

*Review*

## **Overview of Physical Models and Statistical Approaches for Weak Gaseous Plume Detection using Passive Infrared Hyperspectral Imagery**

**Tom Burr \* and Nicolas Hengartner**

Mail Stop F600, Los Alamos National Laboratory, Los Alamos NM 87545, USA

E-mails: [tburr@lanl.gov](mailto:tburr@lanl.gov) and [nickh@lanl.gov](mailto:nickh@lanl.gov)

\* Author to whom correspondence should be addressed; E-mail: [tburr@lanl.gov](mailto:tburr@lanl.gov)

*Received: 22 October 2006 / Accepted: 4 December 2006 / Published: 6 December 2006*

---

**Abstract:** The performance of weak gaseous plume-detection methods in hyperspectral long-wave infrared imagery depends on scene-specific conditions such as the ability to properly estimate atmospheric transmission, the accuracy of estimated chemical signatures, and background clutter. This paper reviews commonly-applied physical models in the context of weak plume identification and quantification, identifies inherent error sources as well as those introduced by making simplifying assumptions, and indicates research areas.

**Keywords:** clutter, generalized least squares, infrared, model averaging, temperature-emissivity separation, errors in predictors, plume detection

---

### **1. Introduction**

Remote detection and identification of chemical plumes using airborne passive infrared (IR) sensors is a challenging problem. Our goal is to detect weak gaseous chemical plumes in cluttered scenes that contain a mixture of background types, such as water, vegetation, asphalt, concrete, buildings, etc. Data from a typical scene consists of the measured, calibrated radiance  $r$  for each of many wavelengths (covering some subset of the visible to the long-wave IR regions using approximately 128 to 256 spectral channels spanning wavelengths from approximately 0.5 to 15  $\mu\text{m}$ ) in each of approximately 500  $\times$  500 spatial pixels (in both the cross-tracking direction and the with-tracking direction of the airplane carrying the IR sensor). The measured signal at each pixel depends on the atmosphere and ground radiance, atmospheric transmission, instrument noise, and whether a

chemical plume lies between the ground and the detector. The relative contributions of these various factors changes with wavelength, so that in principle, the composition of the background and the plume can be estimated.

We seek to determine whether small (up to a few or a few tens of pixels out of tens or hundreds of thousands of pixels) and weak (both in terms of temperature contrast and chemical strength, see section 3) plumes from a known library of possible chemicals might be present in the scene. These chemicals have effects on the measured at-sensor radiance that we refer to as “chemical signatures,” which are based on “known” spectra (the spectral library values are not known perfectly) that must be transformed (introducing an error source). Therefore, the “chemical signatures” are also not known perfectly.

Most real scenes look to the human eye like a mixture of components (Scene A, in Figure 1, [1,2]) and this background mixture is called “clutter.” However, whether formal mixture models consisting of mixing multiple components such as water, asphalt, vegetation, etc. (which each have multivariate distributions) are effective for analysis depends on the goals. A single-component multivariate Gaussian (SCMG) refers to a Gaussian distribution having a single mean vector and covariance matrix. Although modeling clutter as a SCMG may seem too simplistic, it has proven to be surprisingly effective for weak chemical plume detection (see section 7) when compared to other approaches that have been proposed for IR data [3-10].

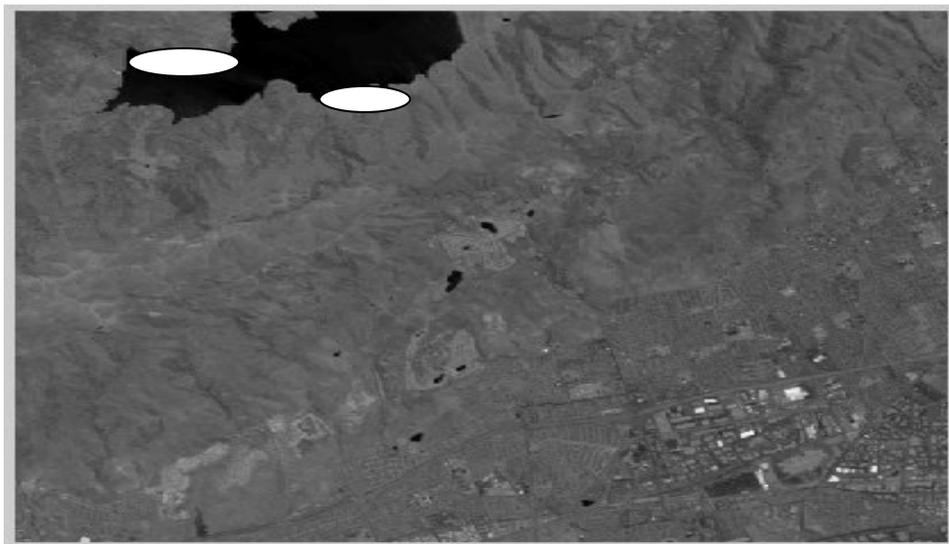
This paper emphasizes issues that impact performance in detecting weak plumes and quantifying the chemicals in a weak plume. Performance is typically defined as the probability of detecting a plume for a given, fixed, and small false alarm probability. Section 2 gives additional background. Section 3 describes the physical model for calibrated, measured IR radiance at aircraft elevation including atmospheric, ground, plume, and sensor effects. Sections 4 and 5 describe model simplifications involving background (nonplume) clutter and linearization of nonlinear plume effects (including the well-known temperature-emissivity separation challenge). This description clarifies inherent error sources and error sources introduced by model simplifications. Section 6 is a problem and solution taxonomy. Section 7 describes current inferences algorithms (linear and nonlinear), associated statistical issues, and factors impacting performance and performance indicators. Section 8 includes related topics and provides additional detail including research areas for some of the topics discussed in previous sections. Section 8 lists key definitions. Section 9 summarizes.

## 2. Background

Figure 2 is a simple schematic of the basic sensor and scene setting. Typically, each pixel represents a ground area covering  $\sim 10$  to  $\sim 100$   $m^2$ . Plume-like pixels are those thought to have a gas plume influencing the signal; background pixels are those thought not to have a plume influencing the signal. Section 3 describes the terms used in the physical models that are indicated in Figure 2.

Detecting gaseous plumes in cluttered scenes using hyperspectral images has received growing attention, dating to approximately 1995 when hyperspectral (having relatively many spectral channels) imaging systems were deployed and analysts began to develop detection and quantification methods [1,2]. This review focuses on identifying weak, rare, but approximately known targets (chemical signatures) in cluttered backgrounds of varying radiance values using one (or one time-averaged)

image consisting of  $p$  wavelengths (spectral channels) in the IR region. Applications include detection of chemical and biological weapons as they are being developed and/or being used, and environmental monitoring. Qualitatively, the main task is to look for evidence of one or more chemical signatures from a “known” library in a scene that is cluttered due to variation in ground materials, ground temperatures and shading, atmospheric effects, viewing angle, and sensor noise. An associated task is to assess our confidence in finding and quantifying such signatures. Confidence assessments often rely on model assumptions and/or computationally-intensive empirical studies.

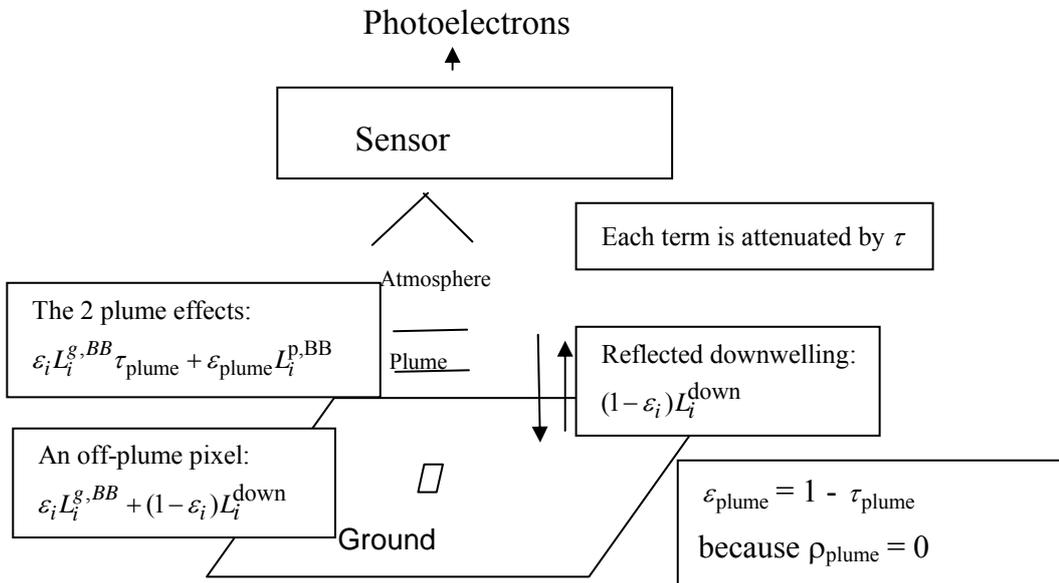


**Figure 1. Scene A.** Typical broadband image of a scene that is cluttered because it includes mountains, buildings, and water, but contains no plumes. Similar images are available from the Jet Propulsion Lab’s AVIRIS website [1]. The two white ovals are hypothetical results of a plume-detection method, indicating either false alarms or the presence of gaseous plumes along the lake shore.

IR hyperspectral imagery can be used to characterize plumes. By “characterize,” we mean either to detect (locate pixels that contain a plume), to detect and identify (identify the chemical components), or to detect, identify, and quantify the plume components (and possibly plume temperature), depending on the context. In addition, background scenes can be segmented using various clustering methods; only when such clustering is part of the plume-characterization [11] will it be within our scope (section 8). More broadly, other targets (including solid or liquid targets) can be discovered using IR, but our scope is gaseous plume characterization [3-10].

IR is challenging, because nearly everything (solids, liquids, and gases) emits and/or absorbs in the IR region of the spectrum. The visible, near IR, short-wave IR, mid-wave IR, and long-wave IR regions each have unique challenges and opportunities, but we consider broad issues that apply generally for the entire near IR and IR regions. We consider only one (or one time-averaged) IR image. Other approaches become available if the ground pixels are viewed more than once or if specific regions of the spectrum are used. For example, it is believed that mid-wave IR data might allow the opportunity to observe plume effects twice, allowing for improved inference and confidence assessments. Using mid-wave IR, the first observed plume effect is explained as in section 3 (Figure 2)

when the plume lies in the direct line of sight between the sensor and the corresponding ground pixels. The second observed plume effect is a shadow effect whereby certain pixels when viewed from directly above will be in the shadow from the sun of a small off-the-ground plume that is some distance removed from the direct line of sight to the sensor.



**Figure 2.** Simple description of on-plume and off-plume pixels. Off-plume pixels contribute  $\epsilon_i L_i^{g,BB} + (1 - \epsilon_i) L_i^{down}$ , where  $L_i^{g,BB}$  is the term describing Planck’s radiation law for an ideal black body,  $\epsilon_i$  is the ground emissivity, which is less than 1 for real surfaces, and  $(1 - \epsilon_i) L_i^{down}$  is the reflected downwelling atmospheric radiation. A plume has two effects: it absorbs some of the emitted ground radiation, and it emits radiation. The emitted ground radiation that transmits through the plume is described using  $\epsilon_i L_i^{g,BB} \tau_{plume} = \epsilon_i L_i^{g,BB} (1 - \epsilon_{plume})$  and the plume emission is  $\epsilon_{plume} L_i^{p,BB}$ . All terms are attenuated by transmission through the atmosphere,  $\tau$ . This paper focuses on small (impacting up to tens of pixels among tens of thousands of pixels) and weak gaseous plumes. Weak plumes in this

context are also called optically thin plumes, meaning that  $\tau_{plume} = e^{-\sum_{k=1}^{n_c} \sigma_k C_k} \approx 1 - \sum_{k=1}^{n_c} \sigma_k C_k$ , where  $C_k$  is the concentration of chemical  $k$  integrated over the pathlength through the plume and  $\sigma_k$  is the chemical spectra for chemical  $k$ . This results in approximately a linear relation between plume absorption (or emission) and either of plume thickness, chemical concentration, or the temperature contrast between the ground and plume.

Figure 1 is the broadband image (the average radiance over all  $p = 224$  spectral channels) of an example scene (614 vertical by 512 horizontal pixels) that is cluttered because it contains mountains, buildings, and water, but contains no plumes. This scene is in the visible/near IR region [1,2], and for our purposes can be considered to have the same features as do IR images. IR images are typically displayed for the human eye using gray scale to represent scalar-valued radiance magnitude at each pixel. The scalar value is typically the broadband image or an average of a subset of the spectral channels.

### 3. Physical Models

Plume detection requires detailed modeling of the physical processes contributing to the measured signal, as we describe next. Figure 2 indicates the main physical processes, as we discuss next.

#### 3.1 Kirchoff's Thermal Equilibrium Law

For an object that is in thermal equilibrium (not heating or cooling), Kirchoff's law expresses conservation of energy [12] as  $\alpha + \tau + \rho = 1$ , where the fractions of incident energy that are absorbed, transmitted, and reflected, are  $\alpha$ ,  $\tau$ , and  $\rho$ , respectively. Kirchoff's law also states that emissivity must equal absorbance for surfaces in thermal equilibrium (good absorbers are good emitters). Therefore, assuming the plume is in thermal equilibrium, and that  $\rho_{\text{plume}} = 0$ , Kirchoff's law for the plume can be written as  $\varepsilon + \tau = 1$ , where emissivity  $\varepsilon$  ( $0 \leq \varepsilon \leq 1$ ) characterizes how closely a real object obeys Planck's radiation law describing an ideal blackbody. For example, if the observed pixel is a patch of mowed grass, then the emitted radiation from the grass is  $\varepsilon_{\text{grass}} L^{BB}$ , where

$$L^{BB} = \frac{c_1 V^3}{\exp\left(\frac{c_2 V}{T}\right) - 1}$$

is Planck's radiation equation and  $\varepsilon_{\text{grass}}$  values other than 1 allow for departure of

mowed grass from an ideal black body obeying Planck's emission law. Similarly,  $\varepsilon_{\text{plume}} L_i^{\text{p, BB}}$  is the radiation emitted by the plume. Because the amounts of emitted, transmitted, and reflected radiation are a function of angle, some treatments include integration over angles. We are not aware of any implemented plume-characterization algorithms that exploit the integral equations, so do not include them here. Also, relatively subtle emissivity differences have been observed from pixels that are thought to represent the same material, such as grass. Material height, moisture, shading, etc., can all contribute potentially important IR emissivity differences among background components. One important aspect of emissivity differences among relatively homogeneous pixels is their impact on methods [13,14] to separate temperature and emissivity (sections 4 and 5).

Concerning notation, we will use  $\alpha$ ,  $\varepsilon$ , and  $\tau$  to denote absorption, emission (emissivity), and transmission, respectively. Reflected energy will have a special symbol (such as  $L_i^{\text{down}}$  in the next section) to denote the energy source. Scalars, vectors, and matrices will be distinguishable by context.

#### 3.2 IR signals

The photons detected by an IR hyperspectral detector associated with background pixel  $i$  can be modeled as

$$S_i^b = (\varepsilon_i L_i^{g, BB} + (1 - \varepsilon_i) L_i^{\text{down}}) \tau + L_i^{\text{up}} + N_i, \quad (1)$$

where  $\varepsilon_i$  is the emissivity,  $L_i^{g, BB}$  is the Planck blackbody function at ground temperature,  $\tau$  is the atmospheric transmission,  $L_i^{\text{up}}$  is upwelling radiance,  $(1 - \varepsilon_i) L_i^{\text{down}}$  is downward atmospheric radiance reflected off the ground, and  $N_i$  includes all omitted effects plus instrument noise [8,15-17]. We typically write, for example,  $\varepsilon_i$  rather than  $\varepsilon_i(\nu_j)$  but caution that nearly all terms depend on frequency

(or equivalently, on wavelength), and that all terms in Eq. (1) are vectors. Various approaches differ partly on the basis of which terms are ignored in Eq. (1) and which terms are assumed to be approximately constant with frequency and/or constant across pixels or a subset of relatively homogeneous pixels. For example, some treatments neglect  $(1-\varepsilon_i)L_i^{\text{down}}$ , the downward atmospheric radiance reflected off the ground because it is relatively small in certain wavelength regions (such as in the long wave IR range) [12,17].

A plume has two effects: it absorbs some of the emitted ground radiation, and it emits radiation. The emitted ground radiation that transmits through the plume is described using  $\varepsilon_i L_i^{g, BB} \tau_{\text{plume}} = \varepsilon_i L_i^{g, BB} (1 - \varepsilon_{\text{plume}})$  and the plume emission is  $\varepsilon_{\text{plume}} L_i^{p, BB}$ . All terms are attenuated by transmission through the atmosphere,  $\tau$ . It is then simple to show that the signal from plume pixel  $i$  is  $S_i^b$  plus terms to model the plume effect, which can be expressed as

$$S_i^p = \alpha_p [L_i^{p, BB} - (\varepsilon_i L_i^{g, BB} + (1 - \varepsilon_i) L_i^{\text{down}})] \tau + S_i^b, \quad (2)$$

where  $\alpha_p = 1 - \tau_p$  is the plume absorption and  $L_i^{p, BB}$  is the Planck function at plume temperature. Eq. (2) neglects scattering effects by the plume and assumes that downward atmospheric radiance reflected off the ground is the same for on-plume and off-plume pixels. That is,  $(1 - \varepsilon_i) L_i^{\text{down}}$  is the downward atmospheric radiance reflected off the ground for both on-plume and off-plume pixels. This implies that downward radiance from the plume that is reflected into the sensor line of sight is neglected. We also assume the plume is close to the ground, so the distance through the atmosphere from ground to sensor is essentially the same as the distance from plume to sensor; therefore,  $\tau$  denotes the atmospheric transmission either from plume to sensor or from ground to sensor. Finally, because we are primarily concerned with the measured difference between on-plume and off-plume pixels,  $L_i^{\text{up}}$  can be ignored.

The  $\varepsilon_i$  terms (emissivities) depend on the properties of the background. Concrete, asphalt, buildings, grass, dirt, water, and other common background features each have their characteristic emissivity. Real and synthetic background scenes of interest are generated from various mixtures of many background emissivities and relatively few plume pixels.

Using the approximation  $\alpha_p = 1 - \tau_p = 1 - e^{-\sum_{k=1}^{n_c} \sigma_k C_k} \approx \sum_{k=1}^{n_c} \sigma_k C_k = \sigma \beta$ , (this is Beer's law [17] for chemical absorption, and assuming a weak, optically thin plume, so that  $1 - e^{-x} \cong x$ ), where  $C_k$  is the concentration of chemical  $k$  integrated over the pathlength through the plume,  $\sigma = (\sigma_1 | \sigma_2 | \dots | \sigma_{n_c})$  is the  $p \times n_c$  matrix of chemical spectra with  $\sigma_k$  the chemical spectra (cross section) for chemical  $k$ ,  $\beta = (C_1, C_2, \dots, C_{n_c})$  is the unknown gas concentration vector,  $n_c$  is the number of chemicals in the plume, and defining  $A_i$  as

$$A_i = \text{Diag} \{ [L_i^{p, BB} - (\varepsilon_i L_i^{g, BB} + (1 - \varepsilon_i) L_i^{\text{down}})] \tau \} \sigma, \quad (3)$$

where  $\text{Diag} \{ \}$  indicates a diagonal matrix, we can rewrite Eq. (2) as

$$S_i^p = A_i \beta + S_i^b. \quad (4)$$

We can write equation (4) generically as

$$r = A_i \beta + z, \quad (5)$$

where  $r$  is the mean-centered, measured, calibrated radiation at pixel  $i$ ,  $\beta$  is the amount of the chemical “signature”  $A_i$  at pixel  $i$  which we want to estimate, and  $z$  is the mean-centered background clutter ( $z = S_i^b - \text{avg}(S_i^b)$ ) [16]. We will refer to  $A_i$  (Eq. (3)) as the “chemical signature” or “target” in the task of distinguishing background from background plus target.

If any of the estimated components in the  $\beta$  parameter is large, this is evidence of a plume at pixel  $i$  (or perhaps over groups of several contiguous pixels at a time, as discussed in sections 7 and 8). There are two main challenges here. First,  $S_i^b$  varies considerably among pixels; this is the “clutter” challenge that we describe in section 4. Second, the target signature  $A_i$  also varies among pixels; this challenge is sometimes included in “temperature-emissivity” separation efforts (sections 5 and 8), and we refer to it as the “misspecified target signature issue.”

In Eq. (5), we typically assume that  $r$  has been mean-centered to have zero sample mean in the scene. In practice, it is not known whether a real scene contains any plumes, so there is typically an iterative procedure that first assumes there are no plumes, then looks for plumes using the covariance estimate from the “off-plume” pixels, then removes “plume-like” pixels and re-estimates the “off-plume” covariance matrix and mean vector. Because we focus on characterizing the background in cases having weak and rare plumes, we will not consider such an iterative procedure. However, plume-detection performance is expected to degrade when on-plume pixels are used to estimate the background covariance matrix and/or mean vector. Performance also degrades if the chemical target directions are similar to the clutter directions (as defined, for example, by the eigenvector directions in the spectral decomposition of the background covariance matrix). It is clear that background clutter (a mixture of materials, viewing angles, viewing illumination, etc.) and target misspecification (“errors in predictors”) combine to make this a challenging inference task, as described in sections 4 and 5, respectively.

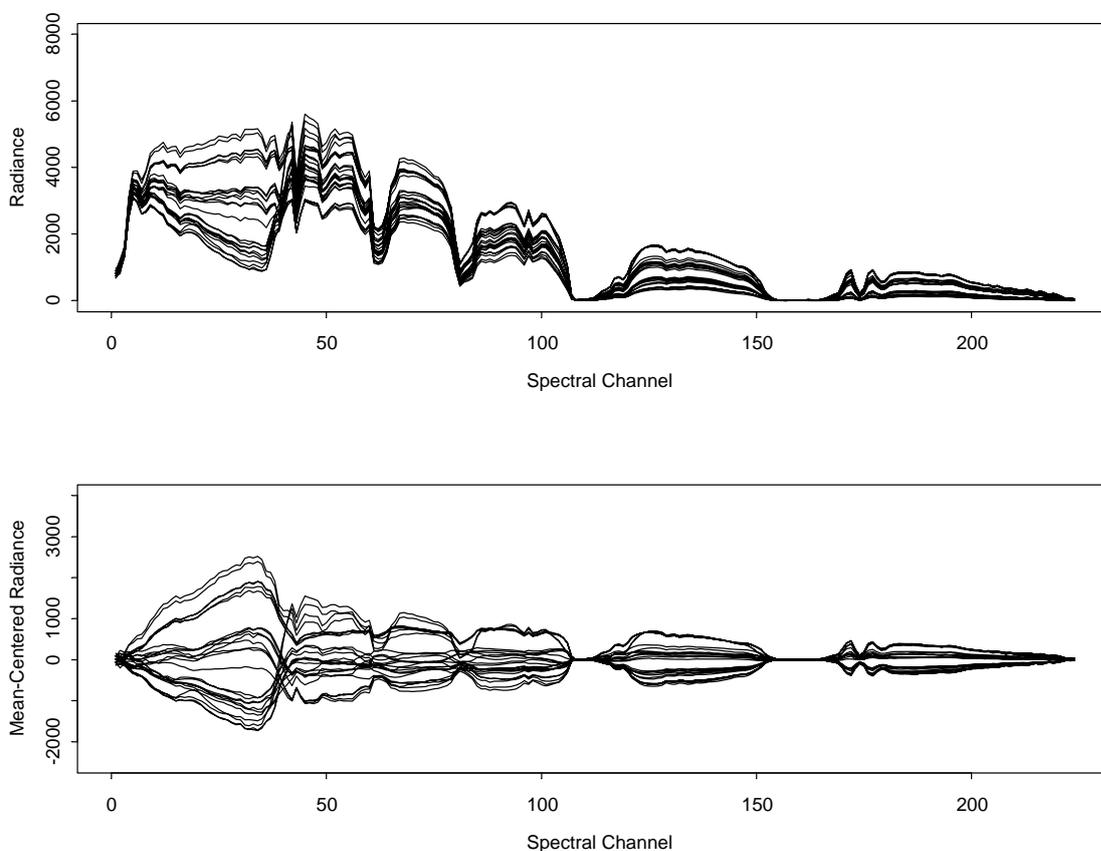
#### 4. Background Clutter $S_i^b$

Several studies have investigated the distribution of background clutter [7,15,18-25] in IR scenes. There is complete agreement that the distribution of  $S_i^b$  is not accurately described as a SCMG. However, which model is most appropriate depends on the goals. As discussed in section 7, for the weak plume detection goal, there also is reasonably good agreement that algorithms based on the SCMG model for clutter are surprisingly effective. For example, references [3-10] conclude either that the simplistic SCMG approach is the overall best method or is not much worse than other methods.

There are several implications of concluding that an SCMG model is effective in our context. First, it is a simple model-based summary of complicated clutter. Second, decision thresholds could be based on SCMG data, and therefore could be computed analytically (analogous, for example, to claiming that a decision threshold of  $\pm 2$  standard deviations corresponds to a 5% false alarm probability by appealing to a Gaussian approximation). Third, it would suggest that it might be difficult to find robust

methods that improve the GLS plume-detection performance (defined as the false negative rate for a given small fixed positive rate). However, if the SCMG model is inadequate, a scene-specific empirical reference distribution could be used to select decision thresholds and associated confidence measures (see Section 8.6).

Figure 3 plots the  $p = 224$  channel response from each of several selected pixels from scene A. The mean-centered radiance (bottom plot) for the selected pixels exhibits long runs of either positive (above the mean) or negative (below the mean) values. This is a type of mixture distribution that might be expected for  $S_i^b$  on the basis of looking at cluttered scenes such as Scene A, which have contiguous patches of lower-than-average or higher-than-average emissivities and is a clear departure from the one-component multivariate Gaussian assumption.

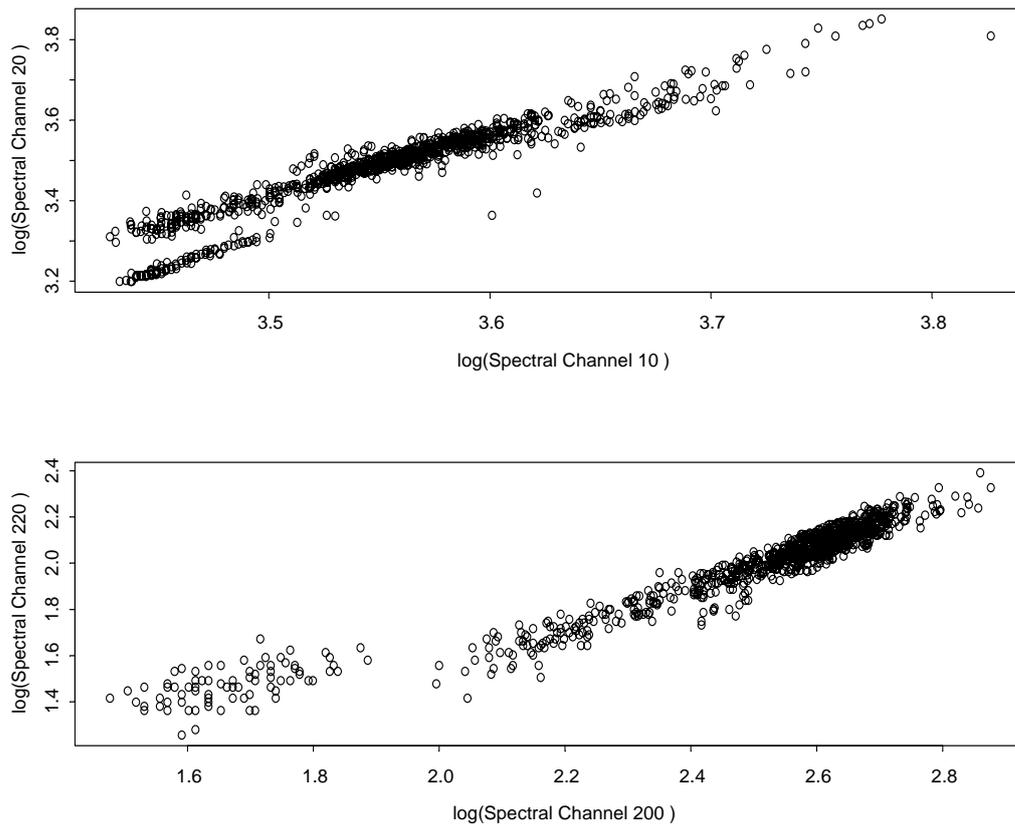


**Figure 3.** Example radiance (top) and mean-centered radiance (bottom) from selected pixels in scene A.

The scatter plots in Figure 4 are two sets of log-transformed (base 10) values from two pairs of spectral channels from each of 1000 randomly-chosen pixels from Scene A. Both plots appear to have at least two or three distinct clusters (probably corresponding to distinct ground components in the scene), which is another indication of non-Gaussian behavior. If the distribution across pixels were more nearly bivariate Gaussian, these scatter plots would be elliptical, without clusters.

Figure 5 (top plot) is a histogram of 40000 randomly-generated values of the squared Mahalanobis distance defined as  $MD^2 = r\hat{\Sigma}^{-1}r^T$  [21]. The  $MD^2$  values in the top plot are distributed as a  $\chi^2(p)$  (chi-

squared with  $p$  degrees of freedom) random variable because the radiance  $r$  values were simulated from the  $N(0, \hat{\Sigma})$  distribution with  $\hat{\Sigma}$  estimated from Scene A. The top plot is substantially different from the bottom plot, which is 40000 randomly-selected  $MD^2$  values from scene A. This also indicates that the  $r$  values from scene A are not well modeled in this context by the  $N(0, \hat{\Sigma})$  distribution. The corresponding  $\chi^2(p)$  probabilities are given by the smooth density function in both the top and bottom plots. It is evident that the  $\chi^2(p)$  probabilities fit the top plot but not the bottom plot.

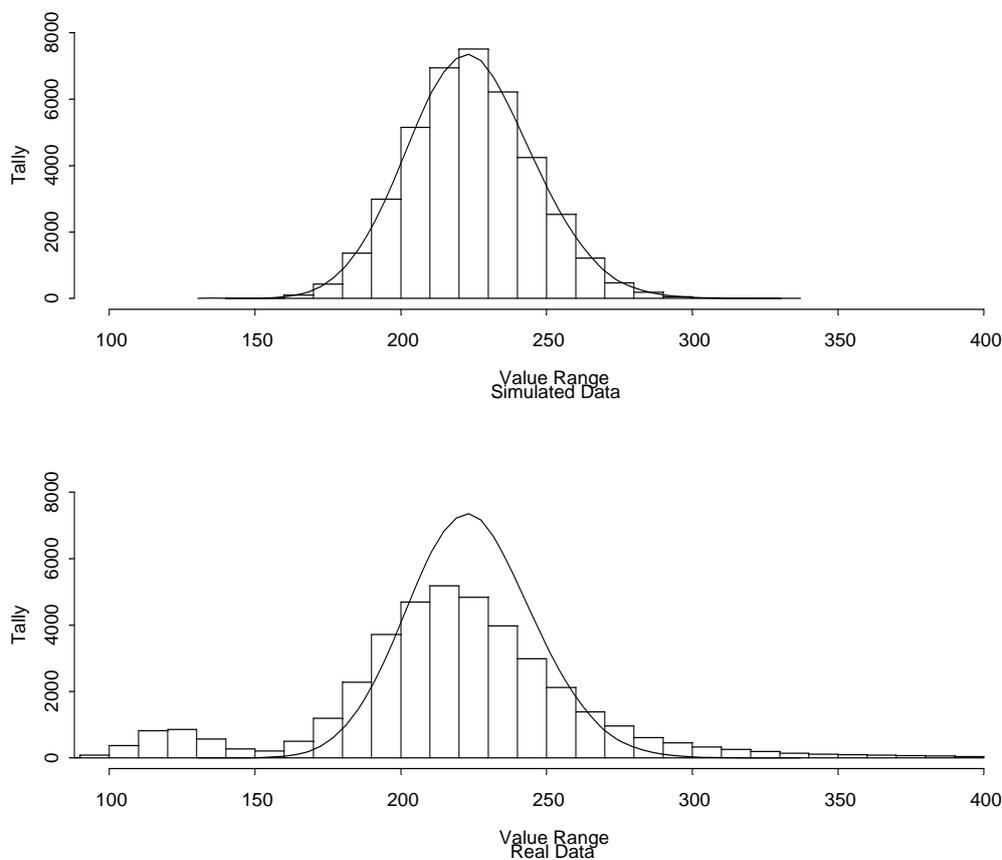


**Figure 4.** Scatter plot of two sets of log-transformed (log base 10) spectral channel pairs, each from the same 1000 randomly-selected pixels from scene A.

Several studies have observed that  $MD^2$  values from IR scenes are not well approximated by a  $\chi^2(p)$  distribution. For example, [21] used plots similar to Figure 5, and also used four quantitative distances for comparing two distributions to demonstrate departure from  $\chi^2(p)$  behavior. Of the four distances, the distance that is most appropriate in our context was the “exceedance metric,” which summed the absolute values of quantiles of the two distributions. The  $p$ th quantile  $F^{-1}(p) = x_p$  of a distribution  $F(x)$  is the value  $x_p$  such that  $100p\%$  of the values generated from  $F$  are less than  $x_p$ , and the exceedance metric between two distributions  $F$  and  $G$  (defined as  $D = \int |F^{-1}(u) - G^{-1}(u)| du$ ) was evaluated using  $K$  equally spaced probabilities (on a logarithm scale). The choice of  $K$  and the spacing of the  $p_i$  is arbitrary, but provided these are chosen before observing the data, this is a reasonable measure of distance between any two distributions. Section 8 describes why another appropriate

distance measure in our context is to compare one or a few specific quantiles or to compare percentages that exceed specified thresholds.

The observed tendency for  $MD^2$  values from IR scenes to be poorly modeled by a  $\chi^2(p)$  distribution has led to consideration of elliptically-contoured (EC) distributions. These EC distributions generalize the multivariate Gaussian in a way that allows for longer and fatter tails; however, as seen in the bottom plot of Figure 5 (and in the “stair-step” plots in [21]), the  $MD^2$  values from real scenes also exhibit clustering, which is not predicted by any EC distribution. This has led to various mixture strategies such as fitting the main body of the  $MD^2$  values with one EC distribution (often close to Gaussian) and the tail of the  $MD^2$  values with a second EC distribution (often not close to Gaussian).



**Figure 5.** The squared Mahalanobis distances from 40000 (top) simulated  $r$  values; (bottom) from randomly-chosen  $r$  values from pixels from Scene A. The smooth curve in both plots is based on the  $\chi^2(p)$  distribution.

For example, [21] applied this strategy to relatively homogeneous subsets of pixels (which were located using a clustering algorithm applied to the  $r$  values). Somewhat surprisingly, even these relatively homogeneous subsets of pixels appear to arise from a mixture distribution. This was also observed in [26] which proposed a combination of local mean subtraction and nonparametric density estimation to estimate the distribution of  $r$  values (Eq. (5)) for AHI (advanced hyperspectral imager) data [27].

Nonparametric density estimation is often effective, but usually only in small dimensions; in large dimensions such as our  $p = 224$  case, the well-known curse of dimensionality [28] suggests that nonparametric density estimates will be low quality. The approach in [26] was in the context of building a model for the background data so that a potential target could be identified on the basis of being anomalous compared to the background model, without having a library of targets. It is not yet known whether the approach in [26] is competitive when a target library is available. In addition, we point out that the simplified model that assumes  $r \sim N(0, \hat{\Sigma})$  together with an estimate of  $\hat{\Sigma}$  is a parametric density estimation approach because it assumes a parametric form for the density (multivariate Gaussian), and the density estimation task reduces to estimation of  $\Sigma$ . This mitigates the well-known curse of dimensionality, but is effective only if the multivariate normality assumption leads to acceptable target detection schemes and associated confidence measures. Section 7 includes more detail involving the number of observations (pixels) required to estimate  $\Sigma$  effectively. For comparison, note that [28] reports that to estimate the density at the mean value of a 10-parameter normal distribution requires 842,000 samples to achieve an average squared error of approximately 0.1.

One motive for fitting any distribution to real data is that it provides insight, as well as a model-based summary. Also, if the model approximates the relevant aspects of  $r$  sufficiently well, it could provide both a way to optimize plume detection and to quantify the confidence in detected plumes. To date, there has been no fully successful demonstration of sufficiently accurate fitting of  $r$  values across a scene using any model ([18, 21, 22] and section 7). In addition, fitted EC distributions focus exclusively on the scalar-valued  $MD^2$  arising from  $r$  values, which only partly captures the relevant aspects of  $r$  values in the context of plume detection. Work in this area is ongoing, and a recent reference [29] developed a method to detect outliers in EC-distributed data. It is currently unknown whether a weak chemical plume would be a sufficiently large outlier to be detected by this outlier-detection approach.

In summary, it is known that one would not choose a SCMG to describe  $S_i^b$  on the basis of the clutter (mixture distribution) that our eyes see in real scenes. However, the appropriate model depends on the goals, and we indicate in section 7 that modeling  $S_i^b$  as a SCMG remains competitive for our plume characterization goals. There has been considerable interest in fitting EC distributions in our context, although EC distributions also do not fit the  $r$  data particularly well. However, individual EC distributions or mixtures of EC distributions fit the distributions of  $MD^2$  values much better than does a  $\chi^2(p)$  distribution, which arises from assuming the  $r$  values arise from a SCMG model. Finally, it is not yet clear whether it will be useful to fit a mixture of EC distributions for the purpose of plume detection [8,15,29].

##### 5. The Target Signature $A_i = \text{Diag}\{[L_i^{p, BB} - (\varepsilon_i L_i^{g, BB} + (1 - \varepsilon_i) L_i^{\text{down}})]\tau\}\sigma$

The terms  $L_i^{p, BB}$  and  $L_i^{g, BB}$  in the target signature  $A_i$  involve Planck's blackbody function,  $L^{BB} = \frac{c_1 v^3}{\exp(\frac{c_2 v}{T}) - 1}$ . The frequency  $v$  corresponding to each wavelength channel can be known to high

accuracy, although instrument resolution and drift in sensor calibration imply that the wavelengths

and corresponding frequencies are not known exactly (see section 8.11). The temperature  $T$  is unknown, and can vary across pixels. The emissivity  $\varepsilon_i$  is a vector of  $p$  unknowns, although  $\varepsilon_i$  can generally be assumed to vary slowly with frequency (see Figure 3). Recall that the term  $L_i^{\text{down}}$  is sometimes neglected, depending on the wavelength region ( $L_i^{\text{down}}$  is small in the LWIR), but generally it contributes to the target signature in varying amounts, thus adding uncertainty to  $A_i$ . The atmospheric transmission  $\tau$  must be estimated (see section 8.5), and the spectral library  $\sigma$  is also estimated (see section 8.11) Because the number of unknowns per pixel exceeds the number of observations (wavelengths) per pixel, various approximations are used for the target signature  $A_i$ , as we now describe.

One common approximation is based on the assumption that the plume temperature is nearly equal to the ground temperature,  $T_p \approx T_g$ , so that, for example,  $L_i^{p, BB}$  can be accurately approximated using only the linear term in a Taylor series expansion [15-17]. Then,  $L_i^{p, BB} \approx L_i^{g, BB} (1 + \frac{c_2 \nu \Delta T}{T_g^2})$ , where  $\Delta T = T_p - T_g$  is the “temperature contrast” between the plume and ground, which can be positive or negative. It is important to recognize that the target signature can also be positive or negative, and sometimes can switch from positive to negative in the same scene even if the multiple plumes consist of the same chemical. Whether the temperature contrast can switch signs (positive to negative or negative to positive) among pixels of the same plume is another unknown that is scene-specific. This implies that inference options that impose constraints on parameter signs via prior information cannot always assume the target signature is nonnegative. Assuming the linear approximation to Planck’s function is adequate, the target signature can be approximated as

$$A_i \approx \text{Diag} \{ [L_i^{g, BB} (1 - \varepsilon_i + \frac{c_2 \nu \Delta T}{T_g^2}) - (1 - \varepsilon_i) L_i^{\text{down}}] \tau \} \sigma, \quad (6)$$

which depends on both pixel and frequency. If the target signature depends on both pixel and frequency in unknown ways due to the dependence on the emissivity  $\varepsilon_i$  and on  $\frac{\Delta T}{T_g^2}$ , then performance claims regarding target detection are difficult.

As a rough approximation, one approach is to assume  $\varepsilon_i = 1$  for all pixels [17], so the approximation for  $A_i$  becomes  $A_i \approx \text{Diag} \{ [L_i^{g, BB} \frac{c_2 \nu \Delta T}{T_g^2}] \tau \} \sigma$ , implying that  $A_i$  is linear in  $\Delta T$ , and the target signature depends on  $\Delta T$  and on  $\text{Diag} \{ \tau \} \sigma$ .

Depending on the particular wavelength, most objects’  $\varepsilon_i$  values vary in a non-negligible way. For example, in the long-wave IR, most  $\varepsilon_i$  values are in the range 0.90 to 1.0 [17]. Therefore, more correctly,  $\varepsilon_i$  varies among pixels, as does the temperature, so there are often attempts to separately estimate temperature and emissivity. Define  $B_1 = (1 - \varepsilon_i)(L_i^{g, BB} - L_i^{\text{down}})$ , and  $B_2 = L_i^{g, BB} \frac{c_2 \nu}{T_g^2}$ . Then,

$$S_i^p = S_i^b + \text{Diag} \{ (B_1 + B_2 \Delta T) \tau \} \sigma, \quad (7)$$

$\text{Diag}\{B_1\tau\} \sigma$  and  $\text{Diag}\{B_2\Delta T\} \sigma$  are the “signatures” that we seek evidence for, which are not dependent on  $T_p$ , but which can be nearly collinear, leading to poor estimation performance.

Another compromise is to let  $\varepsilon_i$  (but not temperature) vary among pixels, and then seek evidence for the “signature”  $A_i$  as approximated in Eq. (6). Because there are more unknowns than observations, various assumptions must be made regarding  $\varepsilon_i$ , such as spatial and/or spectral smoothness. Spatial smoothness assumes that nearby pixels have similar  $\varepsilon_i$  values. Spectral smoothness assumes that  $\varepsilon_i$  varies slowly with wavelength. Various “fit background” methods (section 7) explicitly or implicitly make similar assumptions.

Although it is not necessary to separately estimate emissivity and temperature to estimate  $A_i$ , this section illustrates that errors in  $A_i$  are pixel-dependent because of varying temperature and background emissivity. A referee has pointed out that an effective way to reduce the errors in an estimate of  $A_i$  that does not require separate estimates of temperature and emissivity is as follows. First estimate  $A_i$  and the corresponding  $\beta$  via Eq. (4) using  $A_i = \text{Diag}\{[L_i^{p, BB} - L_i^{\text{clutter}}]\tau\} \sigma$ , where  $L_i^{\text{clutter}} = \varepsilon_i L_i^{g, BB} + (1 - \varepsilon_i) L_i^{\text{down}}$  can initially be estimated using any rough estimate of the background clutter (see section 6.2). This leads to a refined estimate of the background clutter  $L_i^{\text{clutter}}$ , which can then be used to improve the initial estimate of  $A_i$ .

Obtaining high-quality estimates of  $\beta$  using any of these methods is difficult; accuracy depends strongly on the chemicals in the plume and on  $\Delta T$ . Each of the above approaches has merit and has been applied to real and simulated data, particularly in the context of temperature-emissivity separation (TES) which is often treated separately from plume characterization (section 7). However, a tentative finding is that more elaborate methods that account for the error in  $A_i$  are difficult to implement (too many pixels for routine fitting) and do not substantially outperform methods that ignore errors in predictors [6-8,16]. Strategies to deal with these predictor errors (misspecification of the target signature) is the subject of ongoing work, and TES-related issues such as those presented here help to quantify the anticipated errors in  $A_i$ .

To summarize, Eq. (5) is  $r = A_i\beta + z$ , where  $r$  is the mean-centered measured radiation at pixel  $i$ ,  $\beta$  is the amount of target  $A_i$  ( $A_i$  is defined in Eq. (3), but various approximations are typically used, such as the approximation in Eq. (6)) present in the pixel, and  $z$  is the mean-centered background effect ( $z = S_i^b - \text{avg}(S_i^b)$ ) [16]. Regardless of the approach, note that the chemical signature depends on the temperature contrast  $\Delta T$ . This is a characteristic of passive (not active) sensors [17].

## 6. A Taxonomy of Problems and Solution Approaches

In Eq. (5), if the mean-centered radiance  $r$  is measured in  $p$  channels at a given pixel, there will be a minimum of  $p + 1$  unknowns to estimate ( $p$  emissivities and temperature) even in the absence of a chemical target and assuming  $L_i^{\text{down}}$ ,  $\tau$ , and  $L_i^{\text{up}}$  can be estimated (or neglected) using other data or information, leaving only the  $\varepsilon_i L_i^{g, BB}$  term in Eq. (1) that describes  $r$  to be estimated. This is an underdetermined problem even without trying to find evidence for a chemical signature from a transformed chemical library. Therefore, various assumptions regarding how the emissivity  $\varepsilon_i$  and temperature vary across pixels and how the emissivity varies with frequency are made.

### 6.1 A Taxonomy of Problems

This review is limited to weak gaseous plume detection in IR imagery. Within this scope, interpretation of Eq. (5) for the purpose of plume detection indicates that the following questions related to clutter and misspecification of the chemical target(s) determine a range of possible analysis challenges, approaches, and limitations for any given IR image. The need for TES impacts both the clutter and the “misspecification of target” issues.

#### 6.1.1 Questions regarding the scene that relate to how background clutter $S_i^b$ might be modeled

Can the pixels be segmented effectively into relatively homogeneous regions [11,21,26,30], and a separate search for chemical plumes applied to each region? If so, the background clutter model might be impacted.

What causes pixel radiation across the entire image to vary? An obvious effect such as varying background is typically the main explanation, but it is not the only explanation.

What causes pixel radiation within a relatively homogeneous region to vary and why is an EC distribution better than MVN even within homogeneous regions? [30]

Is there a library of background emissivities  $\epsilon$  available that is known to be relevant for the scene? If not, perhaps consider using  $\hat{\Sigma}$  estimated from the entire scene or from pixel subsets to define possible endmembers (section 7) to be used in simultaneously fitting the background and possible chemical targets.

#### 6.1.2 Questions regarding the “misspecification of target”

Is it possible to identify a homogeneous subset of pixels, such as those over identified land features (water, smooth man-made material, etc.) that can improve a TES algorithm? Note that controlled test releases often have objects intended for calibration in the scene, and various options related to TES become available. In this case, can any of the TES methods be effective on a per-pixel basis?

How will the atmospheric transmission  $\tau$  be estimated ( $\tau$  is part of the target signature)? The typical choices [31] are: ignore  $\tau$  (shown to be a bad choice in most plume-characterization studies), estimate  $\tau$  using in-scene methods [32] that rely on knowing something about a subset of the pixels, such as that they represent nearly constant (with respect to frequency) emissivity pixels (water pixels are a good choice here); or estimate  $\tau$  using atmospheric-transmission models such as FASTCODE [33] or MODTRAN [34].

Is a library of possibly chemical spectra available and at what temperature(s)?

Is this a release of a known chemical or a wide-open search for any chemical from a large chemical library? What is known regarding the upper limit for the number of chemicals in a plume?

Does the scene have an associated known chemical plume test release? Controlled chemical releases are occasionally performed. Such “known releases” usually cannot control weather and wind conditions so the chemical concentrations in the “known” plume have nonnegligible measurement error. If error in the “ground truth” is a concern, one strategy is to assume that chemical ratios can be known fairly accurately regardless of weather conditions.

One approach in use today is to follow Foy et al. [17] and assume that the pixel-to-pixel variation in the target signature can be ignored. Alternatively, [6] evaluated the impact of misspecifying the target, but did so in a generic manner by adding varying levels of measurement error (representing pixel-to-pixel variation plus measurement error in the chemical spectra and  $\tau$ ) without a detailed study of the various factors that contribute to the measurement error.

### 6.2 A Taxonomy of Solution Algorithms

Manolakis and Amico [15] divided algorithms for gas plume detection using IR into two broad types: unstructured background and structured background.

The unstructured background leads to a generalized least squares (GLS) approach as follows. Recall that Eq. (5) is  $r = A_i \beta + z$ , where the sample mean of  $z$  can be assumed to be 0, because the sample mean can be subtracted from  $z$  (“mean-centering”). Then if we assume the SCMG model,  $z \sim N(0, \hat{\Sigma})$ , where  $\hat{\Sigma}$  is the  $p \times p$  sample covariance estimated from all background pixels ( $\hat{\Sigma}_{jk} = \sum_{i=1}^n (r_{ij} - \bar{r}_j)(r_{ik} - \bar{r}_k)/(n-1)$ ), the GLS solution is

$$\hat{\beta} = (A^T \hat{\Sigma}^{-1} A)^{-1} A^T \hat{\Sigma}^{-1} r, \quad (8)$$

which is scalar-valued if  $A$  is a single chemical signature and vector-valued if  $A$  includes more than one signature. Here, we have dropped the  $i$  subscript on  $A$  but recognize that  $A$  will in practice be misspecified and pixel-dependent to varying degrees depending on scene-specific conditions (section 6.1.2). This GLS approach is also called the adaptive matched filter (AMF) approach, because it adapts to the data through use of  $\hat{\Sigma}$ . Also, the GLS solution remains the same (and is the minimum variance unbiased estimate of  $\beta$ ) even if we only assume that the clutter has mean 0 and covariance  $\hat{\Sigma}$ , without making the SCMG assumption. However, the SCMG assumption implies that the GLS solution is optimal (having maximum detection probability for a given false positive probability) for distinguishing background from a specified target.

The structured background approach attempts to fit the background, compute the resulting residuals, and search for evidence of the target(s) in the residuals. The background is fit using background “endmembers,”  $A_B$ , writing

$$r = A_T \beta_T + A_B \beta_B + \tilde{z}, \quad \text{where } \tilde{z} \sim N(0, \Sigma_{\text{diagonal}}). \quad (9)$$

The residual matrix  $\Sigma_{\text{diagonal}}$  is expected to have zero off-diagonal elements because of fitting to the endmembers  $A_B$ . Background endmembers can be chosen directly from the scene, chosen from a spectral library of relevant background materials thought to be in the image, or derived from  $\hat{\Sigma}$ . There are several options for choosing them from the scene, including various clustering methods. If derived from the symmetric, positive definite matrix  $\hat{\Sigma}$ , then the principal components, which are linear combinations of the original data  $r$ , with weights given by the components of the eigenvectors in the

spectral decomposition of  $\hat{\Sigma}$  are a common choice [45]. Using realistic simulated data, reference [45] showed promising results with a slightly modified method that involved nonnegativity constraints.

This “fit background” approach is also called orthogonal subspace projection (OSP) in the remote sensing literature. In the statistical literature, the concept is the same as in an “added variable plot” [35]. That is, if  $r$  is first fit to the background endmembers  $A_B$ , then the estimated effect of additional explanatory variables such as the potential target signatures  $A_T$ , can be obtained by a linear regression of the residuals in the  $r$  fit to  $A_B$  regression on the residuals in the  $A_T$  fit to  $A_B$  regression. In the context of variable selection for regression (with the goal of deciding which predictors should be included in a model), the statistical literature [35] also includes some of the results from [5].

The GLS approach has performed well compared to OSP and to all other alternatives attempted to date [3-10,16,19] although performance comparisons depend on specific goals and contexts, including spectral library sizes. Bajorski [5] compared the GLS to the OSP method by viewing the GLS method as arising from variable selection (in which all the background coefficients  $\beta_B$  were set to zero) applied to the OSP method. Not surprisingly, the range of conditions (spectral “angles”) where GLS will outperform OSP can be predicted. This assumes that the model in Eq. (9) ( $r = A_T\beta_T + A_B\beta_B + \tilde{z}$ , where  $\tilde{z} \sim N(0, \Sigma_{\text{diagonal}})$ ) can adequately describe the data (leading to negligible patterns in the residuals) and it ignores errors in  $A_T$ , and/or  $A_B$ . This is a partial candidate explanation for the empirical observation that GLS remains a worthy contender; however, it is not yet known whether errors in  $A_T$  and/or  $A_B$  should be considered.

Often, empirical observations regarding the relative performance of OSP and GLS are made on the basis of adding synthetic plumes to real scenes assuming there is no error in  $A_T$  or in  $A_B$ , and sometimes on the basis of relatively limited numbers of real test releases in which true chemical strengths are difficult to know precisely because of unknown weather conditions impacting the plume.

## 7. Inference Approaches

The inference tasks are to first locate pixels that exhibit plume-like behavior and then to select which chemicals from a chemical library are thought to be present in these plume-like pixels. Some approaches [4,5,7,8,16] combine these tasks by first estimating  $\beta$  in Eq. (5) (or some version of repeated application of Eq. (5) such as the model averaging option described below) for each pixel. This requires an estimate of  $\Sigma$ .

The estimate  $\hat{\Sigma}$  can be based on all pixels or on all pixels except the pixel(s) being evaluated via Eq. (5). For small plumes, there should be little difference in performance arising from using either all pixels or on all pixels except for the one(s) being evaluated. In practice, it is not known whether a real scene contains any plumes, so there is typically an iterative procedure that first assumes there are no plumes, then looks for plumes using  $\hat{\Sigma}$  from the “off-plume” pixels, then removes “plume-like” pixels and re-estimates the “off-plume”  $\hat{\Sigma}$ . As a practical issue, it might require too much time to re-estimate  $\hat{\Sigma}$  if each pixel is to be examined individually via Eq. (5). However, plume-detection performance is expected to degrade when on-plume pixels are used to estimate the background  $\Sigma$  and  $\bar{r}$ . Performance also degrades if the chemical target directions are similar to the clutter directions (as defined, for example, by the eigenvector directions in the spectral decomposition of the background  $\Sigma$  [5]).

Another approach [6,19] applies the two tasks in distinct steps, although iterative estimation of  $\Sigma$  is still required if plume-like pixels are to be omitted from  $\hat{\Sigma}$ . First there is a rapid screening to identify contiguous groups of atypical pixels and then there is an attempt to identify which chemicals are in these plume-like pixel regions. The rapid but computationally demanding screening could use a library of chemical signatures one at a time to identify contiguous pixels that all have evidence of one or more chemical signatures impacting the radiance  $r$ . The white ovals in Figure 1 are a hypothetical depiction of this approach. Because of background clutter and possible overlap between background radiance and chemical (target) signatures, these white ovals could be false positives. This two-step approach will fail to detect single-pixel plumes, but fitting to the pixel-group average has been shown to be effective for finding multiple-contiguous-pixel plumes [6,19].

The remainder of this section describes five inference options for the task of selecting which chemicals from a chemical library are thought to be in a pixel or in a “super-pixel” (a collection of contiguous pixels that was found by the rapid screening for groups of plume-like pixels). A super-pixel radiance is typically the average radiance of contiguous pixels. The contiguous pixels are chosen because they all exhibit departure in the same direction from the background model in Eq. (5), thus providing evidence that the  $\beta$  value for at least one chemical in the library is nonzero. If size of the chemical library is relatively small, say 20 chemicals or less, then Eq. (5) can be fit directly to a pixel or super-pixel allowing the full library to be used in one step. Alternatively, any of the techniques in the variable selection literature [35,36] could be considered (section 8). However, recall that we are focusing on the case where the chemical library is larger than the number of spectral channels, so Eq. (5) is underdetermined without imposing constraints. The constraint we discuss is to assume any plume of interest will contain no more than three chemicals.

### 7.1 GLS and maximum penalized likelihood

The GLS solution to Eq. (5) for a given pixel is  $\hat{\beta}$  as given in Eq. (8), which is an  $p$ -by- $k$  dimensional matrix consisting of  $p$  rows (1 row per spectral channel) of  $k$  concentration estimates [35]. If we assume that real plumes contain at most three chemicals, then  $k \leq 3$ . However, we typically must evaluate many or all possible subsets of 1, 2, or 3 chemicals. Because the number of spectral channels is typically at least 100 or more, this approach restricts attention to the overdetermined  $p > k$  case.

The GLS solution is also a likelihood ratio (LR) solution provided the SCMG assumption holds, so that  $z \sim N(0, \hat{\Sigma})$  in Eq. (5). According to the Neyman Pearson lemma [37], LR solutions have the desirable property of being optimal in the sense of having the smallest false negative rate for a given false positive rate. Furthermore, the  $z \sim N(0, \hat{\Sigma})$  assumption leads to a convenient analytical way to quantify performance without resorting to empirical testing.

Regarding the need to estimate  $\Sigma$ , the signal processing literature [37] refers to a constant false alarm rate detector (CFAR) as one that has a fixed and known false alarm rate. Student's- $t$  distribution was developed for the case of having to estimate the variance in order to test with a CFAR statistic whether a scalar-valued mean was equal to a hypothesized value [35]. Analogously, the need to estimate  $\hat{\Sigma}$  and the desire for a CFAR statistic led Kelly [38] to a small correction factor involving the number of pixels used in the estimate  $\hat{\Sigma}$ . This small correction factor is typically ignored in images

having many pixels. However, it is also known that the estimates of the smallest eigenvalues of  $\hat{\Sigma}$  tend to be underestimated [39], and this could impact OSP-like approaches that use the eigenvectors of  $\hat{\Sigma}$ .

As a step towards addressing this concern, we evaluated  $\hat{\Sigma}$  using varying numbers of pixels (one  $r$  vector per pixel) from Scene A, including 1000, 2000, 3000, 4000, 5000, up to 32000 pixels. We assumed that  $\hat{\Sigma}$  based on 40000 randomly selected pixels was the “true” value (having a smallest eigenvalue of approximately 1.48 and a largest eigenvalue of approximately  $4.5 \times 10^7$ ), and then simulated the varying numbers of pixels. By using simulated pixel values, we could evaluate the impact of sample size alone, without considering pixel heterogeneity. In this case, there was a clear negative bias for the smallest eigenvalue that gradually decreased from approximately 40% relative error for 1000 pixels to approximately 0.5% relative error for 32000 pixels. The mean absolute difference between all  $\hat{\Sigma}$  entries and all  $\Sigma$  entries decreased rapidly as the number of pixels increased from 1000 to 10000 and then very slowly decreased beyond 10000 pixels. Similarly, the sum of the eigenvalues (which equals the sum of the variances) varied erratically with large magnitude that decreased rapidly as the number of pixels increased from 1000 to 10000, and then continued to slowly decrease. The issue of sample size (number of pixels) has not been thoroughly studied in our context, but sample size is nearly always a factor that impacts performance. Qualitatively, we recognize that sample size is made artificially large by pooling heterogeneous pixels to estimate  $\hat{\Sigma}$ . Some [11,20,26] have worked with smaller sample sizes of more homogeneous pixels.

The GLS estimate for  $\hat{\beta}$  must be converted to a decision regarding which chemicals are present.

One option for doing so involves the use of a penalized negative log likelihood,  $PL = \frac{RSS_j}{RSS_{\min}} + \lambda S_j$ ,

where  $RSS_j$  is the residual sum of squares for subset  $j$ ,  $RSS_{\min}$  is the minimum RSS over all fitted models,  $\lambda$  is a tuning parameter, and  $S_j$  is the number of predictors (chemicals) in subset  $j$  [6]. The residual is the error remaining after fitting Eq. (5) using  $\hat{\beta}$  from Eq. (8) corresponding to subset  $j$ . Readers who are familiar with variable selection literature will recognize that the penalty-for-large-models value  $\lambda = 2$  corresponds to the well-known Aikike information criterion and  $\lambda = \ln(p)$  corresponds to the Bayesian information criterion [6, 36]. The “pick winner” method (as we will refer to it below) simply chooses that chemical subset that has the smallest PL value (smallest is best because this is a negative log likelihood). Among the contending models is the null model having no chemicals, so it is possible that the prediction is “no chemical present.” In practice,  $\lambda$  can be chosen empirically, so that a desired false alarm rate is maintained.

## 7.2 Bayesian Model Averaging

In order to decide which chemicals are present in a candidate plume, GLS can be applied to each subset of 1, 2, or 3 chemicals, and then Bayesian Model Averaging (BMA) [40] can be used to estimate the probability that each chemical from the library is present. These probability estimates are impacted by nonGaussian behavior [23, 41], implying that although the BMA outputs are in the 0 to 1 range, they do not necessarily behave as well-calibrated probabilities [41]. Here we briefly describe BMA for subset selection.

For a given data set  $D$  and probability model for the data, it would be ideal if we could calculate the exact probability of each subset. By Bayes theorem,  $P(M_1 | D) = P(D | M_1)P(M_1)$ , where  $P(M_1)$  is the prior probability for model (subset)  $M_1$ . This calculation requires calculation of the expression  $P(D | M_1) = \int P(D | M_1, \beta_1)\pi(\beta_1 | M_1)d\beta_1$ , where  $\beta_1$  is the coefficient vector for the chemicals in model  $M_1$ . Such integrals are notoriously difficult in most real problems requiring either numerical integration, analytical approximation, or Markov Chain Monte Carlo methods [41-43]. Even if the integral could be computed accurately, we would rarely know the exact subset probabilities because real data never follows any probability model exactly. Therefore, various approximations are in common usage, with the BIC (Bayesian Information Criterion) being perhaps the most common [40].

Following [40],  $P(M_j | D)$  can be approximated using the approximate result that  $P(M_j | D) \propto e^{(-BIC_j/2)}$ , where  $BIC_j = p \ln(RSS_j / p) + (n_j + 1) \ln(p)$ ,  $n_j$  is the number of predictor variables in model  $j$  (the number of chemicals from the library that are being fit in model  $j$ ),  $RSS_j = \sum_{j=1}^p (\hat{\Sigma}^{-1/2} r_j - \hat{\Sigma}^{-1/2} A_i \hat{\beta}_i)^2$  is the residual sum of squares for model  $j$ ,  $p$  is the number of spectral channels per pixel, and  $\hat{\beta}_j$  is the ordinary least squares estimate of  $\beta_j$ . The BIC expression is derived using the Laplace method for approximating the integral required to calculate  $P(M_j | D)$ , and assuming a flat prior (over the region where the integrand is nonnegligible) for the value of  $\beta_j$ . Model  $M_j$  is defined by the subset of chemicals from the full chemical library being used in a particular fit, so we use the terms “model” and “subset” interchangeably here.

The probability that chemical  $C$  is present is  $P(C|D) = \sum_{j=1}^M I(C \in M_j)P(M_j | D)$ , where  $I(\cdot)$  equals 1 if its argument is true. That is, to estimate  $P(C|D)$ , we simply sum the model probabilities for each subset that includes chemical  $C$ . Although these probabilities have varying accuracy, depending on the data, BMA is one of the most effective strategies for chemical subset selection, particularly when prior information such as some chemical combinations being highly likely or unlikely is available [40].

On the basis of one small study [6], it appears that the “pick-winner” that uses the penalized likelihood approach performs approximately the same as the BMA method, both with and without “errors in predictors,” in the absence of any prior information. However, because both methods evaluate many subsets (or in some cases, evaluate all subsets), there is almost no computational advantage in using the “pick-winner”  $PL$  method; in addition, BMA performs much better [40] if there is prior information such as restrictions on the magnitudes and/or signs of the  $\beta$  coefficients. Note that in our case, both methods use the strong prior information that no more than three chemicals are in a plume.

### 7.3 Fitting the Background Options

Recall that Eq. (5) could be replaced with Eq. (9),  $r = A_T \beta_T + A_B \beta_B + \tilde{z}$ , where the background is explicitly modeled by fitting background endmembers rather than using  $\hat{\Sigma}$  as estimated from all background pixels. The basis vectors chosen as endmembers can be obtained using principle components or, Foy and Theiler [46] illustrated promising results using independent components. Regardless of how the endmembers are chosen, this is an OSP-type approach [44,45] and Bajorski [5]

provides one of the most definitive treatments of when it should outperform GLS-type approaches based on Eq. (5). If the chemical spectra considered by Bajorski [5] had been transformed via the “whitening” transform involving multiplication by  $\hat{\Sigma}^{-1/2}$ , then a more direct comparison of the GLS and OSP approaches using libraries of transformed chemical spectra could have been made. Also, most of the results in Bajorski [6] are available from the variable selection literature [35,36]. This “fit background” approach is the basis for the only currently-implemented nonlinear regression methods as described next.

#### 7.4 Nonlinear Bayesian Regression

Heasler et al. [7] report mixed results with an elegant nonlinear Bayesian regression (NLBR) approach that has OSP-like features as well as an attempt to estimate pixel temperature, using a library of 20 possible chemicals. This could also be applied in our setting of a large library where many small subsets of chemicals must each be separately evaluated. The OSP-like features essentially try to fit the background emissivity at each pixel. No linearization simplifications are used, so the nonlinear Planck function and plume absorption are retained. Prior mean and variances must be assigned for all parameters. The parameters include temperature, fits to the background (“endmember”) principal components, and fits to the 20 chemicals at each pixel. There has not yet been a published attempt to apply NLBR to superpixels; however, in a small experiment by the authors using known releases with well-characterized chemical ratios, NLBR applied to superpixels did worse than BMA applied to the same superpixels.

Heasler et al. [7] discuss possible reasons the GLS approach remains competitive compared to this elegant NLBR approach. One reason is the same as mentioned in section 5 involving TES. That is, TES can lead to highly collinear predictors, which is never good for performance or interpretation. A second reason involves the fact that NLBR must also assume a distribution for the error vector that remains after fitting pixel temperature, and fitting to the background endmembers and to the chemical signatures. Apparently, in the real scenes evaluated, this error distribution exhibits spatial clustering, possibly due to a shortcoming of the “fit background” approach or possibly due to sensor artifacts and/or drifting calibration.

Overall, NLBR should be the best approach if all important effects are well modeled, and it can clearly specify which of the physical effects in the background and/or target signature will be modeled and how. However, it has not yet found wide usage. This is partly due to the mixed results (roughly comparable to GLS results), plus it is more complicated and computationally demanding to implement. If the error distributions are specified approximately correctly, and all important effects are well modeled, then NLBR has the advantage of providing error estimates and associated confidence estimates as part of the inference. At the very least, the NLBR approach might be able to provide realistic input for how to structure the error in the signature  $A$  (an approximation to  $A$  is defined in Eq. (6), and other approximations to  $A$  are also commonly used) to extend the generic errors in predictors study initiated in [6].

#### 7.5 Machine Learning

Foy et al. [8] report on “the unreasonable effectiveness of the AMF (GLS)” as compared to a modified support vector machine (SVM) from the machine or statistical learning literature. The experiment involved “matched pairs,” in which chemical signatures were added to each pixel to create data to train the SVM. The experiment ignored error in the chemical signatures because the authors believe that background clutter is the major factor that limits and determines performance. Interestingly, only a two-dimensional vector from each pixel (both with and without the synthetic chemical added) was used to train the SVM, consisting of the GLS estimate  $\hat{\beta}$  and the associated sum of squared residuals (SSR). Interestingly, the humble GLS estimate  $\hat{\beta}$  did well compared to the SVM applied to  $\hat{\beta}$  plus the associated SSR. Other pattern recognition methods could be attempted, using perhaps the entire  $r$  vector rather than a two-dimensional summary.

The use of “matched pairs” in this context is not new, but typically, synthetic chemical is added to a few pixels rather than to each pixel. Adding synthetic chemical effects to each pixel creates a large training set that is useful for characterizing how well one should detect and characterize plumes in a scene that is known to have no plumes. In the analysis of a scene that might contain plumes, some caution regarding this approach is needed, unless it is also assumed as we do here, that any plume of interest is weak and small.

## 8. Related Topics

This section mentions related topics, adds detail to a few topics already mentioned, and indicates research areas.

### 8.1 Band selection

Perhaps surprisingly, more is not always better. In the context of spectral bands (wavelengths), several studies have attempted to detect any of a few selected chemicals that have strong effects (absorption or emission, depending on temperature contrast) in only a selected few bands. Typically, it is better to omit the other bands because they contribute only noise to the inference. Studies involving larger chemical libraries typically rely on stochastic searches for good bands. Of course, which bands are best to retain for analysis depends on the inference algorithm, and generally, it is best to optimize each particular algorithm using some type of band selection if possible [47,48]. The option to combine bands is also worth considering, but is also computationally demanding, assuming that analytical evaluation of performance (plume detection probability for a small false alarm probability) of various band selections is not feasible (it usually isn't).

### 8.2 Clustering

It has been shown that EC distributions describe the  $MD^2$  values better than a Gaussian distribution, even in the case of restricting attention to relatively homogeneous pixels chosen on the basis of a clustering method [20-22]. If clustering pixels [11] is to improve plume detection performance, it must be that a sufficient number of relatively homogeneous pixels can be found in order to estimate  $\Sigma$ .

Section 7 mentioned our sample size experiment applied to data from Figure 1. Such experiments should accompany any image-segmentation strategy to ensure sufficient numbers of pixels. This is a version of the ubiquitous “variance-bias” tradeoff which in this setting refers to the use of many heterogeneous pixels to reduce variance at the expense of increased bias versus the other extreme of using a few homogeneous pixels to reduce bias at the expense of increased variance [35].

### 8.3 Plume modeling

Synthetic plumes can be added to each pixel, one at a time, following Foy et al. [8]. However, for realistic performance claims and algorithm optimization, this requires accurate modeling and no synthetic plume can mimic all the features of real plumes. There are two levels of modeling. The more ambitious level is to begin with a plume dispersion model, weather conditions, and a source term (flow rate of gas up the release stack for example) and then predict plume location, size, and chemical strength. The less ambitious level is to simply assume a certain plume location, size, and chemical strength, and invoke models such as those involving the target signature in section 5 to mimic the anticipated effect on a pixel in the scene of interest. Concerning the first, [49] showed that plume features could be predicted fairly accurately (predicting plume brightness temperature to within approximately 0.2 degrees K), which is encouraging. Concerning the second, it is believed that misspecifying the target signature cannot be avoided, so its effect should be studied more [6, 25], but currently, understanding the impact of background clutter is more heavily emphasized.

Concerning real plume releases, recall that “known releases” usually cannot control weather and wind conditions so the chemical concentrations in the “known” plume have nonnegligible measurement error. If error in the “ground truth” is a concern, recall that one strategy is to assume that chemical ratios can be known fairly accurately regardless of weather conditions. This is because weather conditions typically impact all released chemicals in the same way, thus preserving the ratio that can be predicted if release rates are carefully controlled and measured.

### 8.4 Background modeling

Section 4 described the main models (Gaussian or EC or mixtures thereof) in use for the background pixels. Another potentially useful approach is to use nonparametric density estimation. Recall that [26] used a standard kernel approach, applied to local-mean adjusted relatively homogeneous pixels. As mentioned, the curse of dimensionality [28] is potentially problematic here, so caution is in order. Also, the appropriate background model depends on the context. There have been no published attempts to apply nonparametric density estimation to the background pixels when there is a library of potential chemical signatures. Instead, [26] used density estimation to develop a method to identify unusual background pixels. Perhaps in the context of using the entire  $r$  vector (see the machine learning subsection 7.5) in the final decision regarding whether a chemical is present, it would be useful to apply a density-estimate based rule [23]. As a step toward using more than the scalar-valued GLS estimate when searching for evidence for one chemical, references [8] and [16] show that using the GLS value together with the residual sum of squares has possible merit.

### 8.5 Atmospheric correction

The atmospheric transmission term  $\tau$  can be estimated [31] using in-scene method [32] or computer-intensive atmospheric models [33,34]. Although at least one group working entirely with simulated data reported that accurate atmospheric correction estimates are not important [31], current thinking is the opposite. That is, recently studies have shown the importance of accurate atmospheric corrections [19, 25]. Our experience is that  $\tau$  can be estimated using MODTRAN [34] and/or FASTCODE [33] to within no better than approximately 10% relative error on a per spectral band basis. The 10% is a lower bound because it arises simply from varying inputs (related to atmospheric constituents) to FASTCODE to reflect our uncertainty in the atmospheric constituents while ignoring error in the FASTCODE model itself. Although [31] used modern simulation capability, it seems likely that some features in real data are missing in their simulated data. In addition, perhaps the chosen metric (spectral angle, or correlation, between the “true” and estimated  $\tau$  vector) for evaluating the impact of imperfect atmospheric corrections [31] does not capture all the relevant performance-related information.

To date, there has been no indication of “systematic modeling errors,” that would tend to make the estimated  $\tau$  either higher or lower than the true  $\tau$  for most or all of the spectral bands. Interestingly, atmospheric modeling parameters in MODTRAN are sometimes adjusted until specific features in the image have a desired property [31]. This could lead to systematic modeling errors arising essentially from the phenomenon of “right answer, wrong reason.” That is, because all models are wrong (by definition of the word “model”) to varying degrees, parameter predictions from some models could worsen when adjustments are made to improve other predictions. This issue deserves further study. If there were systematic modeling errors, this would be bad news for inference performance. It is currently assumed that the measurement errors in  $\tau$  are random across bands, leading to less performance degradation than if errors in  $\tau$  were systematic.

### 8.6 Variable selection, model selection, and confidence measures

By “variable selection,” we simply mean to choose the subset of chemicals from the chemical library that are thought to be in the plume. By “model selection” we mean the broader task of selecting a model for the background and the target, and an associated inference approach. Unfortunately, “Bayesian Model Averaging” (BMA) refers to averaging over subsets of candidate chemicals so the jargon is not completely consistent, but should be clear from context. Both variable and model selection continue to generate research [36] in a wide variety of applications.

Confidence measures are always desired, and sometimes required. One approach to assessing confidence is computationally demanding [19,23]. It invokes the concept of developing a reference distribution by repeating the inference approach (using, for example, BMA and/or the “pick winner” approach with penalized likelihood) for many randomly-selected subsets of the same number of contiguous pixels as are in the purported plume region. For example, if the BMA-based estimate of the

probability of chemical C is 0.95, we compare 0.95 to the reference distribution of chemical probability estimates. We refer here to probability “estimates” because sections 4-7 and our empirical observations suggest that the BMA-based probability estimates might not be sufficiently well calibrated to interpret in the natural fashion as a probability. For example, if 0.95 is the  $p = 0.80$  quantile, then there is a 20% false alarm rate in using 0.95 as a decision threshold (rather than the desired interpretation of a 5% false alarm rate).

Recall that calculating the exceedance metric for comparing two probability distributions involved many ( $K$ ) quantiles [21]. Suppose instead we choose a particular false alarm rate such as 0.01 (or a particular threshold such as 3.0 in similar contexts). Then, for example, the empirical distribution of chemical probability estimates based on all analyzed pixels could be compared to a simulated reference distribution assuming a particular distribution (often Gaussian unless a maximum is taken over the entire chemical library) for these estimates to determine whether the simulated distribution was adequate for selecting thresholds. In effect, we simply choose  $K = 1$  and apply the exceedance metric or a version of the metric that compares fractions exceeding specified thresholds rather than quantiles.

More qualitatively, in our experience, potential false positives such as the white ovals in Figure 1 can sometimes be associated with ground features, and thus rejected without formal statistical evaluation. However, the performance (the false negative rate for a given low fixed positive rate) of such a strategy is difficult to quantify.

One relatively simple strategy to provide an estimated confidence measure that we have not seen in the literature is as follows. First locate a candidate plume region in the scene, then add and/or subtract the estimated plume effect and compute the corresponding residuals. Commonly-used residual summaries such as the residual sum of squares could then provide a confidence measure. Foy et al. [8] is the first to consider using both the residual sum of squares and the GLS estimate  $\hat{\beta}$  for inferring which pixels contain plume effects. We believe that using the entire residual vector could also be valuable for inference and for associated confidence measures.

### 8.7 Errors in predictors

Reference [6] included an empirical study of the impact of error in the predictor matrix  $A$  in a narrow setting having a four-chemical library, without attempting to adjust parameter estimates on the basis of knowledge regarding the magnitude of errors in  $A$ .

It would be useful to model the error in  $A$  in Eq. (5) so that techniques from the errors in predictors literature [50] could be applied. Reference [7] made initial progress toward this goal by assigning prior probabilities to the plume temperature and background fit parameters, in a type of TES strategy via NLBR. Recall that NLBR has not yet proven itself in IR inferences, possibly because it is more sensitive to model and error distribution assumptions than are simpler methods such as the GLS.

### 8.8 Temperature-emissivity separation

There are several reviews available for the TES task ([13], [14], and the references in [14]), so we will not review TES here. Note however, that Figure 4 indicates a difficulty in separating temperature

and emissivity because of the lack of a well-defined single linear relation between the pairs of spectral channels across pixels. Such a relation is required for many existing TES approaches.

Currently, the only data-driven, scene-specific approach for assessing how well the TES task is performing in our context of plume detection seems to be the very computationally demanding approach of injecting synthetic plumes into real scenes and evaluating the reasonableness of the various estimates. Because the need for TES introduces errors into the predictors (chemical signatures), these errors need to be accounted for in performance assessments. At the least, when adding synthetic plumes with signatures, the assumed signature values used in GLS or any other approach should differ from the values used to modify the background radiance  $r$ . The magnitude of the difference should be determined by the estimated error magnitudes arising from the TES task.

### 8.9 Constrained regression

Gallagher et al. [45] report promising results using constrained regression to fit the background endmembers, followed by constrained regression to fit the chemical signatures (assuming either absorption or emission but not both). The NLBR approach is more general and easily accommodates such constraints. Although NLBR has not yet been developed into an effective strategy, it could also enforce such constraints.

### 8.10 Outliers/robustness

In our experience, “bad pixels” (due to sensor malfunctions or unknown causes) tend to occur in rows or columns of contiguous pixels. Several reasonable approaches for outlier detection have been implemented and have improved performance, simply by omitting any bad pixels.

Robustness to outliers is a large topic, and it is useful to consider robust-to-outlier methods for estimating  $\Sigma$  that describes the background clutter [51]. Such methods are not in common use in IR analysis. Several options are not overly computationally expensive, so should be considered [51].

### 8.11 Chemical spectral data bases and background emissivity data bases

There are several commonly-used libraries of chemical spectra, typically provided with accuracy estimates [52]. It is usually necessary to transform the tabled chemical spectral values to the resolution and spectral channels that are relevant for the particular instrument under study; this introduces another error source. Regarding background emissivity data bases, it is important to realize that the chemical spectra are temperature-dependent and that the same material (mowed grass) can have a substantially different emissivity depending on object height, moisture content, viewing angle, shading, etc. Therefore, it can be extremely time consuming to choose appropriate endmembers from such libraries for the “fit background” OSP-like inference option on the basis of what is known about the particular scene. In all cases, errors in these spectra are nonnegligible, although perhaps are negligible compared to the background clutter issue.

## 9. Definitions

AMF: adaptive matched filter, which we have written as  $\hat{\beta} = (A^T \hat{\Sigma}^{-1} A)^{-1} A^T \hat{\Sigma}^{-1} r$ . The covariance  $\hat{\Sigma}^{-1}$  is estimated from the scene.

CFAR: constant false alarm rate

Clutter: spatial heterogeneity in multi-dimensional radiance measurements, as in Figure 1.

GLS: generalized least squares. This is the same as the AMF.

Optically thin plume: a linear relation exists between plume absorption (or emission) and either of plume thickness, chemical concentration, or the temperature contrast between the ground and plume.

Brightness temperature: the temperature value in Planck's blackbody radiation function that corresponds to the observed radiance

Radiance: A measure that describes the amount of electromagnetic radiation that passes through or is emitted from a particular area, and falls within a given solid angle in a specified direction. The SI unit of radiance is watts per steradian per square metre ( $\text{W}\cdot\text{sr}^{-1}\cdot\text{m}^{-2}$ ).

TES: temperature-emissivity separation is an attempt to separately estimate emissivity and temperature for a given pixel. For example, for a pixel containing grass, the ground radiance is

$\varepsilon_{\text{grass}} L^{BB}$ , where  $L^{BB} = \frac{c_1 V^3}{\exp(\frac{c_2 V}{T}) - 1}$  is the Planck blackbody radiation, which depends on ground

temperature  $T$ . Although  $\varepsilon_{\text{grass}}$  has as many unknowns as observations (spectral channels),  $T$  is a single unknown, so provided various assumptions regarding spectral and/or spatial smoothness in  $\varepsilon_{\text{grass}}$  values are adequate, TES can be fairly successful in some situations.

## 10. Summary

This review focused on cases where real or synthetic plumes contain one, two, or at most three gases from a library of hundreds to thousands of possible chemicals. Typically there are approximately 10000 to 500,000 pixels in a scene that has one to several weak (optically thin), small (impacting only a relatively few pixels) plumes. The number of pixels corresponding to a given plume is tens to hundreds. It is important to recognize that only chemical plumes having nonzero thermal contrast with the ground can be detected using passive IR.

Even in this relatively confined context, many technical issues regarding physical models, simplifying assumptions, and error sources arise that impact inference performance, as discussed. At the time of this review, the humble GLS approach remains surprisingly effective compared to methods having a more explicit "fit background" aspect. Most studies that come to this conclusion regarding GLS have considered one chemical at a time, representing a context where it is known that if a plume is present, then it consists of the chosen chemical.

Many believe that background clutter (arising from emissivity variation corresponding to different ground components such as water, grass, asphalt, trees, etc., and from ground temperature variation among pixels) is the main factor that limits performance. This is the reason that EC distributions have

been proposed for hyperspectral IR data. EC distributions do not fit all aspects of the radiance values very well, but do fit the Mahalanobis distances that are based on radiance values better than the multivariate Gaussian does.

### Acknowledgements

The authors gratefully acknowledge support from the U.S. Department of Energy and acknowledge helpful conversations with Brian McVey, Herb Fry, Bernie Foy, and James Theiler from Los Alamos National Laboratory and with Pat Heasler and Lawrence Chilton from Pacific Northwest National Laboratory.

### References

1. AVIRIS Free Standard Data Products, Jet Propulsion Laboratory (JPL), National Aeronautics and Space Administration (NASA). <http://aviris.jpl.nasa.gov/html/aviris.freedata.html>.
2. Vane, G.; Green, R.; Chrien, T.; Enmark, H.; Hansen, E.; Porter, W. The Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), *Remote Sensing of the Environment*. **1993**, *44*, 127-143.
3. Bernhardt M.; Heather, J.; Smith, M. New Models for Hyperspectral Anomaly Detection and Un-Mixing, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XII, edited by Sylvia S. Shen, Paul E. Lewis. *Proc. SPIE*. **2005**, *5806*, 720-730.
4. Manolakis, D.; Shaw, G. Detection Algorithms for Hyperspectral Imaging Applications. *IEEE Signal Processing Magazine*. **2002**, *19*(1), 29-43.
5. Bajorski, P. Analytical Comparison of the Matched Filter and Orthogonal Subspace Projection Detectors in Structured Models for Hyperspectral Images, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XII, edited by Sylvia S. Shen, Paul E. Lewis, *Proc. SPIE*. **2006**, *6233*, 1-12.
6. Burr, T.; Fry, H.; McVey, B.; Sander, E. Chemical Identification using Bayesian Model Selection, 2002 *Proceedings of the American Statistical Association, Section on Physical and Engineering Sciences*[CD-ROM], Ann Arbor, Michigan: American Statistical Association, Los Alamos National Laboratory Unclassified Report, LA-UR-02-7281, available from T. Burr upon request.
7. Heasler, P.; Hylden, J. Technical Letter Report for Non-Linear methods Task: Fits of Nonlinear Bayesian Regression (NBLR) to Tanasi and Polecat Images, Pacific Northwest National Laboratory Official Use Only Report, 2005.
8. Foy, B.; Theiler, J.; Fraser, A. Unreasonable Effectiveness of the Adaptive Matched Filter, to appear in: *Proc. MSS (Military Sensing Symposia) Passive Sensors Conference*, **2006**.
9. Manolakis, D.; Siracusa, C.; Marden, D.; Shaw, G., Hyperspectral Adaptive Matched Filter Detectors: Practical Performance Comparison. *Proc. SPIE*. **2001**, *4381*, 18-33.
10. Messinger, D. Gaseous Plume Detection in Hyperspectral Images: a Comparison of Methods, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery X. *Proc. SPIE*. **2004**, *5425*, 592-603

11. Funk, C.; Theiler, J.; Roberts, D.; Borel, C. Clustering to Improve Matched Filter Detection of Weak Gas Plumes in Hyperspectral Thermal Imagery. *IEEE Trans. Geoscience and Remote Sensing*. **2001**, *39*, 1410-1420.
12. Schott, J. *Remote Sensing*, Oxford Press: New York, 1997.
13. Kealy, P.; Hook, S. Separating Temperature and Emissivity in Thermal Multispectral Scanner Data: Implications for Recovering Land Surface Temperatures. *IEEE Transactions on Geoscience and Remote Sensing*. **1993**, *31*(6): 1155-1164.
14. Lausten, K.; Resmini, R. A New Approach to Infer Surface Emissivity Parameters from Longwave Infrared Hyperspectral Measurements, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XII. *Proc. SPIE*. **2006**, *6233*, 1-10.
15. Manolakis, D.; Amico, F. A Taxonomy of Algorithms for Chemical Vapor Detection with Hyperspectral Imaging Spectroscopy. *Proc. SPIE*. **2005**, *5795*, 125-133.
16. Theiler, J.; Foy, B.; Fraser, A. Nonlinear Signal Contamination Effects for Gaseous Plume Detection in Hyperspectral Imagery, Los Alamos National Laboratory Unclassified Report LAUR-06-1996, to appear. *Proc. SPIE*. **2007**.
17. Foy, B.; Petrin, R.; Quick, R.; Shimada, T.; Tsee, J. Comparisons Between Hyperspectral Passive and Multispectral Active Sensor Measurements. *Proc. SPIE*. **2002**, *4722*, 98-109.
18. Theiler, J.; Foy, B.; Fraser, A. Characterizing non-Gaussian Clutter and Detecting Weak Gaseous Plumes in Hyperspectral Imagery, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI. *Proc. SPIE*. **2005**, *5806*, 182-193.
19. McVey, B.; Burr, T.; Fry, H. Distribution of Chemical False Positives for Hyperspectral Image Data, Los Alamos National Laboratory Restricted Release Technical Report LA-CP-02-521, cited with permission, 2003.
20. Manolakis, D.; Marden, D.; Kerekes, J.; Shaw, G. On the Statistics of Hyperspectral Imaging Data. *Proc. SPIE*. **2001**, *4381*, 308-316.
21. Marden, D.; Manolakis, D. Using Elliptically Contoured Distributions to Model Hyperspectral Imaging Data and Generate Statistically Similar Synthetic Data. *Proc. SPIE*. **2004**, *5425*, 558-572.
22. Marden, D.; Manolakis, D. Modeling Hyperspectral Imaging Data. *Proc. SPIE*. **2003**, *5093*, 253-262.
23. Burr, T.; Fry, H.; McVey, B. Characterizing Clutter in the Context of Detecting Weak Gaseous Plumes using the SEBASS Sensor, Los Alamos National Laboratory Restricted Release Report, LA-CP-06-0982, cited with permission, 2006.
24. Mitchell, H.; Jellison, G.; Miller, D.; Salvaggio, C.; Miller, C. The Importance of Background in the Detection and Identification of Gas Plumes Using Emissive Infrared Hyperspectral Sensing. *Proc. SPIE*. **2003**, *5093*, 206-217.
25. Sheen, D.; Gallagher, N.; Sharpe, S.; Anderson, K.; Schultz, J. Impact of Background and Atmospheric Variability on Infrared Hyperspectral Chemical Detection Sensitivity. *Proc. SPIE*. **2003**, *5093*, 218-229.
26. Clare, P.; Bernhardt, M.; Oxford, W.; Murphy, S.; Godfree, P.; Wilkinson, V. New Approach to Anomaly Detection in Hyperspectral Images, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XII. *Proc. SPIE*. **2003**, *5093*, 17-28.

27. Lucey P.; Williams, T.; Mignard, M.; Julian, J.; Kobubun, D.; Allen, G.; Hampton, D.; Schaff, W.; Schlangen, M.; Winter, E.; Kendall, W.; Stocker, A.; Horton, K.; Bowman A. AHI- an Airborne Long Wave Infrared Hyperspectral Imager, Airborne Reconnaissance XXII. *Proc. SPIE*. **1998**, 3431, 36-43.
28. Silverman, B. *Density Estimation for Statistics and Data Analysis*, 1st Ed, Chapman and Hall Ltd: London and New York, 1986.
29. Garcia-Escudero, L.; Gordaliza, A. Generalized Radius Processes for Elliptically Contoured Distributions. *Journal of the American Statistical Association* **2005**, 100(471), 1036-1045.
30. Manolakis, D.; Rossacci, M.; Cipar, J.; Lockwood, R.; Cooley, T.; Jacobson, J. Statistical Characterization of Natural Hyperspectral Backgrounds Using *t*-Elliptically Contoured Distributions, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI. *Proc. SPIE*. **2005**, 5806, 56-65.
31. Miller, B.; Messinger, D. The Effects of Atmospheric Compensation Upon Gaseous Plume Signatures, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI. *Proc. SPIE*. **2005**, 5806, 229-236.
32. Young, S.; Johnson, B.; Hackwell J. An In-scene Method for Atmospheric Compensation of Thermal Hyperspectral Data. *J. Geophys. Res. Atm.* **2002**, 107, 4774-4774.
33. FASCODE for atmospheric transmission calculations, see <http://www.vs.af.mil/ProductLines/IR-Clutter/fascode.aspx>
34. Berk, A.; Bernstein, L.; Robertson, D. MODTRAN: A Moderate Resolution Model for LOWTRAN 7, GL-TR-890122,AD-A214-337, Geophysics Laboratory, Hamscom Air Force Base, Mass, 1989.
35. Weisberg, S. *Applied Linear Regression Analysis*, 2<sup>nd</sup> edition, Wiley: New York, 1985.
36. Casella, G.; Moreno, E. Objective Bayesian Variable Selection. *Journal of the American Statistical Association* **2006**, 101 (473), 157-167.
37. Scharf, L. *Statistical Signal Processing*, Addison-Wesley, New York 1991.
38. Kelly, E. An Adaptive Detection Algorithm. *IEEE Trans. Aerospace and Electronic Systems* **1986**, 22, 115-127.
39. Friedman, J. Regularized Discriminant Analysis. *Journal of the American Statistical Association* **1989**, 84 (405), 165-175.
40. Raftery, A.; Madigan, D.; Hoeting J. Bayesian Model Averaging for Linear Regression Models. *Journal of the American Statistical Association* **1997**, 92, 179-191.
41. Burr, T.; Fry, H.; McVey, B.; Sander, E.; Cavanaugh, J.; Neath, A. Performance of Variable Selection Methods in Regression using Variations of the Bayesian Information Criterion, Los Alamos National Laboratory Unclassified Report LAUR-05-6324, submitted, 2005, available from T. Burr upon request.
42. DiCiccio, T.; Kass, R.; Raftery, A.; Wasserman, L. Computing Bayes Factors by Combining Simulation and Asymptotic Approximations. *Journal of the American Statistical Association* **1997**, 92, 903-915.
43. Kass R.; Raftery, A. Bayes Factors. *Journal of the American Statistical Association* **1995**, 90, 773-795.

44. Harsanyi, J.; Chang, C. Hyperspectral Image Classification and Dimensionality Reduction: an Orthogonal Subspace Projection Approach. *IEEE Trans. Geoscience and Remote Sensing*. **1994**, *32*, 779-785.
45. Gallagher, N.; Sheen, D.; Shaver, J.; Wise, B.; Shultz, J. Estimation of Trace Vapor Concentration-Pathlength in Plumes for Remote Sensing Applications From Hyperspectral Images. *Proc. SPIE*. **2003**, *5093*, 184-194.
46. Foy, B.; Theiler, J. Scene Analysis and Detection in Thermal Infrared Remote Sensing Using Independent Component Analysis. *Proc. SPIE*. **2004**, *5439*, 131-139.
47. Nakariyakul, S.; Casasent, D. Adaptive Branch and Bound Algorithm for Use on Hyperspectral Data, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI. *Proc. SPIE*. **2006**, *6233*, 175-184.
48. Paskaleva, B.; Hayat, M.; Tyo, J.; Wang, Z.; Martinez, M. Feature Selection for Spectral Sensors with Overlapping Noising Spectral Bands, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI. *Proc. SPIE*. **2006**, *6233*, 773-782.
49. Griffin, M.; Czerwinski, R.; Upham, C.; Wack, E.; Burke, H. A Procedure for Embedding Effluent Plumes into LWIR Imagery, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI. *Proc. SPIE*. **2005**, *5806*, 78-87.
50. Carroll, R.; Ruppert, D.; Stefanski, L. *Measurement Error in Nonlinear Models*, Chapman and Hall, London, 1996.
51. Rousseeuw, P.; van Zomeren, B. Unmasking Multivariate Outliers and Leverage Points (with discussion). *Journal of the American Statistical Association* **1990**, *85*(411), 633-651.
52. Johnson, T.; Sams, R.; Sharpe, S. The PNNL Quantitative Infrared DataBase for Gas-Phase Sensing: Spectral Library for Environmental, Hazmat, and Public Safety Standoff Detection, Chemical and Biological Point Sensors for Homeland Defense. *Proc. SPIE*. **2004**, *5269*, 159-167.