral reflectance must be processed to obtain the radiance spectrum. See [22], [23] for details.

III. Covariance-Matrix-Based Solutions

A number of solutions to the above problem have been developed. If the covariance matrix $\boldsymbol{R}_b$ is known exactly, the optimum detector for (1) with unknown signal amplitude is the matched filter (MF) [7]

$$\frac{|\boldsymbol{s}^T\boldsymbol{R}_b^{-1}\boldsymbol{x}|^2}{\boldsymbol{s}^T\boldsymbol{R}_b^{-1}\boldsymbol{s}} \underset{H_0}{\overset{H_1}{\gtrless}} t_{\mathrm{MF}} \tag{3}$$

where t_{MF} denotes the MF threshold. The MF detector is obtained by a GLR approach (e.g., [5]), by which the ML estimate of the unknown amplitude a is first estimated and then substituted back into the likelihood ratio to form a test statistic. In practice, the MF detector *cannot* be implemented since $\boldsymbol{R}_b$ is typically unknown. However, it provides a baseline for performance comparison when considering any realizable detection scheme.

[3]The spectral signature may vary due to variations in atmospheric conditions and other factors, and the uncertainty can be captured by a linear mixing model [2]. We do not consider such spectral variations since our focus is effective cancellation of the background.

In practice, the unknown $\boldsymbol{R}_b$ can be replaced by some estimate, such as the sample covariance matrix obtained from the training pixels

$$\hat{\boldsymbol{R}}_b = \frac{1}{N}\sum_{n=1}^{N} \boldsymbol{x}_n \boldsymbol{x}_n^T. \tag{4}$$

Using $\hat{\boldsymbol{R}}_b$ in (3) leads to the so-called AMF detector [7]

$$\frac{\left|\boldsymbol{s}^T \hat{\boldsymbol{R}}_b^{-1} \boldsymbol{x}\right|^2}{\boldsymbol{s}^T \hat{\boldsymbol{R}}_b^{-1} \boldsymbol{s}} \underset{H_0}{\overset{H_1}{\gtrless}} t_{\mathrm{AMF}} \tag{5}$$

where t_{AMF} denotes the AMF threshold.

Alternatively, one can treat both a and $\boldsymbol{R}_b$ as unknowns and estimate them successively by ML. Such a GLR approach was pursued by Kelly [6], which gives the following Kelly test:

$$\frac{\left|\boldsymbol{s}^T \hat{\boldsymbol{R}}_b^{-1} \boldsymbol{x}\right|^2}{\left(\boldsymbol{s}^T \hat{\boldsymbol{R}}_b^{-1} \boldsymbol{s}\right)\left(N + \boldsymbol{x}^T \hat{\boldsymbol{R}}_b^{-1} \boldsymbol{x}\right)} \underset{H_0}{\overset{H_1}{\gtrless}} t_{\mathrm{Kelly}} \tag{6}$$

where t_{Kelly} denotes the corresponding threshold.

Another popular detector is the ACE test [8], [9]

$$\frac{\left|\boldsymbol{s}^T \hat{\boldsymbol{R}}_b^{-1} \boldsymbol{x}\right|^2}{\left(\boldsymbol{s}^T \hat{\boldsymbol{R}}_b^{-1} \boldsymbol{s}\right)\left(\boldsymbol{x}^T \hat{\boldsymbol{R}}_b^{-1} \boldsymbol{x}\right)} \underset{H_0}{\overset{H_1}{\gtrless}} t_{\mathrm{ACE}} \tag{7}$$

which is obtained by a GLR procedure that takes into account not only the unknown amplitude a and background covariance matrix $\boldsymbol{R}_b$, but also the variability of the variance of the background under H_0 and H_1. Interestingly, the ACE test is the AMF test (5) normalized by the signal energy weighted by the covariance matrix inverse $\hat{\boldsymbol{R}}_b^{-1}$. By the Schwartz inequality, one can see that the ACE test statistic is bounded between zero and one.

The AMF, Kelly, and ACE tests have constant false alarm rate (CFAR). However, they entail a large training requirement. The covariance matrix $\boldsymbol{R}_b$ has a dimension of $L \times L$. Typical values for L in real HSI systems are in the range of hundreds. An accurate estimate of the covariance matrix would require a large number of target-free training pixels, which may not be available, especially in nonhomogeneous environments. In addition, the computational complexity of these detectors is high, since $\boldsymbol{R}_b$ has to be estimated and inverted frequently.

IV. Proposed Approach

In this section, we present a class of parametric adaptive signal detectors with reduced training requirement. The proposed detectors, which are detailed in Section IV-B, relies on an NS-AR model introduced in Section IV-A, an ML parameter estimation algorithm derived in Section IV-C, a model order selection method discussed in Section IV-D, and a training screening technique presented in Section IV-E.

A. Parametric Modeling of HSI Data

It is well known that the interference suppression ability of the detectors discussed in Section III comes from a whitening procedure. Consider, for example, the AMF detector (5). The whitening operation takes as inputs the signature vector $\boldsymbol{s}$ and test pixel $\boldsymbol{x}$, and outputs whitened versions

$$\tilde{\boldsymbol{s}} \triangleq \hat{\boldsymbol{R}}_b^{-1/2} \boldsymbol{s}, \quad \tilde{\boldsymbol{x}} \triangleq \hat{\boldsymbol{R}}_b^{-1/2} \boldsymbol{x} \tag{8}$$

where $\hat{\boldsymbol{R}}_b^{-1/2}$ denotes the matrix square-root of $\hat{\boldsymbol{R}}_b^{-1}$. Following the whitening, the AMF detector reduces to simple correlation of the whitened outputs

$$\frac{\left|\tilde{\boldsymbol{s}}^T \tilde{\boldsymbol{x}}\right|^2}{\tilde{\boldsymbol{s}}^T \tilde{\boldsymbol{s}}} \underset{H_0}{\overset{H_1}{\gtrless}} t_{\mathrm{MF}}. \tag{9}$$

If the whitening operation can be designed or approximated via a parametric model without explicitly estimating $\boldsymbol{R}_b$, then it is conceivable that fewer training pixels are needed, provided that the parametric model is parsimonious enough (without an extraordinary number of parameters). This is the essence of our parametric-model-based methods. Next, we consider two different parametric modeling approaches.

1) AR Modeling: AR models have been popular choices for parametric modeling in spectral analysis, speech coding, wireless channel modeling, seismic signal processing, among others (e.g., [24]). Parametric adaptive detection based on multichannel AR models has been considered in [11]–[14], [25], [26] for airborne radar systems equipped with multiple antennas. It was shown that significant saving in training and complexity can be achieved by fitting the interference and radar clutter into suitable multichannel AR models.

For the problem under study, the $L \times 1$ background vector $\boldsymbol{b}$, or equivalently the observed signal $\boldsymbol{x}$ under H_0, may be assumed to be a scalar AR process which produces the L samples of $\boldsymbol{b}$. If an AR model is appropriate for HSI data, then the detection problem amounts to first estimating the AR coefficients from training data, whitening the signals by a whitening filter constructed from the AR coefficient estimates, and computing the decision statistic from the whitened signals followed by thresholding. For brevity, the above approach is referred to as the *parametric adaptive matched filter* (PAMF),[4] or *normalized PAMF* (NPAMF) [25] if the decision variable is normalized, similar to the normalization imposed by the ACE detector of (7).

We have tested the above AR-based PAMF/NPAMF detectors with real HIS data using fixed AR parameters across the spectral domain and found they suffer a performance loss compared to the methods proposed here. The reason is that AR models are not a suitable parametric model for HSI data. In particular, we find that HSI data are *nonstationary in the spectral dimension*, whereas fixed parameter AR models characterize *stationary* random processes. To see this, we have computed the sample covariance matrix $\hat{\boldsymbol{R}}_b$ from a total of $K = 24 \times 46 = 1104$ training pixels drawn from a homogeneous region of the HSI data described in Section V. Fig. 1 depicts the main and three subdiagonals of $\hat{\boldsymbol{R}}_b$, which correspond to the autocorrelation function (ACF) at spectral lag 0 (i.e., variance), lag 1, lag 2 and lag 3, respectively, versus the spectral bands. Clearly, the signal is not stationary since the variance and ACF at other lags vary significantly across the spectral bands.

[4]Details of the PAMF and NPAMF detectors can be inferred from the proposed NS-PAMF and NS-NPAMF detectors discussed in Section IV-B, as the former are special cases of the latter.

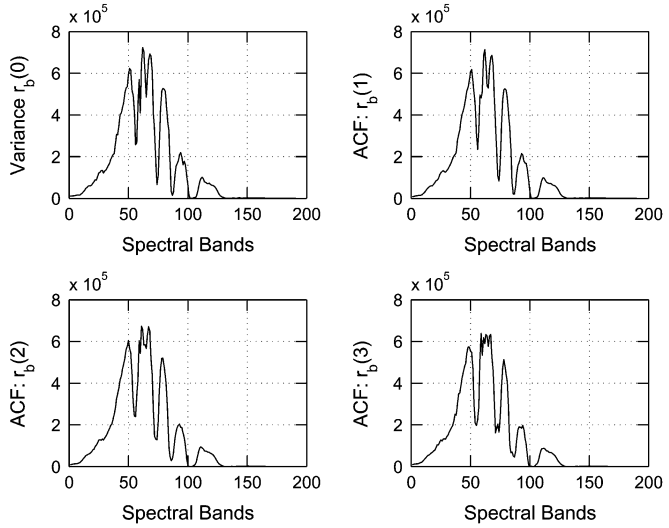


Fig. 1. Sample estimates of the autocorrelation function at spectral lag 0 (variance), lag 1, lag 2, and lag 3 across the spectral bands.

2) NS-AR Modeling: Although HSI data is nonstationary (NS) across the entire spectral dimension, it may be considered approximately stationary over a sufficiently small number of adjacent spectral bands. This can seen from Fig. 1, where the variation of the sample statistics over a few adjacent spectral bands is considerably smaller compared with that over the entire spectral bands. In the following, we consider a NS-AR modeling approach by taking into account such *local stationarity* of HSI data. Specifically, let $x_n(l)$ denote the spectral response at the lth spectral band of the nth training pixel $\boldsymbol{x}_n$, that is, $\boldsymbol{x}_n \triangleq [x_n(0), \ldots, x_n(L-1)]^T$. Then, we slice $\boldsymbol{x}_n$ into $L - L_s + 1$ overlapping subvectors

$$\boldsymbol{x}_{n,l} \triangleq [x_n(l), \ldots, x_n(l + L_s - 1)]^T, \quad l = 0, \ldots, L - L_s \quad (10)$$

where $L_s \leq L$ denotes the length of the subvectors. Equivalently, these subvectors can be thought of as being obtained by windowing $\boldsymbol{x}_n$ using a sliding window of size L_s. For sufficiently small L_s, each subvector $\boldsymbol{x}_{n,l}$ can be modeled as an Mth-order AR process

$$x_n(k) = -\sum_{m=1}^{M} a_l(m) x_n(k-m) + w_{n,l}(k), \quad k = l, l+1, \ldots, l+L_s-1; \quad n = 1, \ldots, N \quad (11)$$

where $w_{n,l}(k)$ denotes the modeling residual for the lth subvector $\boldsymbol{x}_{n,l}$. The residual is Gaussian (since $x_n(k)$ is so) with zero mean and variance σ_l^2, and spectrally white so that $\{w_{n,l}(k)\}$ are independent with respect to k and n [24]. Note that the lth set of the AR coefficients $a_l(1), \ldots, a_l(M)$ is associated with the lth subvector $\boldsymbol{x}_{n,l}$ and that different subvectors are associated with different sets of AR coefficients. For simplicity, we consider fixed AR model order M (also see discussions in Section IV-D).

From the estimation perspective, the choice of M and window size L_s should be made with tradeoffs among the bias, variance, and stationarity of the modeling approach. A large M might be desirable since it can provide better fitting (lower bias) to the HSI data. Increasing M, however, would require the window size L_s to increase accordingly since more parameters are to be estimated, and, therefore, more data should be provided within each subvector to reduce the variance of parameter estimates. If L_s is too large, the assumption of stationarity within the subvector may be violated, which can cause significant degradation. From the application aspect, however, these parameters are related to the HSI sensor characteristics, such as the operating spectral range, spectral resolution, etc. For the HSI data used in this paper, we found a window size $8 \leq L_s \leq 15$ is generally appropriate for modeling. Once L_s is selected, we can use information-criterion-based model order selection techniques to determine M. We leave the details to Section IV-D.

Instead of the above sliding-window-based NS-AR modeling approach, one can consider an alternative NS-AR model that models the HSI data across all the spectral bands

$$x_n(l) = -\sum_{m=1}^{M} b_l(m) x_n(l-m) + v_n(l), \quad l = 0, \ldots, L-1 \quad (12)$$

where $b_l(m)$ denotes the shift-varying AR coefficient and $v_n(l)$ the fitting error of the lth sample. Note that the above model differs from (11) in that the AR coefficients are varying from sample to sample, whereas in (11), they are assumed to remain fixed within a sliding subvector of L_s samples. An additional parametric model for the shift-varying AR coefficients $\{b_l(m)\}$ is necessary to ensure they can be estimated. This *doubly* parametric approach is more sensitive to the choice of the parameters, whose estimation is also considerably more involved. In the following, we consider only the sliding-window-based NS-AR modeling approach.

B. NS-AR Model Based Parametric Adaptive Detectors

If the above NS-AR model (11) is appropriate for modeling target-free HSI data (i.e., the background), then a time-series-based (as opposed to the previous covariance-matrix-based) whitening process can be developed without explicitly estimating $\boldsymbol{R}_b$. This leads to a class of parametric adaptive detectors that are summarized below:

- *Step 1—Parameter Estimation:* Estimate the NS-AR coefficients $\{a_l(m)\}$ in (11) and the variance $\{\sigma_l^2\}$ of the residual from the training pixels $\{\boldsymbol{x}_n\}_{n=1}^N$ by using an ML-based estimation algorithm detailed in Section IV-C. Let $\{\hat{a}_l(m), \hat{\sigma}_l^2\}$ denote the coefficient estimates.
- *Step 2—Whitening:* Form a shift-varying moving-average (MA) whitening filter from the parameter estimates $\{\hat{a}_l(m), \hat{\sigma}_l^2\}$, and whiten the test pixel $\boldsymbol{x}$ and target signature $\boldsymbol{s}$ as follows:

$$\breve{x}(l) = \frac{1}{\hat{\sigma}_l}\left[x(l) + \sum_{m=1}^{M} \hat{a}_{l-L_s}(m) x(l-m)\right]$$
$$\breve{s}(l) = \frac{1}{\hat{\sigma}_l}\left[s(l) + \sum_{m=1}^{M} \hat{a}_{l-L_s}(m) s(l-m)\right], \quad l = L_s - 1, \ldots, L-1 \quad (13)$$

where $\breve{x}(l)$ and $\breve{s}(l)$ denote the lth output sample of the whitening filter when the input is the test pixel $\boldsymbol{x}$ and target signature $\boldsymbol{s}$, respectively. It should be noted from

(13) that each set of the NS-AR parameter estimates, i.e., $\{\hat{a}_l(m)\}_{m=1}^{M}$ and $\hat{\sigma}_l$, is used to compute one pair of output samples $\breve{x}(l)$ and $\breve{s}(l)$; as the sliding window shifts to the next position, we use the next set of parameter estimates for whitening. In effect, (13) implements the whitening operation (8) in a time-series fashion by taking into account the NS nature of the signal. For an input of L spectral samples, the time-series-based whitening filter outputs $L - L_s$ whitened samples due to initialization of the whitening filter. Although such dimensionality reduction may affect the detection performance, the impact is negligible for small L_s and large L, which is typical in HSI systems.

- *Step 3—Detection:* The outputs of the shift-varying whitening filter corresponding to the test pixel $\boldsymbol{x}$ and target signature $\boldsymbol{s}$, respectively, are used to form the decision statistic. Depending upon how the decision statistic is formed, we have a class of parametric detectors. For example, the parametric counterparts of the covariance-matrix-based AMF (5) and ACE (7) detectors are given by

$$\frac{\left|\sum_{l=L_s-1}^{L-1}\breve{s}(l)\breve{x}(l)\right|^2}{\sum_{l=L_s-1}^{L-1}\breve{s}^2(l)} \underset{H_0}{\overset{H_1}{\gtrless}} t_{\mathrm{NS-PAMF}} \tag{14}$$

$$\frac{\left|\sum_{l=L_s-1}^{L-1}\breve{s}(l)\breve{x}(l)\right|^2}{\left(\sum_{l=L_s-1}^{L-1}\breve{s}^2(l)\right)\left(\sum_{l=L_s-1}^{L-1}\breve{x}^2(l)\right)} \underset{H_0}{\overset{H_1}{\gtrless}} t_{\mathrm{NS-NPAMF}} \tag{15}$$

which are referred to as the *NS-PAMF* and *NS-NPAMF* detectors, respectively. A time-series-based, Kelly-like test can also be obtained in a similar fashion.

Finally, it is noted that the above NS-PAMF and NS-NPAMF detectors reduce to the PAMF and NPAMF detectors, respectively, which are briefly discussed in Section IV-A-1), when $L_s = L$, that is, the sliding window reaches the maximum value and includes the entire spectral bands. In that case, the NS-AR model in (11) reduces to the standard stationary AR model.

C. ML Estimation of NS-AR Coefficients

Parameter estimation plays a critical role for the proposed parametric detectors. In this section, we present an ML estimator to estimate the NS-AR coefficients in (11) using training pixels $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$. Our ML estimator is an extension of that in [24] for fixed AR models to NS-AR processes.

Consider the vector of AR coefficients of model order M:$\boldsymbol{a}_l \triangleq [a_l(1), \ldots, a_l(M)]^T$. According to the statistical assumptions made in Section II and the NS-AR model (11), the lth set of subvectors $\boldsymbol{x}_{1,l}, \ldots, \boldsymbol{x}_{N,l}$ formed from the N training pixels are i.i.d. multivariate Gaussian whose joint probability density function (pdf) is parameterized by the NS-AR coefficients $\boldsymbol{a}_l$ and variance σ_l^2. Then, the ML estimates of $\boldsymbol{a}_l$ and σ_l^2 are obtained by maximizing the joint pdf $p(\boldsymbol{x}_{1,l}, \ldots, \boldsymbol{x}_{N,l}; \boldsymbol{a}_l, \sigma_l^2)$. Exact maximization of the joint pdf with respect to the unknown parameters turns out to be highly involved computationally [27]. Instead, we seek to optimize a conditional pdf, which produces an asymptotic ML estimate of the parameters for large date size [24]. Specifically, let

$$\boldsymbol{x}_{n,l}^{(1)} \triangleq [x_n(l), \ldots, x_n(l+M-1)]^T \tag{16}$$

$$\boldsymbol{x}_{n,l}^{(2)} \triangleq [x_n(l+M), \ldots, x_n(l+L_s-1)]^T \tag{17}$$

which collect the first M and, respectively, the last $L_s - M$ samples of $\boldsymbol{x}_{n,l}$. Thus, we have $\boldsymbol{x}_{n,l} = [\boldsymbol{x}_{n,l}^{(1)T}, \boldsymbol{x}_{n,l}^{(2)T}]^T$. Our asymptotic ML estimator seeks to maximize the joint conditional pdf $p(\boldsymbol{x}_{1,l}^{(2)}, \ldots, \boldsymbol{x}_{N,l}^{(2)} | \boldsymbol{x}_{1,l}^{(1)}, \ldots, \boldsymbol{x}_{N,l}^{(1)}; \boldsymbol{a}_l, \sigma_l^2)$ with respect to $\boldsymbol{a}_l$ and σ_l^2. We will write the conditional pdf as $p(\boldsymbol{x}_l^{(2)} | \boldsymbol{x}_l^{(1)}; \boldsymbol{a}_l, \sigma_l^2)$ for brevity.

To find an explicit form of the above conditional pdf, we observe from (11) that

$$w_{n,l}(k) = x_n(k) + \sum_{m=1}^{M} a_l(m) x_n(k-m),$$
$$k = l, l+1, \ldots, l+L_s-1; \quad n = 1, \ldots, N. \tag{18}$$

Since $\{w_{n,l}\}$ are i.i.d. Gaussian with zero mean and variance σ_l^2, we have (e.g., [24])

$$p\left(\boldsymbol{x}_l^{(2)} \middle| \boldsymbol{x}_l^{(1)};\ \boldsymbol{a}_l, \sigma_l^2\right) = (2\pi\sigma_l^2)^{-N(L_s-M)/2} \times \exp\left\{-\frac{1}{2\sigma_l^2}\sum_{n=1}^{N}\cdot\sum_{k=l+M}^{l+L_s-1}\left[x_n(k) + \sum_{m=1}^{M} a_{m,l} x_n(k-m)\right]^2\right\}. \tag{19}$$

Maximizing the above conditional pdf is equivalent to minimizing the negative log likelihood function

$$V(\boldsymbol{a}_l, \sigma_l^2) \triangleq -\ln p\left(\boldsymbol{x}_l^{(2)} \middle| \boldsymbol{x}_l^{(1)}; \boldsymbol{a}_l, \sigma_l^2\right). \tag{20}$$

Define an $(L_s - M) \times M$ matrix

$$\boldsymbol{X}_{n,l} \triangleq \begin{bmatrix} x_n(l+M-1) & \cdots & x_n(l) \\ \vdots & \vdots & \vdots \\ x_n(l+L_s-2) & \cdots & x_n(l+L_s-M-1) \end{bmatrix}. \tag{21}$$

Then, $V(\boldsymbol{a}_l, \sigma_l^2)$ can be more compactly expressed as

$$V(\boldsymbol{a}_l, \sigma_l^2) = C_1 + \frac{1}{2}N(L_s - M)\ln\sigma_l^2 + \frac{1}{2\sigma_l^2}\sum_{n=1}^{N}\left\|\boldsymbol{x}_{n,l}^{(2)} + \boldsymbol{X}_{n,l}\boldsymbol{a}_l\right\|^2 \tag{22}$$

where[5]

$$C_1 \triangleq \frac{1}{2}N(L_s - M)\ln(2\pi). \tag{23}$$

[5]We keep the constant term C_1 which depends on M for model order selection in Section IV-D.

Taking the derivative of $V(\boldsymbol{a}_l, \sigma_l^2)$ with respect to σ_l^2 and setting it to zero yield

$$\hat{\sigma}_l^2(\boldsymbol{a}_l) = \frac{1}{N(L_s - M)} \sum_{n=1}^{N} \left\| \boldsymbol{x}_{n,l}^{(2)} + \boldsymbol{X}_{n,l} \boldsymbol{a}_l \right\|^2. \tag{24}$$

Substituting $\hat{\sigma}_l^2(\boldsymbol{a}_l)$ back into (22) reduces the cost function to

$$V(\boldsymbol{a}_l, \hat{\sigma}_l^2) = C_1 + C_2 + \frac{1}{2} N(L_s - M) \ln \hat{\sigma}_l^2(\boldsymbol{a}_l) \tag{25}$$

where

$$C_2 = \frac{1}{2} N(L_s - M). \tag{26}$$

Therefore, the ML estimate of $\boldsymbol{a}_l$ is obtained by minimizing $\hat{\sigma}_l^2(\boldsymbol{a}_l)$, the variance of the NS-AR modeling residual. The solution is obtained by least-squares fitting

$$\hat{\boldsymbol{a}}_l = -\left(\sum_{n=1}^{N} \boldsymbol{X}_{n,l}^T \boldsymbol{X}_{n,l} \right)^{-1} \left(\sum_{n=1}^{N} \boldsymbol{X}_{n,l}^T \boldsymbol{x}_{n,l}^{(2)} \right), \quad l = 0, 1, \ldots, L - L_s. \tag{27}$$

The matrix within the first pair of brackets is assumed nonsingular. A necessary condition for nonsingularity is that the number of training pixels N is such that

$$N \geq \frac{M}{L_s - M}. \tag{28}$$

This is because the above matrix inverse can be expressed as $(\boldsymbol{X}_l^T \boldsymbol{X}_l)^{-1}$, where

$$\boldsymbol{X}_l \triangleq [\boldsymbol{X}_{1,l}^T, \ldots, \boldsymbol{X}_{N,l}^T]^T \tag{29}$$

is a tall matrix when the above condition is satisfied. On the other hand, when $N \geq (M)/(L_s - M)$, $\boldsymbol{X}_l^T \boldsymbol{X}_l$ is full rank almost surely due to the random nature of the HSI data.

Finally, substituting the ML estimate (27) back into (24) yields the minimum variance of the residual

$$\hat{\sigma}_l^2 = \frac{1}{N(L_s - M)} \boldsymbol{x}_l^{(2)T} \boldsymbol{P}_{\boldsymbol{X}_l}^{\perp} \boldsymbol{x}_l^{(2)} \tag{30}$$

where $\boldsymbol{x}_l^{(2)} \triangleq [\boldsymbol{x}_{1,l}^{(2)T}, \ldots, \boldsymbol{x}_{N,l}^{(2)T}]^T$ and $\boldsymbol{P}_{\boldsymbol{X}_l}^{\perp}$ is the projection matrix onto the null space of $\boldsymbol{X}_l$

$$\boldsymbol{P}_{\boldsymbol{X}_l}^{\perp} = \boldsymbol{I} - \boldsymbol{X}_l \left(\boldsymbol{X}_l^T \boldsymbol{X}_l \right)^{-1} \boldsymbol{X}_l^T \tag{31}$$

where $\boldsymbol{I}$ is an identity matrix.

D. NS-AR Model Order Selection

In this section, we develop information-criterion-based model order selection techniques to determine the NS-AR model order M in (11). Although in principle it is possible to select a different M for each subvector $\boldsymbol{x}_{n,l}, l = 0, \ldots, L - L_s$, by a separate fitting of M to the information criterion, this is a tedious process. In the following, we use a fixed M for all l.

Model order selection for parametric models is a classical topic and has been investigated by various researchers for various models (e.g., [24] and [28] and references therein). We examine herein the model order selection problem for the NS-AR model in (11) for the HSI application, which appears not to have been addressed elsewhere. Specifically, we consider a *generalized Akaike information criterion* (GAIC), which chooses the model order M that minimizes

$$W(M) = \sum_{l=0}^{L-L_s} [V_l(M) + \gamma(M)] \tag{32}$$

where $V_l(M)$ is the minimum cost associated with the lth set of subvectors $\boldsymbol{x}_{1,l}, \ldots, \boldsymbol{x}_{N,l}$, and $\gamma(M)$ is a penalty term that penalizes increasing model order [28]. Specifically, the minimum cost is derived in Section IV-C (cf. (25))

$$V_l(M) = C_1(M) + C_2(M) + \frac{1}{2} N(L_s - M) \ln \hat{\sigma}_l^2(M) \tag{33}$$

where $C_1(M), C_2(M)$, and $\hat{\sigma}_l^2(M)$ are given by (23), (26), and (30), respectively, and the dependence on M is made explicit. On the other hand, the penalty term typically takes the form [28]

$$\gamma(M) = \alpha(M + 1) \ln(NL_s) \tag{34}$$

or

$$\gamma(M) = \alpha(M + 1) \ln[\ln(NL_s)] \tag{35}$$

where $M + 1$ is the total number of unknowns for each set of subvectors $\{\boldsymbol{x}_{n,l}\}_{n=1}^N$, NL_s is the number of data samples contained in $\{\boldsymbol{x}_{n,l}\}_{n=1}^N$, and $\alpha \geq 2$ is a parameter of user choice. Note that the above GAIC reduces to the standard AIC [29] when the $(L - L_s + 1)$-term summation in (32) vanishes and $\gamma(M) = 2(M + 1)$. It is known that AIC is not a consistent model order estimator [24]. Choosing a penalty term proportional to $\ln(NL_s)$ or $\ln[\ln(NL_s)]$ is an effective way of obtaining consistent order estimates [28].

E. Training Screening

One assumption made in Section II is that the N training pixels $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ are target free. This assumption is reasonable in *homogeneous* environments where targets are rare or sparsely populated, but usually violated in *heterogeneous* or *dense-target* environments. In the latter case, the performance of all training-based detectors, including those covariance-matrix-based detectors discussed in Section III, degrade considerably. Training screening to eliminate "bad" training data in such cases has been examined in a number of recent studies for radar target detection (e.g., [30]–[32] and references therein). In this section, we discuss screening of heterogeneous HSI training data. Rather than treating it as an independent process, we cast training screening within the proposed NS-AR framework.

For covariance-matrix-based detectors in Section III, one screening approach according to statistical ranking and selection theory is to compute the following metric from the training set [30]:

$$T_n = \boldsymbol{x}_n^T \hat{\boldsymbol{R}}_b^{-1} \boldsymbol{x}_n, \quad n = 1, \ldots, N. \tag{36}$$

Then, the metric is used to partition the training set $\mathcal{S} \triangleq \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ into two disjoint sets $\mathcal{S}_1$ and $\mathcal{S}_2$ (see [30] for details), of which the former contains the refined training data while the latter contains *outliers* that are discarded.

The above training screening approach relies on an estimate of a full-rank sample covariance matrix $\hat{\boldsymbol{R}}_b$. To circumvent this, we note that $\boldsymbol{x}_n^T \hat{\boldsymbol{R}}_b^{-1} \boldsymbol{x}_n = \|\tilde{\boldsymbol{x}}_n\|^2$, where $\tilde{\boldsymbol{x}} \triangleq \hat{\boldsymbol{R}}_b^{-1/2} \boldsymbol{x}_n$, i.e., the "whitened" version of $\boldsymbol{x}_n$. The whitening operation can be equivalently implemented in a time-series fashion by an MA whitening filter without the need to estimate $\hat{\boldsymbol{R}}_b$. This alternative screening approach is proposed in [32] and referred to as the *innovation power sorting* (IPS) method, since the output of the MA whitening filter is often called the *innovation* of the input (e.g., [33]).

The IPS can be extended and cast within the NS-AR framework. Specifically, we first use the ML estimator in Section IV-C to estimate the NS-AR parameters $\{\hat{a}_l(m), \hat{\sigma}_l^2\}$ from the original training set $\mathcal{S}$. Next, we form a shift-varying MA whitening filter from these parameter estimates and, similarly to (13), whiten the training set as follows:

$$\breve{x}_n(l) = \frac{1}{\hat{\sigma}_l} \left[x_n(l) + \sum_{m=1}^{M} \hat{a}_{l-L_s}(m) x_n(l-m) \right], \quad l = L_s - 1, \ldots, L-1; \quad n = 1, \ldots, N. \tag{37}$$

Finally, we compute the following metric

$$T_n = \sum_{l=L_s-1}^{L-1} \breve{x}_n^2(l), \quad n = 1, \ldots, N \tag{38}$$

which is used to replace (36) for the partition of $\mathcal{S}$ into $\mathcal{S}_1$ and $\mathcal{S}_2$.

V. Experimental Results

In this section, we present experimental results to illustrate the performance of our proposed techniques. For comparison, we consider the covariance matrix-based *ACE* test (7), the AR-model-based *NPAMF* detector (see Section IV-A-1)), our NS-AR-model-based *NS-NPAMF* detector (15), and a modified version called *NS-LP-NPAMF* that is briefly explained below. We do not compare with the AMF (5) or Kelly (6) tests which were found to perform similarly to the ACE test in our experiments. Meanwhile, the ACE, NPAMF, NS-NPAMF, and LS-NS-NPAMF are all normalized tests whose test statistics range between 0 and 1, which makes comparison more convenient.

The modification made in the NS-LP-NPAMF detector is due to an observation that HSI spectral data exhibit small oscillations. As an example, Fig. 2 depicts the original HSI data of a randomly chosen pixel from the HSI data set described below. Such oscillations along the spectral dimension do not contribute much to detection, meanwhile making parameter estimates more noisy. It was found that passing the HSI data through a lowpass (LP) filter to first remove those oscillations before applying the proposed NS-AR modeling, estimation, and detection techniques is helpful. Our NS-NPAMF detector

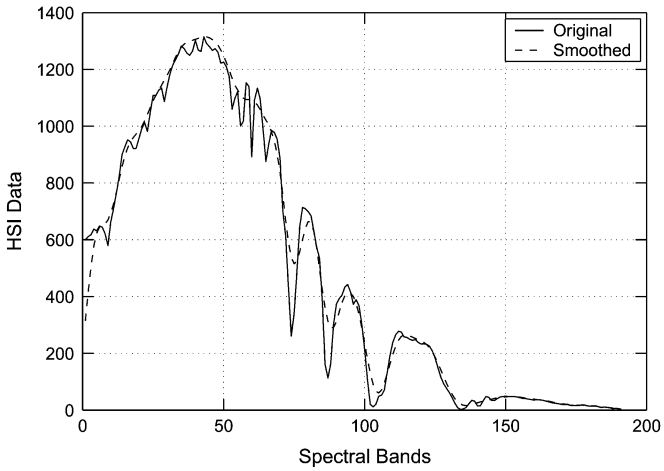


Fig. 2. Original (solid line) and smoothed (dotted line) HSI data.

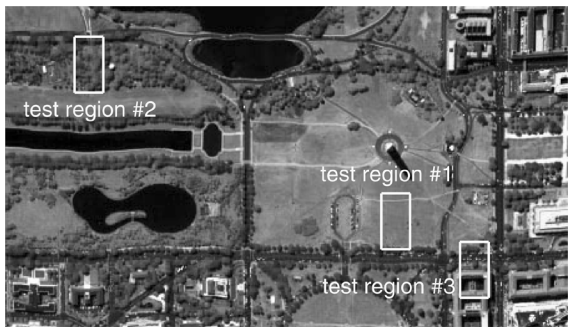


Fig. 3. HSI image of the Washington DC Mall with $L = 191$ spectral bands. Three test regions are highlighted in yellow.

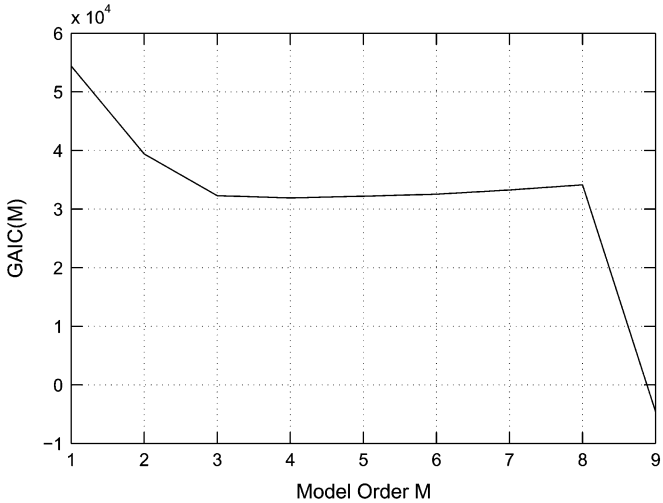


Fig. 4. NS-AR model order selection.

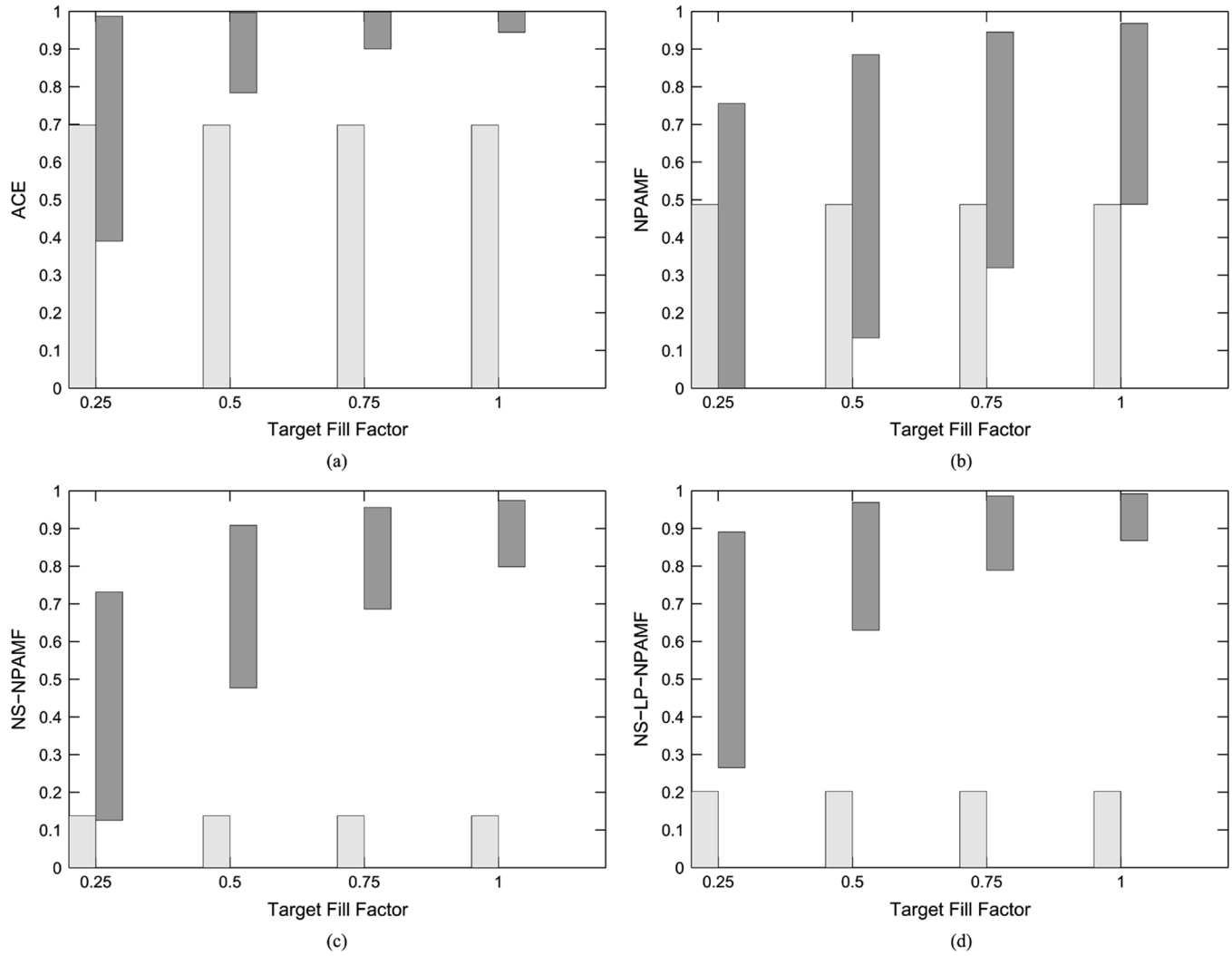


Fig. 5. Test region 1: target-background separation versus target fill factor, where the dark bars correspond to the range of test statistics under H_1, while the light bars show the counterpart under H_0. (a) ACE. (b) NPAMF. (c) NS-NPAMF. (d) NS-LP-NPAMF.

(15) with such a modification is called NS-LP-NPAMF. For lowpass filtering, we use a simple moving-average filter with impulse response given by a Kaiser window, whose length is equal to the sliding window size L_s and the shape parameter is 3. It should be noted that LP filtering is applied to all signals involved in detection, including the training pixels, test pixel, and target signature.

The HSI data employed in our studies is provided on the CD that accompanies [1]. Fig. 3 is a color infrared (IR) image from a portion of the data set, which shows a view of an airborne hyperspectral data flightline over the Washington DC area. The sensor system used in this case measured the spectral response in 210 spectral bands in the 0.4 to 2.4 μm region of the visible and IR spectrum. Bands in the 0.9- and 1.4-μm region where the atmosphere is opaque have been omitted from the data, leaving $L = 191$ spectral bands. Additional information on the data set can be found in [1]. The image shown in Fig. 3 was made using bands 60, 27, and 17 for the red, green, and blue colors, respectively. Three test regions are highlighted in Fig. 3. Test region 1 is relatively homogeneous and formed by grass, test region 2 is less homogeneous with tree and road, and test region 3 corresponds to a heterogeneous environment. To simulate the H_1 condition, we superimpose a target signal to the test pixel. The target signal corresponds to the spectral signature of a man-made object (taken from a pixel in Fig. 3) and is scaled according to particular target fill factors [2]. Each test data set is first demeaned using an approach as described in [34].

A. Model Order Selection

We first use the GAIC developed in Section IV-D to determine the model order M of the NS-AR model. Fig. 4 depicts $W(M)$ in (32) as a function of M for $L_s = 10$ and $N = 8$, and the result is obtained by averaging over the pixels in test region 1. Results obtained with the other two test regions are similar. It is seen that $W(M)$ decreases quickly as M increases from 1 to 3, reaches its minimum and remains relatively flat between 3 and 5, then increases slightly from 5 to 8, and finally drops drastically for $M = 9$. The pattern of decrease followed by increase of $W(M)$ is standard for most model selection techniques [24]. To understand why $W(M)$ drops again at $M = 9$, we note that (28) is violated with $M = 9$. As a result, $\boldsymbol{X}_l$ in (29) does not have full column rank, and there are numerous solutions for the

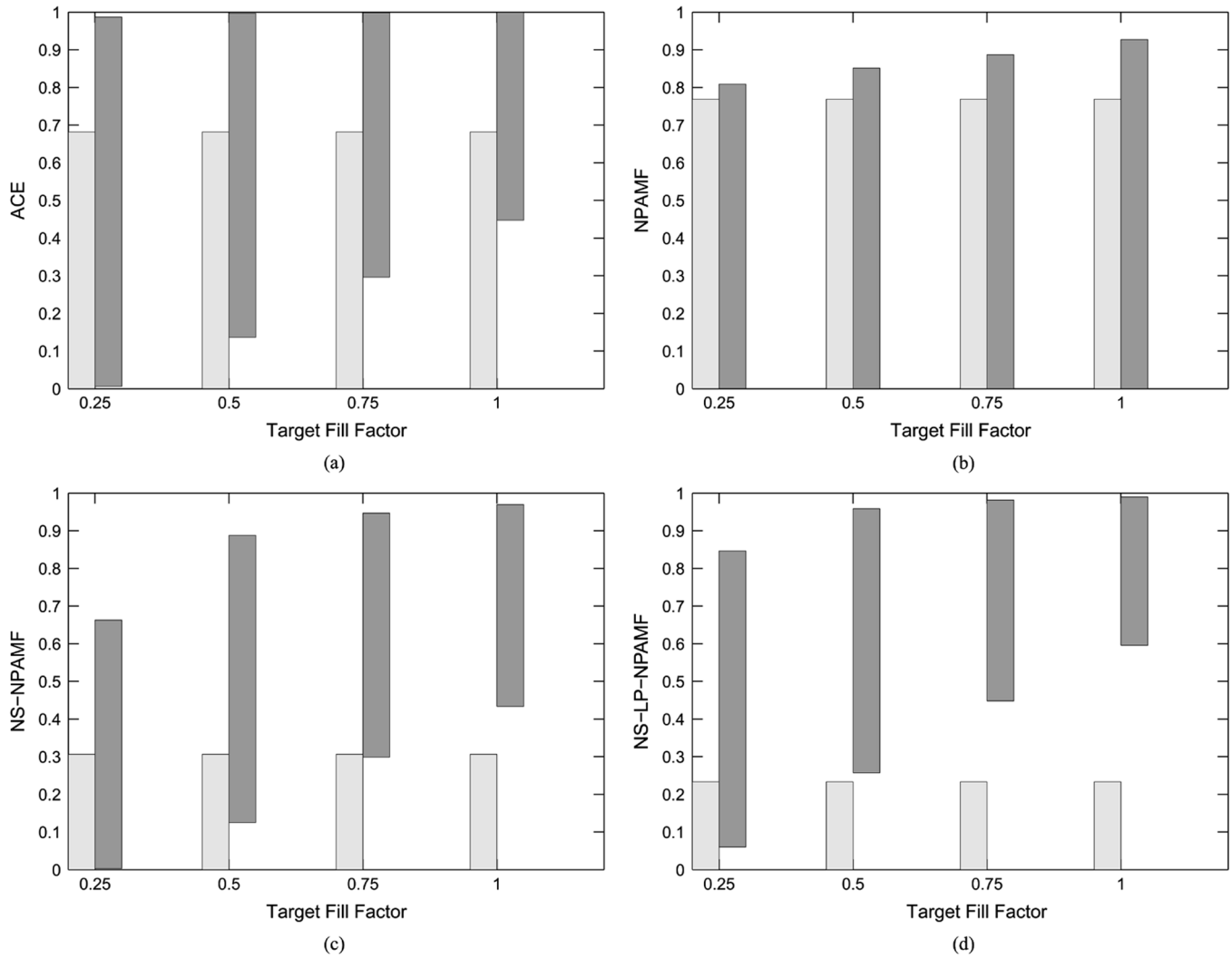


Fig. 6. Test region 2: target-background separation versus target fill factor, where the dark bars correspond to the range of test statistics under H_1, while the light bars show the counterpart under H_0. (a) ACE. (b) NPAMF. (c) NS-NPAMF. (d) NS-LP-NPAMF.

NS-AR coefficients $\{a_l(m)\}$ that lead to zero residual in the NS-AR model. In the following, we choose $M = 5$.

B. Detection in Homogeneous Environments

To illustrate detection performance in homogeneous environments, the figure of merit employed here is the *separation* of test statistics under H_0 and H_1, which is also used in [2]. For all methods, we use $N = 8$ training pixels, which corresponds to a 3×3 region without counting the center pixel (i.e., test pixel), for sample covariance matrix or parameter estimation. The sample covariance matrix $\hat{\boldsymbol{R}}_b$ is rank deficient in this case. As suggested in [2], we use the approximation $\hat{\boldsymbol{R}}_b^{-1} \approx \boldsymbol{I} - \boldsymbol{U}_1\boldsymbol{U}_1^T$, where $\boldsymbol{U}_1$ is formed by the principle eigenvectors of $\hat{\boldsymbol{R}}_b$, for the ACE detector. The subvector length (i.e., sliding window length) is $L_s = 10$ for NS-NPAMF and NS-LP-NPAMF.

First, consider test region 1. Fig. 5(a)–(d) depicts the test statistic separation of the four detectors, respectively, as a function of the target fill factor. We note that NPAMF is the worst of all detectors, which corroborates our earlier observation that stationary AR modeling is not suitable for HSI data. However, both NS-NPAMF and NS-LP-NPAMF outperform the ACE test, with NS-LP-NPAMF being slightly better than NS-NPAMF. Specifically, we see that the former achieves full target-background separation when the fill factor is 0.25, while the latter does not.

Fig. 6(a)–(d) depicts the counterpart results when the detectors are applied to test region 2, which is less homogeneous than test region 1. It is seen that all four detectors experience some degradation relative to the previous results. However, the proposed NS-NPAMF and NS-LP-NPAMF detectors, especially the latter, still significantly outperform the others.

C. Detection in Heterogeneous Environments

We now consider detection in heterogeneous environments. To this end, we embed 5 targets at randomly chosen locations in test region 3. We run the ACE and NS-LP-NPAMF detectors throughout the test region pixel by pixel, with and without training screening. If training screening is not applied, we use the $N = 8$ pixels surrounding the test pixel for training. Otherwise, we first compute metric (36) for the ACE detector and, respectively, metric (38) for the NS-LP-NPAMF detector using all pixels within the test region, and then the metrics are used

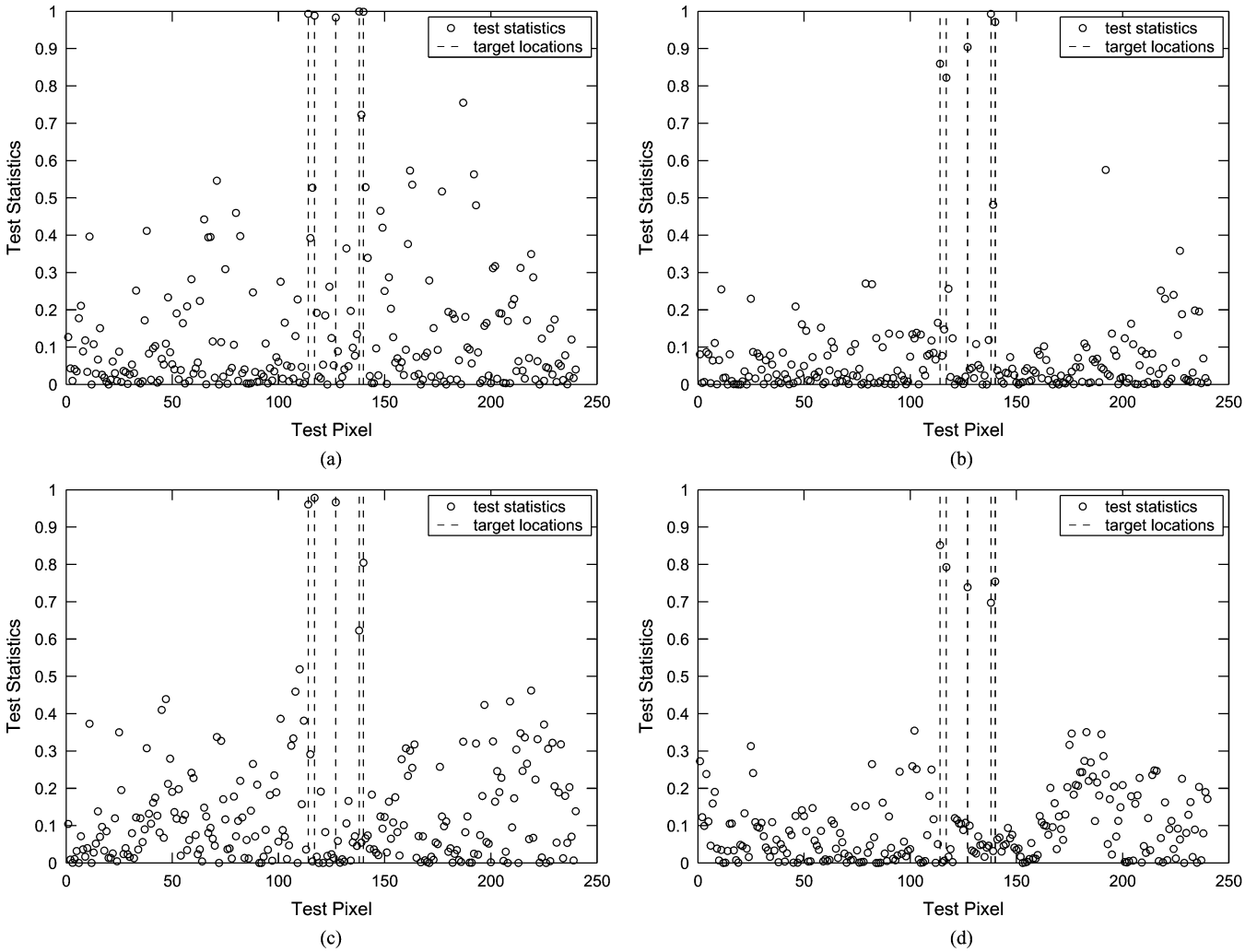


Fig. 7. Test statistics of ACE and NS-LP-NPAMF of test pixels in the test region 3 with five embedded targets. (a) ACE without training screening. (b) NS-LP-NPAMF without training screening. (c) ACE with training screening. (d) NS-LP-NPAMF with training screening.

to select $N = 8$ new training pixels to refine the parameter/covariance matrix estimate. Fig. 7(a)–(d) depicts the test statistics of the two detectors, with and without training screening, versus the index of the pixels within the test region. The dotted lines in these plots indicate the indices/locations of the embedded targets. By comparing the results, it is seen that training screening helps both detectors. It is also seen that the proposed NS-LP-NPAMF detector outperforms the ACE detector with or without training screening.

Finally, we consider a dense-target scenario by embedding not only five targets but also outliers in test region 1. In particular, about 20% of the pixels at random locations in the region are embedded with outliers that have a different spectral signature from that of the target. Fig. 8(a)–(d) shows the test statistics of the ACE and NS-LP-NPAMF detectors with and without training screening. It is seen that the NS-LP-NPAMF detector overall achieves a better performance than the other.

VI. Conclusion

In this paper, we have exploited parametric modeling of HSI data and investigated its application for subpixel target detection in HSI systems. We have shown that HSI data are nonstationary in the spectral dimension, which makes parametric adaptive modeling and detection more challenging than earlier studies for stationary data. To deal with nonstationarity, we have proposed a sliding-window-based NS-AR model tailored for HSI data. We have developed parametric adaptive detectors by exploiting the NS-AR model and addressed a range of issues, including model order selection, training screening, parameter estimation, time-series-based signal whitening, and detection. We have examined the performance of the proposed detectors and compared with covariance-matrix-based techniques using real HSI data. It has been shown that the proposed parametric detectors are more efficient in training data usage and outperform the covariance-matrix-based methods when training is limited.

Our approach implicitly assumes that HSI data is spectrally correlated. In most cases, HSI sensors oversample the spectral signal [35], which brings in spectral correlation in HSI data. The covariance-matrix-based detectors, however, can be applied in the absence of spectral correlation (as in earlier multispectral systems with a few spectral bands [17]). While the covariance-matrix-based AMF, ACE and GLRT detectors have a CFAR behavior, it is unclear whether the proposed detectors retain

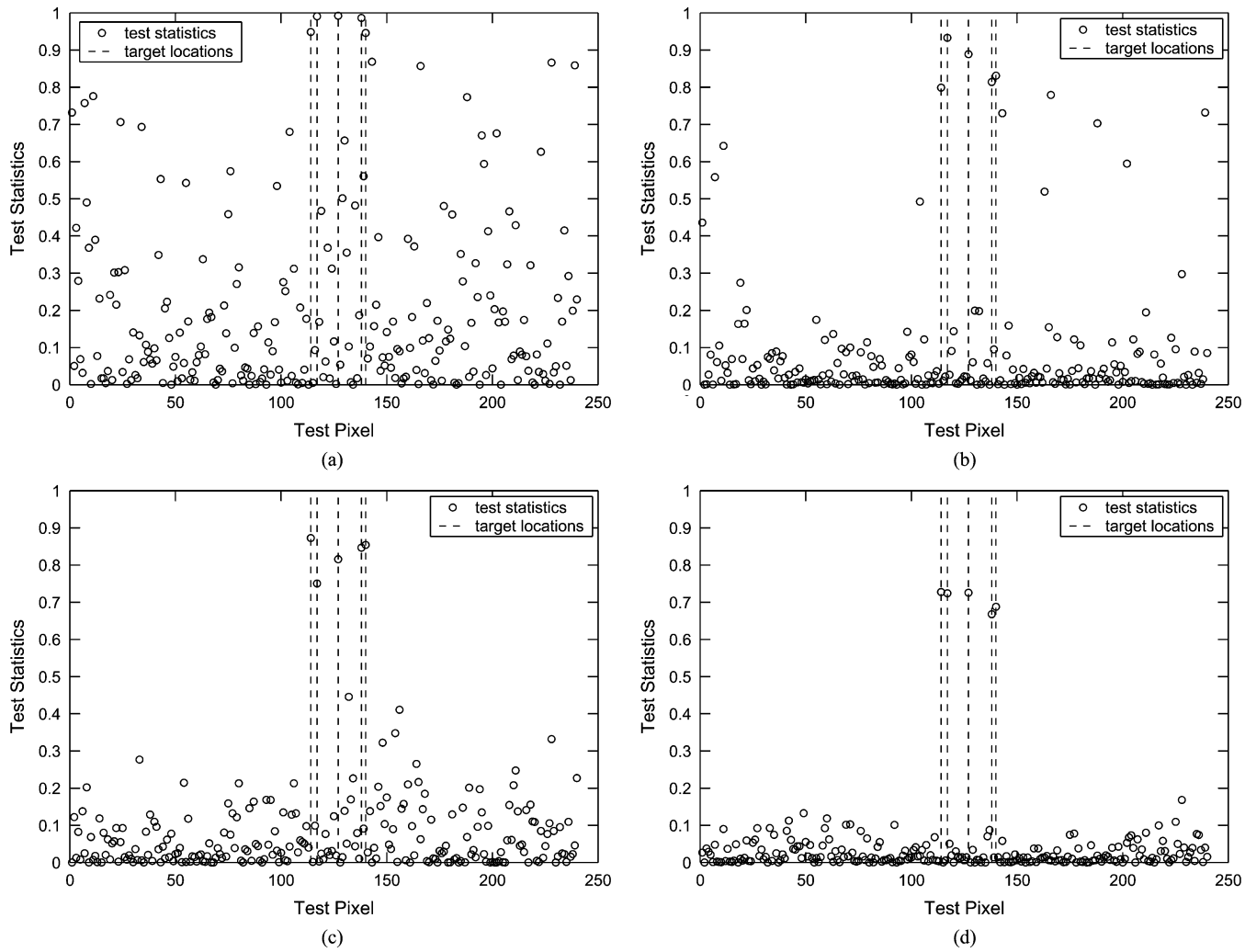


Fig. 8. Test statistics of ACE and NS-LP-NPAMF of test pixels in the test region 1 with five embedded targets and more than 20% of the pixels are embedded with outliers. (a) ACE without training screening. (b) NS-LP-NPAMF without training screening. (c) ACE with training screening. (d) NS-LP-NPAMF with training screening.

the same property. This remains an issue to be examined in the future. Other research along the proposed direction includes analytical study of the proposed detectors and exploration of alternative nonstationary parametric models for HSI target detection.

REFERENCES

[1] D. A. Landgrebe, *Signal Theory Methods in Multispectral Imaging*. Hoboken, NJ: Wiley, 2003.

[2] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," in *IEEE Signal Process. Mag.*, vol. 19, Jan. 2002, pp. 29–43.

[3] D. Manolakis, C. Siracusa, and G. Shaw, "Hyperspectral subpixel target detection using the linear mixing model," *IEEE Trans. Geosci. and Remote Sens.*, vol. 39, no. 7, pp. 1392–1409, Jul. 2001.

[4] T. K. Sarkar, S. Park, J. Koh, and R. A. Schneible, "A deterministic least squares approach to adaptive antennas," in *Dig. Signal Process.*, 1996, pp. 185–194.

[5] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Upper Saddle River, NJ: Prentice-Hall, 1998.

[6] E. J. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 22, pp. 115–127, Mar. 1986.

[7] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 1, pp. 208–216, Jan. 1992.

[8] L. T. McWhorter and L. L. Scharf, "Adaptive matched subspace detectors and adaptive coherence estimators," in *Proc. 30th Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, Nov. 1996, pp. 1114–1117.

[9] S. Kraut and L. L. Scharf, "The CFAR adaptive subspace detector is a scale-invariant GLRT," in *IEEE Trans. Signal Process.*, vol. 47, Sep. 1999, pp. 2538–2541.

[10] J. Ward, "Space–time adaptive processing for airborne radar," Lincoln Laboratory, MIT, Cambridge, Tech. Rep. 1015, Dec. 1994.

[11] M. Rangaswamy and J. H. Michels, "A parametric multichannel detection algorithm for correlated non-Gaussian random processes," in *Proc. 1997 IEEE Nat. Radar Conf.*, Syracuse, NY, May 1997, pp. 349–354.

[12] J. H. Michels *et al.*, "Multichannel parametric adaptive matched filter receiver," U.S. Patent 6 226 321, May 2001.

[13] ——, "Normalized parametric adaptive matched filter receiver," U.S. Patent 6 771 723, Aug. 2004.

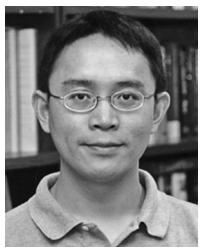
[14] J. R. Román, M. Rangaswamy, D. W. Davis, Q. Zhang, B. Himed, and J. H. Michels, "Parametric adaptive matched filter for airborne radar applications," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 2, pp. 677–692, Apr. 2000.

[15] A. L. Swindlehurst and P. Stoica, "Maximum likelihood methods in radar array signal processing," *Proc. IEEE*, vol. 86, pp. 421–441, Feb. 1998.

[16] J. Li, G. Liu, N. Jiang, and P. Stoica, "Moving target feature extraction for airborne high-range resolution phased-array radar," *IEEE Trans. Signal Process.*, vol. 49, no. 2, pp. 277–289, Feb. 2001.

[17] 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. 19, pp. 1760–1770, Oct. 1990.

[18] D. Manolakis and D. Marden, "Non gaussian models for hyperspectral algorithm design and assessment," in *Proc. 2002 IEEE Int. Geoscience Remote Sensing Symp. (IGARSS'02)*, vol. 3, Jun. 24–28, 2002, pp. 1664–1666.
[19] D. Marden and D. Manolakis, "Modeling hyperspectral imaging data," in *Proc. SPIE—Algorithms Technologies for Multispectral, Hyperspectral Ultraspectral Imagery*, vol. 5093, 2003, pp. 253–262.
[20] D. Marden and D. Manolakis, "Using elliptically contoured distributions to model hyperspectral imaging data and generate satistically similar synthetic data," in *Proc. SPIE—Algorithms Technologies for Multispectral, Hyperspectral Ultraspectral Imagery X*, vol. 5425, Aug. 2003, pp. 558–572.
[21] A. D. Stocker and A. P. Schaum, "Application of stochastic mixing models to hyperspectral detection problems," in *Proc. SPIE—Algorithms for Multispectral Hyperspectral Imagery III*, vol. 3071, Aug. 1997, pp. 47–60.
[22] G. Healey and D. Slater, "Models and methods for automated material identification in hyperspectral imagery acquired under unknown illumination and atmospheric conditions," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 6, pp. 2706–2717, Nov. 1999.
[23] B. Thai and G. Healey, "Invariant subpixel material detection in hyperspectral imagery," *IEEE Trans. Geosci. and Remote Sens.*, vol. 40, no. 3, pp. 599–608, Mar. 2002.
[24] S. M. Kay, *Modern Spectral Estimation: Theory and Application*. Englewood Cliffs, NJ: Prentice-Hall, 1988.
[25] J. H. Michels, B. Himed, and M. Rangaswamy, "Performance of STAP tests in Gaussian and compound-Gaussian clutter," *Dig. Signal Process.*, vol. 10, no. 4, pp. 309–324, Oct. 2000.
[26] J. H. Michels, M. Rangaswamy, and B. Himed, "Performance of parametric and covariance based STAP tests in compound-Gaussian clutter," *Dig. Signal Process.*, vol. 12, no. 2, 3, pp. 307–328, Apr., July 2002.
[27] G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*. San Francisco, CA: Holden-Day, 1970.
[28] T. Söderström and P. Stoica, *System Identification*. London, U.K.: Prentice-Hall Int., 1989.
[29] H. Akaike, "A new look at statistical model identification," *IEEE Trans. Autom. Control*, vol. AC-19, pp. 716–723, 1974.
[30] P. Chen, W. Melvin, and M. Wicks, "Screening among multivariate normal data," *J. Multivariate Anal.*, vol. 69, pp. 10–29, 1999.
[31] W. Melvin, "Space–time adaptive radar performance in heterogeneous clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 2, pp. 621–633, Apr. 2000.
[32] J. H. Michels, B. Himed, and M. Rangaswamy, "Robust STAP detection in a dense signal airborne radar environment," *Signal Process. (Special Issue on New Trends and Findings in Antenna Array Processing for Radar)*, vol. 84, pp. 1625–1636, 2004.
[33] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 1996.
[34] J. Y. Chen and I. S. Reed, "A detection algorithm for optical targets in clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 23, no. 1, pp. 46–59, Jan. 1987.
[35] G. Shaw and D. Manolakis, "Signal processing for hyperspectral image exploitation," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 12–16, Jan. 2002.



Hongbin Li (M'99) received the B.S. and M.S. degrees from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 1991 and 1994, respectively, and the Ph.D. degree from the University of Florida, Gainesville, in 1999, all in electrical engineering.

From July 1996 to May 1999, he was a Research Assistant in the Department of Electrical and Computer Engineering at the University of Florida. He was a Summer Visiting Faculty Member at the Air Force Research Laboratory, Rome, NY, during summers 2003 and 2004. Since July 1999, he has been an Assistant Professor in the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ. His current research interests include wireless communications, statistical signal processing, and radars.

Dr. Li is a member of Tau Beta Pi and Phi Kappa Phi. He received the Harvey N. Davis Teaching Award in 2003 and the Jess H. Davis Memorial Award for excellence in research in 2001 from Stevens Institute of Technology, and the Sigma Xi Graduate Research Award from the University of Florida in 1999. He is an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS.

James H. Michels (F'05) received two M.S. degrees from Syracuse University, Syracuse, NY, in both physics and electrical engineering, and the Ph.D. degree in electrical engineering in 1991.

His Ph.D. dissertation in multichannel signal processing provided the foundation for in-house and contractual research programs in his former position at the Air Force Research Laboratory (AFRL). These included the development of multichannel model-based methods for radar space–time adaptive processing (STAP). He has over 50 papers and six patents in these technical areas. Two of these patents have made contributions to biomedical applications involving the detection and extraction of physiological features and conditions via multichannel electrocardiogram (ECG) processing. From August 1998 to August 1999, he was a visiting Scientist at the former Defence Evaluation and Research Agency (DERA), Malvern, U.K., sponsored by the Air Force Office of Scientific Research (AFOSR) Engineer and Scientist Exchange Program (ESEP). Currently, he is a Research Faculty Member at Syracuse University, where he also assists a graduate student research group. His current research interests include detection, estimation, multichannel adaptive signal processing, time series analyses, space–time adaptive processing (STAP), change detection, particle filtering, and stochastic resonance.

Dr. Michael twice received the Charles Ryan Award for contributions to basic research and received the Chief Scientist Award for Best Published Paper in 1995 at AFRL. In 1999, he was an IEEE Region I Award winner for "contributions to multi-channel signal processing."