Veritas: an admissible detector for targets of unknown strength

James Theiler

Space Data Science and Systems Group,
Los Alamos National Laboratory,
Los Alamos, NM 87545, USA

ABSTRACT

Inspired by the comic faux-Latin aphorism *"veritas duo sigma"* (truth at two sigmas), an approach to target detection is proposed based on a likelihood ratio test in which the unknown target strength is treated as known, with strength chosen to correspond to a minimal level of detectability. For Gaussian distributions, this strength typically corresponds to two or three sigmas. This detector is *admissible*, which means that there is no other detector that is uniformly superior to it. The simplicity of the *veritas* detector permits closed-form solutions to be derived for a variety of signal detection problems. In a series of numerical experiments, these simple detectors are compared to traditional detectors, such as the locally most powerful detector and the generalized likelihood ratio test detector.

Keywords: Adaptive matched filter, Clutter, Clairvoyant fusion, Composite hypothesis testing, Elliptically-contoured distribution, Generalized likelihood ratio test, Hyperspectral imagery, Target detection

1. INTRODUCTION

For additive signals on Gaussian clutter, the linear matched filter[2–5] is the uniformly most powerful detector (UMP). It is the optimal detector for any strength of additive signal. For even mildly complicated scenarios, however, for non-Gaussian clutter, or for replacement-model or Beer's Law signals, there is no UMP detector, and optimum detection depends on the target signal strength.

If the target strength were known, then the so-called clairvoyant detector,[6] which is based on a simple likelihood ratio test, would provide optimal detection. Our interest here, however, is in scenarios in which we do not know the target strength. This is a composite hypothesis testing problem, and in most scenarios, it does not admit a single unambiguously optimal solution.

The traditional approach to the general composite hypothesis testing problem is the generalized likelihood ratio test (GLRT). Here a maximum likelihood estimate is made of the unknown strength, and that estimate is used in the clairvoyant detector. There is no guarantee that the GLRT is optimal, however, and examples can be constructed (*e.g.*, see [7]) for which the GLRT is dominated by a Bayesian detector. Two generalizations of the GLRT, penalized likelihood[8,9] and clairvoyant fusion,[10,11] unfortunately share the sub-optimality property.

A less common alternative to the GLRT is sometimes called the locally most powerful (LMP) detector,[6] which identifies the optimal detector in the limit that the target strength goes to zero. The motivation for using this statistic is that weak targets are hardest to find; for strong targets, which are easier to find, we can get away with less-than-optimal detectors.

The approach suggested here follows the same line of reasoning, but instead of optimizing on the very weakest target strengths, it optimizes on target strengths that are, in some sense, just detectable. Again, for stronger targets, because they are easier to find, we can get away with suboptimal detectors. And for weaker targets, we aren't going to find them anyway. So we aim for what we hope is a sweet spot. Where that sweet spot is (and whether it is a single spot, or maybe some kind of average over a range) will depend on the the application at hand. But a reasonable place to start is with a desired false alarm rate that we would like to achieve. Targets

Email: jt@lanl.gov

are sought whose strength just permits them to be detected at the desired false alarm rate; and the detector is optimized for those targets.

While this approach has wider application, the motivating example here is pixel-wise target detection in hyperspectral imagery. For each pixel in a hyperspectral image, we ask: is there a target present?

2. GENERIC FORMULATION FOR A TARGET IN CLUTTER

Let $\mathbf{x} \in \mathbb{R}^d$ represent the value (reflectance, usually, but possibly also radiance or even uncalibrated "digital number") for a pixel in an image with d spectral channels. Let $\boldsymbol{\xi}$ be a function that describes the effect of target on background. If $\mathbf{z}$ is the background spectrum (*i.e.*, the spectrum of a pixel with no target), and $\mathbf{x}$ is the spectrum of that pixel when a target is present, then we write $\mathbf{x} = \boldsymbol{\xi}(\mathbf{z})$.

We can write $P_{\text{bkg}}(\mathbf{z})$ as the probability density associated with the background pixels $\mathbf{z}$. Using the usual formula for change-of-variables in probability distributions, we can say that

$$P_{\text{target}}(\mathbf{x}) = P_{\text{bkg}}(\boldsymbol{\xi}^{-1}(\mathbf{x})) \left| \frac{d\boldsymbol{\xi}}{d\mathbf{x}} \right|^{-1}, \tag{1}$$

and from this the likelihood ratio is given by

$$\mathcal{L}(\mathbf{x}) = \frac{P_{\text{target}}(\mathbf{x})}{P_{\text{bkg}}(\mathbf{x})} = \frac{P_{\text{bkg}}(\boldsymbol{\xi}^{-1}(\mathbf{x}))}{P_{\text{bkg}}(\mathbf{x})} \left| \frac{d\boldsymbol{\xi}}{d\mathbf{x}} \right|^{-1}. \tag{2}$$

This is in some sense *the* fundamental equation for target detection. If we have a model for the background clutter (encoded in the distribution P_{bkg}) and a model for target-background interaction (encoded in the function $\boldsymbol{\xi}$), then Eq. (2) provides the optimal detector for that target in that background. The issue we face in this exposition is that $\boldsymbol{\xi}$ is *not* known; it depends on some measure of target strength (or abundance), which we will hereafter denote a.

3. ADDITIVE TARGET MODEL

For an additive target, we write $\mathbf{x} = \boldsymbol{\xi}(\mathbf{z}) = \mathbf{z} + a\mathbf{t}$, where $\mathbf{t} \in \mathbb{R}^d$ is the known target signature and a is the unknown scalar-valued measure of target strength. The detector in Eq. (2) becomes

$$\mathcal{L}(a, \mathbf{x}) = \frac{P_{\text{bkg}}(\mathbf{x} - a\mathbf{t})}{P_{\text{bkg}}(\mathbf{x})}. \tag{3}$$

3.1 Additive target in Gaussian clutter

For a Gaussian background with mean $\boldsymbol{\mu}$ and covariance R, we can write

$$2 \log \mathcal{L}(a, \mathbf{x}) = \mathcal{A}(\mathbf{x}) - \mathcal{A}(\mathbf{x} - a\mathbf{t}) \tag{4}$$

where

$$\mathcal{A}(\mathbf{x}) = (\mathbf{x} - \boldsymbol{\mu})' R^{-1} (\mathbf{x} - \boldsymbol{\mu}) \tag{5}$$

is the Mahalanobis distance to the mean, widely used as the "RX" anomaly detector.[12] Then

$$2 \log \mathcal{L}(a, \mathbf{x}) = 2a\mathbf{t}' R^{-1} (\mathbf{x} - \boldsymbol{\mu}) - a^2 \mathbf{t}' R^{-1} \mathbf{t} \tag{6}$$

and a further monotonic transform* gives the familiar matched filter

$$\mathcal{D}(\mathbf{x}) = \frac{2 \log \mathcal{L}(a, \mathbf{x}) + a^2 \mathbf{t}' R^{-1} \mathbf{t}}{2a} = \mathbf{t}' R^{-1} (\mathbf{x} - \boldsymbol{\mu}), \tag{7}$$

*The transform is monotonic as long as $a > 0$.

which does *not* depend on a. This matched filter is the optimal detector for all positive target strength values a; thus it is the uniformly most powerful (UMP) detector.

Note that in the absence of target, this detector has mean zero and variance given by

$$\mathrm{Var}[\mathcal{D}(\mathbf{x})] = \left\langle \mathbf{t}' R^{-1}(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})R^{-1}\mathbf{t} \right\rangle = \mathbf{t}' R^{-1}\mathbf{t} \tag{8}$$

Meanwhile, in the presence of target with strength a, the mean value of the detector is given by $\mathrm{Mean}[\mathcal{D}(\mathbf{x})] = a\mathbf{t}'R^{-1}\mathbf{t}$. Thus, the detector on average achieves a "one sigma" (*i.e.*, one standard deviation) detection when this mean value is equal to the square root of the variance, and this occurs when $a = 1/\sqrt{\mathbf{t}'R^{-1}\mathbf{t}}$. We refer to this as the characteristic target strength:

$$a_o = 1/\sqrt{\mathbf{t}'R^{-1}\mathbf{t}} \tag{9}$$

3.2 Additive target in elliptically-contoured clutter

A slighly more complicated case arises when the background is not Gaussian. Here, we consider the multivariate t-distribution, which shares many properties of the Gaussian (such as a mean $\boldsymbol{\mu}$ and covariance R) but it has a fatter tail, and is often considered a better model for hyperspectral data.[13] This distribution is given by

$$P_{\mathrm{bkg}}(\mathbf{x}) = c[\nu - 2 + \mathcal{A}(\mathbf{x})]^{-(\nu+d)/2} \tag{10}$$

where c is a constant, and ν is a paramter that characterizes how fat the tail is. For $\nu \to \infty$, the distribution converges to a Gaussian; as $\nu \to 2$, the distribution gets fatter and fatter tails until for $\nu \leq 2$, the second moment does not exist, and so the covariance of the distribution is not formally defined.

For a known target strength a, we can compute the clairvoyant detector by taking an appropriate monotonic transform of the likelihood ratio. Here,

$$\mathcal{L}(a,\mathbf{x}) = \frac{[\nu - 2 + \mathcal{A}(\mathbf{x} - a\mathbf{t})]^{-(\nu+d)/2}}{[\nu - 2 + \mathcal{A}(\mathbf{x})]^{-(\nu+d)/2}} \tag{11}$$

so

$$1 - \mathcal{L}(a,\mathbf{x})^{-2/(\nu+d)} = \frac{2a\mathbf{t}'R^{-1}(\mathbf{x}-\boldsymbol{\mu}) - a^2\mathbf{t}'R^{-1}\mathbf{t}}{\nu - 2 + \mathcal{A}(\mathbf{x})} \tag{12}$$

and one form of the clairvoyant is given by

$$\mathcal{D}(a,\mathbf{x}) = (\nu-1)[1 - \mathcal{L}(a,\mathbf{x})^{-2/(\nu+d)}] = \mathcal{F}_\nu^2(\mathbf{x})\left[2a\mathbf{t}'R^{-1}(\mathbf{x}-\boldsymbol{\mu}) - a^2/a_o^2\right] \tag{13}$$

where a_o is defined in Eq. (9) and

$$\mathcal{F}_\nu(\mathbf{x}) = \sqrt{\frac{\nu-1}{\nu - 2 + \mathcal{A}(\mathbf{x})}}. \tag{14}$$

We cannot actually use the clairvoyant formula in practice, because it requires that we know a. For the Gaussian in the previous section, we sidestepped this issue by performing a monotonic transform that eliminated the dependence on a. But that is not possible here.

The *veritas* solution is to replace a in Eq. (13) with a small multiple of the *characteristic* signal strength a_o. This leads to

$$\mathcal{D}_{veritas}(n,\mathbf{x}) = \frac{\mathcal{D}(na_o,\mathbf{x})}{2n} = \mathcal{F}_\nu^2(\mathbf{x})\left[\frac{\mathbf{t}'R^{-1}(\mathbf{x}-\boldsymbol{\mu})}{\sqrt{\mathbf{t}'R^{-1}\mathbf{t}}} - \frac{n}{2}\right] \tag{15}$$

where n is the "number of sigmas" you seek in your detector. If you want good performance at a very small false alarm rate, then you have to be looking for stronger targets, and so n will be larger.

The LMP solution seeks the detector that is optimal in the $a \to 0$ limit. This is obtained from Eq. (15) with $n = 0$. This is not recommended for practical purposes, however, because such weak targets are essentially undetectable; to say it another way, detecting such weak targets would require a very high (and typically unacceptable) false alarm rate.

Table 1. Additive model with multivariate t-distributed background

Detector	Expression
Clairvoyant	$\mathcal{D}(a,\mathbf{x}) = \mathcal{F}_\nu^2(\mathbf{x})\left[a\mathbf{t}'R^{-1}(\mathbf{x}-\boldsymbol{\mu}) - \frac{1}{2}a^2\mathbf{t}'R^{-1}\mathbf{t}\right]$
Veritas	$\mathcal{D}(n,\mathbf{x}) = \mathcal{F}_\nu^2(\mathbf{x})\left[\mathbf{t}'R^{-1}(\mathbf{x}-\boldsymbol{\mu}) - \frac{1}{2}n\sqrt{\mathbf{t}'R^{-1}\mathbf{t}}\right]$
LMP	$\mathcal{D}(\mathbf{x}) = \mathcal{F}_\nu^2(\mathbf{x})\left[\mathbf{t}'R^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]$
GLRT	$\mathcal{D}(\mathbf{x}) = \mathcal{F}_\nu(\mathbf{x})\left[\mathbf{t}'R^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]$
AMF	$\mathcal{D}(\mathbf{x}) = \mathbf{t}'R^{-1}(\mathbf{x}-\boldsymbol{\mu})$
ACE	$\mathcal{D}(\mathbf{x}) = \mathbf{t}'R^{-1}(\mathbf{x}-\boldsymbol{\mu})/\sqrt{\mathcal{A}(\mathbf{x})}$
RX	$\mathcal{A}(\mathbf{x}) = (\mathbf{x}-\boldsymbol{\mu})'R^{-1}(\mathbf{x}-\boldsymbol{\mu})$

The large a limit is also potentially of interest. For this case, the additive target on multivariate t background, the $a \to \infty$ limit leads to $\mathcal{D}(\mathbf{x}) = \lim_{n\to\infty} 2\mathcal{D}_{veritas}(n,\mathbf{x})/n = -\mathcal{F}_\nu(\mathbf{x})^2$ which is equivalent to $\mathcal{A}(\mathbf{x})$. Thus, the RX anomaly detector becomes the optimal target detector for very strong targets on an EC background.

The more traditional approach to the problem of unknown target strength is the generalized likelihood ratio test (GLRT). Since we don't know the actual a, we estimate it by maximizing $P_{\text{target}}(\mathbf{x}) = P_{\text{bkg}}(\mathbf{x} - a\mathbf{t})$. This leads to an estimator

$$\widehat{a} = \mathbf{t}'R^{-1}(\mathbf{x}-\boldsymbol{\mu})/\mathbf{t}'R^{-1}\mathbf{t}. \tag{16}$$

which we can substitute into Eq. (13)[†], take the square root as a monotonic transform[‡], and obtain the GLRT solution reported in [14]:

$$\mathcal{D}_{GLRT}(\mathbf{x}) = \sqrt{\mathcal{D}(\widehat{a},\mathbf{x})} = \mathcal{F}_\nu(\mathbf{x})\left[\frac{\mathbf{t}'R^{-1}(\mathbf{x}-\boldsymbol{\mu})}{\sqrt{\mathbf{t}'R^{-1}\mathbf{t}}}\right] \tag{17}$$

4. REPLACEMENT TARGET MODEL

For solid subpixel targets, the strength of the target signal is proportional to the area of the target, and since the background over this area is occluded by the target, its magnitude is correspondingly reduced. In particular, if a is the area of the target relative to the area of a pixel, then

$$\mathbf{x} = \boldsymbol{\xi}(\mathbf{z}) = (1-a)\mathbf{z} + a\mathbf{t}. \tag{18}$$

Here, the likelihood ratio is given by

$$\mathcal{L}(a,\mathbf{x}) = \frac{P_{\text{target}}(\mathbf{x})}{P_{\text{bkg}}(\mathbf{x})} = (1-a)^{-d}\frac{P_{\text{bkg}}\left(\dfrac{\mathbf{x}-a\mathbf{t}}{1-a}\right)}{P_{\text{bkg}}(\mathbf{x})} \tag{19}$$

4.1 Replacement model on Gaussian background

The log likelihood for a Gaussian background is given by

$$\begin{aligned}\log\mathcal{L}(a,\mathbf{x}) &= -d\log(1-a) + \log P_{\text{bkg}}\left(\frac{\mathbf{x}-a\mathbf{t}}{1-a}\right) - \log P_{\text{bkg}}(\mathbf{x}) \\ &= -d\log(1-a) - \tfrac{1}{2}\mathcal{A}\left(\frac{\mathbf{x}-a\mathbf{t}}{1-a}\right) + \tfrac{1}{2}\mathcal{A}(\mathbf{x})\end{aligned} \tag{20}$$

[†]Note, to get the GLRT solution, it is important that we substitute $\widehat{a}$ into an expression that is not only a monotonic transform of the likelihood, it must be a monotonic transform that does not depend on a.

[‡]In theory, one should worry about negative values; in practice, Eq. (17) works fine even when $\mathbf{t}'R^{-1}(\mathbf{x}-\boldsymbol{\mu}) < 0$.

Using

$$\mathcal{A}\left(\frac{\mathbf{x}-a\mathbf{t}}{1-a}\right) = \frac{1}{(1-a)^2}\mathcal{A}(\mathbf{x}-a(\mathbf{t}-\boldsymbol{\mu})) = \frac{1}{(1-a)^2}\left[\mathcal{A}(\mathbf{x}) - 2a(\mathbf{t}-\boldsymbol{\mu})'R^{-1}(\mathbf{x}-\boldsymbol{\mu}) + a^2\mathcal{A}(\mathbf{t})\right], \tag{21}$$

we can write

$$\log\mathcal{L}(a,\mathbf{x}) = -d\log(1-a) - \frac{\frac{1}{2}\left[\mathcal{A}(\mathbf{x}) - 2a(\mathbf{t}-\boldsymbol{\mu})'R^{-1}(\mathbf{x}-\boldsymbol{\mu}) + a^2\mathcal{A}(\mathbf{t})\right]}{(1-a)^2} + \tfrac{1}{2}\mathcal{A}(\mathbf{x}) \tag{22}$$

$$= -d\log(1-a) + \frac{a(\mathbf{t}-\boldsymbol{\mu})'R^{-1}(\mathbf{x}) - \frac{1}{2}a^2\mathcal{A}(\mathbf{t}) - (a-\frac{1}{2}a^2)\mathcal{A}(\mathbf{x})}{(1-a)^2}. \tag{23}$$

So now we can obtain a detector by taking a monotonic transform:

$$\mathcal{D}(a,\mathbf{x}) = \frac{(1-a)^2\left[\log\mathcal{L}(a,\mathbf{x}) + d\log(1-a)\right]}{a} + \tfrac{1}{2}a\mathcal{A}(\mathbf{t}) \tag{24}$$

$$= (\mathbf{t}-\boldsymbol{\mu})'R^{-1}(\mathbf{x}-\boldsymbol{\mu}) - (1-\tfrac{1}{2}a)\mathcal{A}(\mathbf{x}). \tag{25}$$

This clairvoyant detector has a very simple form, and can be informally interpreted as the mean-subtracted matched filter minus some fraction of the Mahalanobis anomaly function, where the fraction depends on the target abundance a.

As before, we can obtain the LMP detector by taking $a = 0$ in Eq. (25).

The $a \to 1$ limit is also potentially of interest for this problem; that is obtained from Eq. (25) simply by setting $a = 1$.

4.2 Replacement model on multivariate t background

For the multivariate t distribution given in Eq. (10), the likelihood ratio given in Eq. (19) becomes

$$\mathcal{L}(a,\mathbf{x}) = (1-a)^{-d}\left(\frac{(\nu-2) + \dfrac{\mathcal{A}(\mathbf{x}-a(\mathbf{t}-\boldsymbol{\mu}))}{(1-a)^2}}{(\nu-2)+\mathcal{A}(\mathbf{x})}\right)^{-(\nu+d)/2} \tag{26}$$

and from this, a clairvoyant can be produced:

$$\mathcal{D}(a,\mathbf{x}) = \frac{(\nu-1)(1-a)^2}{2a}\left(1 - \left[(1-a)^d\mathcal{L}(a,\mathbf{x})\right]^{-2/(\nu+d)}\right) \tag{27}$$

$$= \mathcal{F}_\nu^2(\mathbf{x})\left(\mathcal{D}_o(a,\mathbf{x}) - \tfrac{1}{2}a\mathcal{A}(\mathbf{t})\right) \tag{28}$$

with $\mathcal{D}_o(a,\mathbf{x})$ given in Eq. (25).

To compute the GLRT, we need to find $\widehat{a}$ that maximizes $\mathcal{L}(a,\mathbf{x})$; to do this, we must take the derivative with respect to a, and then find the value of a for which that derivative is zero. This generally leads to more complicated formulations, and although closed-form solutions are not guaranteed in general, they have been found for the replacement model with Gaussian clutter (the finite target matched filter (FTMF) of Schaum and Stocker[15]) and with multivariate t-distributed background clutter.[16]

4.3 Counting sigmas with the replacement model

For the additive model, the notion of a characteristic target strength (defined in Eq. (9)) was relatively straightforward to define. For the replacement model, the notion is a little more complicated.

In the absence of target, the detector in Eq. (25) has mean given by $-(1-\frac{1}{2}a)d$ and variance given by a more complicated formula: $\mathcal{A}(\mathbf{t}) + 2d(1-\frac{1}{2}a)^2$, which is derived using the fact that $\mathcal{A}(\mathbf{x})$ is chi-squared distributed with d degrees of freedom. In the presence of a target of abundance a, the mean of the distribution becomes

$-(1-\frac{1}{2}a)\left[(1-a)^2 d + a^2\mathcal{A}(\mathbf{t})\right] + a\mathcal{A}(\mathbf{t})$. Equating square of mean difference and variance leads to characteristic target strength a_o. The expression will be complicated in general, but for small a ($a \ll 1$), we get

$$a_o = 1/\sqrt{2d + \mathcal{A}(\mathbf{t})}. \tag{29}$$

Since the target abundance is bounded between zero and one, the choice for our n-sigma *veritas* detector would be given by Eq. (25) or Eq. (28) with $a = \min(1, na_o)$.

5. ABSORPTIVE PLUME MODEL

For an absorptive plume, we have from Beer's Law that the radiance observed at some wavelength λ is given by $x_\lambda = z_\lambda \exp(-at_\lambda)$, where z_λ is the radiance that would observed in the absence of plume, t_λ is the absorption coefficient of the plume gas, and a is the plume strength (which depends on the concentration of the gas and the thickness of the plume). For a sensor with d wavelengths, we can express this in vector form, with d-dimensional vectors $\mathbf{x}$ and $\mathbf{z}$, whose components are x_λ and z_λ, respectively:

$$\mathbf{x} = \boldsymbol{\xi}(\mathbf{z}) = \exp(-aT)\mathbf{z}, \tag{30}$$

where T is a diagonal matrix whose diagonal elements are the absorption coefficients t_λ.

It bears remarking that a common approximation here is to take $\widehat{\boldsymbol{\xi}}(\mathbf{x}) = \mathbf{z} - aT\langle\mathbf{z}\rangle = \mathbf{z} - aT\boldsymbol{\mu}$, which corresponds to the additive model with $\mathbf{t} = -T\boldsymbol{\mu}$. It is this approximation that enables the use of a standard[17–19] or EC-based[14] matched filter for gas detection. But one can obtain closed-form solutions without making this approximation.[20] To do this, we use $|d\boldsymbol{\xi}/d\mathbf{x}| = \exp(-a\tau)$ where $\tau = \mathrm{Trace}(T)$, and substitute into the generic expression in Eq. (2), to obtain

$$\mathcal{L}(a,\mathbf{x}) = \frac{P_{\text{target}}(\mathbf{x})}{P_{\text{bkg}}(\mathbf{x})} = \frac{P_{\text{bkg}}(\exp(aT)\mathbf{x})\,\exp(a\tau)}{P_{\text{bkg}}(\mathbf{x})}. \tag{31}$$

5.1 Absorptive plume model on Gaussian background

For a Gaussian background, Eq. (31) leads to a clairvoyant detector of the form

$$\begin{aligned}\mathcal{D}(a,\mathbf{x}) &= 2\log\mathcal{L}(a,\mathbf{x}) - 2a\tau = 2\log P_{\text{bkg}}(\exp(aT)\mathbf{x}) - 2\log P_{\text{bkg}}(\mathbf{x})\\ &= (\exp(aT)\mathbf{x}-\boldsymbol{\mu})'R^{-1}(\exp(aT)\mathbf{x}-\boldsymbol{\mu}) + (\mathbf{x}-\boldsymbol{\mu})'R^{-1}(\mathbf{x}-\boldsymbol{\mu})\\ &= \mathcal{A}(\mathbf{x}) - \mathcal{A}(\exp(aT)\mathbf{x})\end{aligned} \tag{32}$$

5.2 Absorptive plume model on multivariate t background

For a multivariate t-distributed background, the expression is a little more complicated. Here, Eq. (31) becomes

$$\mathcal{L}(a,\mathbf{x}) = \frac{P_{\text{target}}(\mathbf{x})}{P_{\text{bkg}}(\mathbf{x})} = \frac{P_{\text{bkg}}(\exp(aT)\mathbf{x})\exp(a\tau)}{P_{\text{bkg}}(\mathbf{x})} = \frac{\left[\nu-2+\mathcal{A}(\exp(aT)\mathbf{x})\right]^{-(\nu+d)/2}}{\left[\nu-2+\mathcal{A}(\mathbf{x})\right]^{-(\nu+d)/2}} \times \exp(a\tau). \tag{33}$$

Thus,

$$\mathcal{L}(a,\mathbf{x})^{-2/(\nu+d)} = \frac{(\nu-2)+\mathcal{A}(\exp(aT)\mathbf{x})}{(\nu-2)+\mathcal{A}(\mathbf{x})} \times \exp\left(\frac{-2a\tau}{\nu+d}\right) \tag{34}$$

$$\begin{aligned}&= \left(1 + \frac{\mathcal{A}(\exp(aT)\mathbf{x}) - \mathcal{A}(\mathbf{x})}{(\nu-2)+\mathcal{A}(\mathbf{x})}\right) \times \exp\left(\frac{-2a\tau}{\nu+d}\right)\\ &= \left(1 + \frac{\mathcal{F}_\nu(\mathbf{x})^2\left[\mathcal{A}(\exp(aT)\mathbf{x}) - \mathcal{A}(\mathbf{x})\right]}{\nu-1}\right) \times \exp\left(\frac{-2a\tau}{\nu+d}\right).\end{aligned} \tag{35}$$

So then

$$\mathcal{D}(a,\mathbf{x}) = (\nu-1)\left(1 - \exp(\frac{2a\tau}{\nu+d})\mathcal{L}(a,\mathbf{x})^{-2/(\nu+d)}\right) \tag{36}$$

as the monotonic transform, and then use Eq. (35) for $\mathcal{L}(\mathbf{x})^{-2/(\nu+d)}$ to obtain

$$\begin{aligned}\mathcal{D}(a,\mathbf{x}) &= (\nu-1)\left(1-\exp(\frac{2a\tau}{\nu+d})\left\{\left(1+\frac{\mathcal{F}_\nu(\mathbf{x})^2\left[\mathcal{A}(\exp(aT)\mathbf{x})-\mathcal{A}(\mathbf{x})\right]}{\nu-1}\right)\times\exp(\frac{-2a\tau}{\nu+d})\right\}\right)\\ &= \mathcal{F}_\nu^2(\mathbf{x})\big[\mathcal{A}(\mathbf{x})-\mathcal{A}(\exp(aT)\mathbf{x})\big]. \qquad (37)\end{aligned}$$

with $\mathcal{F}_\nu$ defined in Eq. (14).

5.3 Counting sigmas with the absorptive plume model

For weak plumes, the additive approximation leads to a characteristic strength of $a_o = 1/\sqrt{\boldsymbol{\mu}'T'R^{-1}T\boldsymbol{\mu}}$.

6. SIMULATIONS AND EVALUATION

For additive targets on a multispectral t-distributed background, numerical experiments were performed to compare their performance as a function of target strength. The detectors that were evaluated are listed in Table 1. For the *veritas* detector we optimzed at $n = a/a_o = 4$ sigmas.

A moderate dimension of $d = 20$, corresponding to twenty multispectral channels, was chosen, and a moderate fatness of $\nu = 10$ was chosen. A sample of $N = 10^8$ points were drawn from the t distribution with zero mean and unit covariance matrix. For each target strength a, a corresponding set of $N = 10^8$ points were generated using the additive model: $\mathbf{x} = \mathbf{z} + a\mathbf{t}$.

To these data, the detectors were applied, and ROC curves were computed in order to evaluate the performance of the the different detectors. Three statistics were computed from the ROC curves: Area under the ROC curve (AUC), Detection rate (DR) at a threshold corresponding to a false alarm rate of 0.0001, and False alarm rate (FAR) associated with a detection rate of ninety percent.

Results are shown in Figs. 1-4. The statistics in these figures are plotted so that better performance corresponds to lower values. In all cases, we observe that the clairvoyant detector provides (just as theory predicts) a bound on the best possible performance of a detector. We also observe (again as expected) that the *veritas* detector is optimal for target strengths $a/a_o = 4$. We observe, however, that this optimality does not extend over the entire range of target strengths.

Indeed, by using different metrics over a range of target values, we see quite an interplay among the various detectors. Observe, for example, the behavior of the AMF detector in Fig. 2; for weaker targets, it is one of the worst performers (only RX performs worse) but for very strong targets it is the best detector. Furthermore, this behavior is really only observed in the DR@FAR=1e-4 (detection rate corresponding to false alarm rate of 0.0001) metric; the AUC and false alarm metric do not exhibit the same kind of variation for AMF. Indeed, one should be careful about drawing general conclusions about comparative performance of detectors from a limited number of empirical comparisons.

One result that is hinted at in all of the plots, but is most evident in Fig. 4, is that the *veritas* solution becomes not only "less optimal" as $a/a_o \gg n$ (where n is the number of sigmas for which the detector has been optimized), but it can also become objectively worse. In some sense, the *veritas* detector parallels the LMP detector: both are optimal at the target strength for which they are optimized, but the informal notion that they will still be "pretty good" for even stronger targets does not quite hold up.

Although the RX detector, pretty much as expected, performs poorly compared to the other detectors, we do see its relative perfomance improve considerably as target strengths are increased. This is most evident in Fig. 4, where the RX detector beats the *veritas* detector at around fifteen sigmas.

7. CONCLUSIONS

In this paper, closed-form expressions were derived for the *veritas* detectors in six distinct scenarios: for three different target models on two different background distributions. Because the *veritas* detector is based directly on the likelihood ratio, without the need to integrate (as in Bayesian methods) or optimize (as in GLRT methods), it is possible to obtain closed-form solutions even for relatively complicated target detection scenarios and/or background distributions.

While modern computers make the need for closed-form solutions less important than they may be been in the past, they are still useful. Low computation may be all that is available in an on-orbit situation.[21,22] And in terms of explainability and understanding, a "formula" is almost always better than a subroutine.

For background distributions that are entirely unknown, and target models that may be very complicated (as long as the model is well enough specified that the effect of targets on background can be simulated), then the *veritas* idea leads directly to the matched-pair machine learning approach.[23–25]

Although *veritas* detectors are optimal at the target strengths that are deemed of most interest, they are not not optimal in the sense of uniformly most powerful (UMP). Indeed, Fig. 4 shows a case in which the *veritas* detector's non-optimal behavior for strong targets is particularly noteable.

A potential extension is to use a small number of discrete target strengths, and to build a prior from a sum of delta functions. This Bayesian approach to target detection is guaranteed to produce admissible detectors, and since the prior is discrete, the dreaded integration that is usually required of Bayesian detectors is replaced with the sum of a few terms, and the result is a slightly more complicated but still closed-form solution. How best to specify what these target strengths should be, and what weights should be given to each: that is an optimization problem that is beyond (but only just beyond) the scope of this paper.

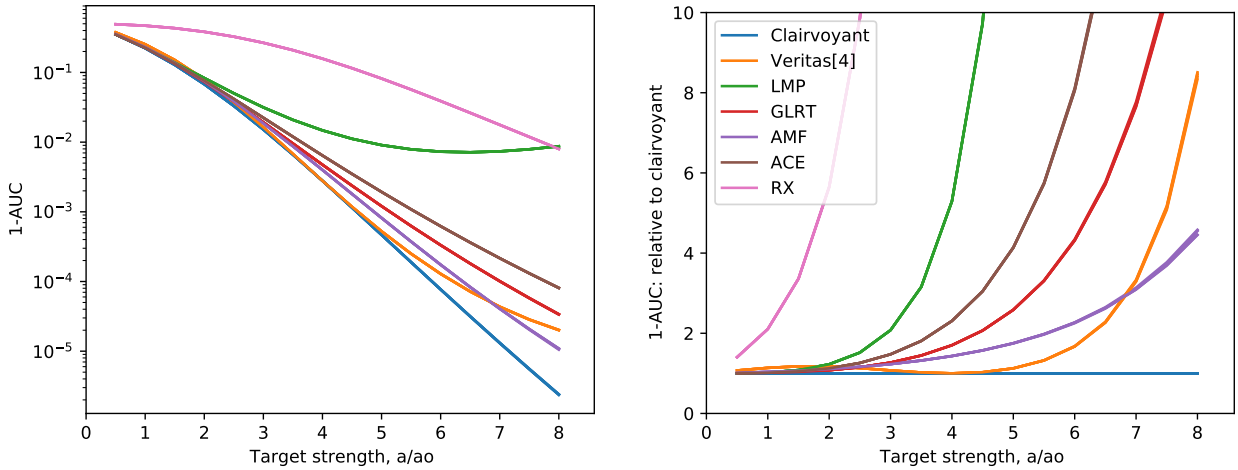
8. ACKNOWLEDGMENTS

The author acknowledges support from the NASA-sponsored project on "Compact high-resolution trace-gas hyperspectral imagers, with agile on-board processing, for trace gas monitoring" (N7-INVEST17-0010), led by Steven P. Love at Los Alamos National Laboratory.

REFERENCES

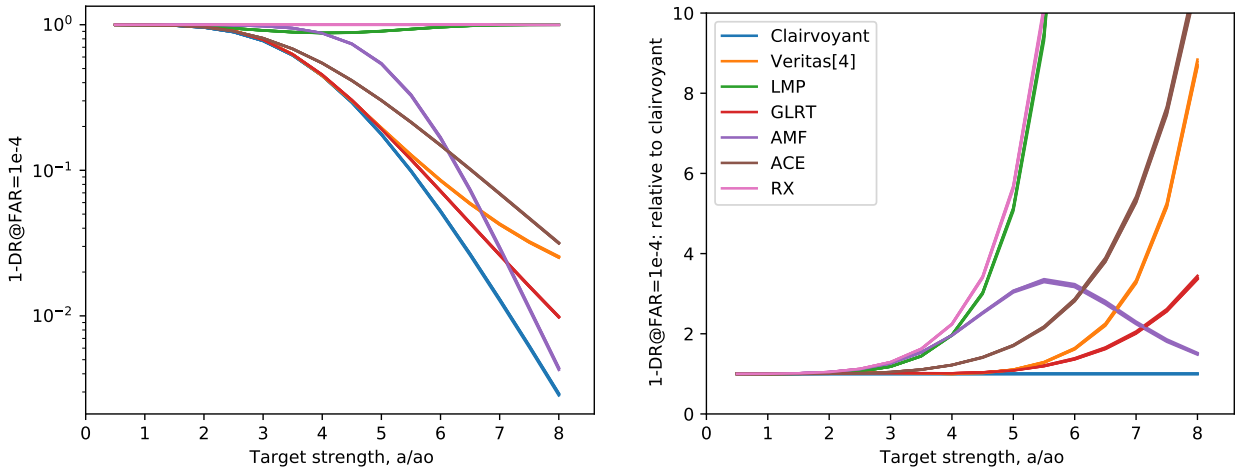
1. D. J. Helfand, E. C. Moran, B. Oppenheimer, and R. L. White, "Veritas in duo sigma – look at all those sources!," *Bulletin of the American Astronomical Society* **25**, p. 1298, 1993.
2. I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. Aerospace and Electronic Systems* **10**, pp. 853–863, 1974.
3. E. J. Kelly, "Performance of an adaptive detection algorithm: rejection of unwanted signals," *IEEE Trans. Aerospace and Electronic Systems* **25**, pp. 122–133, 1989.
4. F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerospace and Electronic Systems* **28**, pp. 208–216, 1992.
5. L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Trans. Signal Processing* **42**, pp. 2146–2156, 1994.
6. S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, vol. II, Prentice Hall, New Jersey, 1998.
7. J. Theiler, "Confusion and clairvoyance: some remarks on the composite hypothesis testing problem," *Proc. SPIE* **8390**, p. 839003, 2012.
8. J. Chen, "Penalized likelihood-ratio test for finite mixture models with multinomial observations," *Canadian Journal of Statistics* **26**, pp. 583–599, 1998.
9. A. Vexler, C. Wu, and K. F. Yu, "Optimal hypothesis testing: from semi to fully Bayes factors," *Metrika* **71**, pp. 125–138, 2010.
10. A. Schaum, "Continuum fusion: a theory of inference, with applications to hyperspectral detection," *Optics Express* **18**, pp. 8171–8181, 2010.

11. A. Schaum, "Clairvoyant fusion: a new methodology for designing robust detection algorithms," *Proc. SPIE* **10004**, p. 100040C, 2016.
12. I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoustics, Speech, and Signal Processing* **38**, pp. 1760–1770, 1990.
13. D. Manolakis, D. Marden, J. Kerekes, and G. Shaw, "On the statistics of hyperspectral imaging data," *Proc. SPIE* **4381**, pp. 308–316, 2001.
14. J. Theiler and B. R. Foy, "EC-GLRT: Detecting weak plumes in non-Gaussian hyperspectral clutter using an elliptically-contoured generalized likelihood ratio test," in *Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, p. I:221, 2008.
15. A. Schaum and A. Stocker, "Spectrally selective target detection," in *Proc. ISSSR (International Symposium on Spectral Sensing Research)*, p. 23, 1997.
16. J. Theiler, B. Zimmer, and A. Ziemann, "Closed-form detector for solid sub-pixel targets in multivariate t-distributed background clutter," in *Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 2773–2776, 2018.
17. A. Hayden, E. Niple, and B. Boyce, "Determination of trace-gas amounts in plumes by the use of orthogonal digital filtering of thermal-emission spectra," *Applied Optics* **35**, pp. 2802–2809, 1996.
18. D. G. Manolakis and F. M. D'Amico, "A taxonomy of algorithms for chemical vapor detection with hyperspectral imaging spectroscopy," *Proc. SPIE* **5795**, pp. 125–133, 2005.
19. D. R. Thompson, I. Leifer, H. Bovensmann, M. Eastwood, M. Fladeland, C. Frankenberg, K. Gerilowski, R. O. Green, S. Kratwurst, T. Krings, B. Luna, and A. K. Thorpe, "Real-time remote detection and measurement for airborne imaging spectroscopy: a case study with methane," *Atmospheric Measurement Techniques* **8**, pp. 4383–4397, 2015.
20. J. Theiler and A. Schaum, "Some closed-form expressions for absorptive plume detection," in *Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 1786–1789, 2020.
21. S. P. Love, L. A. Ott, J. Theiler, B. R. Foy, C. L. Safi, M. E. Dale, C. G. Peterson, A. A. Guthrie, N. A. Dallman, K. G. Boyd, P. S. Stein, J. A. Wren, M. C. Proicou, and M. K. Dubey, "High-resolution hyperspectral imaging of dilute gases from cubesat platforms," in *American Geophysical Union, Fall Meeting*, 2018. Abstract #A41K-3107.
22. J. Theiler, B. R. Foy, C. Safi, and S. P. Love, "Onboard cubesat data processing for hyperspectral detection of chemical plumes," *Proc. SPIE* **10644**, p. 1064405, 2018.
23. J. Theiler, "Matched-pair machine learning," *Technometrics* **55**, pp. 536–547, 2013.
24. J. Theiler, "Transductive and matched-pair machine learning for difficult target detection problems," *Proc. SPIE* **9088**, p. 90880E, 2014.
25. A. Ziemann, M. Kucer, and J. Theiler, "A machine learning approach to hyperspectral detection of solid targets," *Proc. SPIE* **10644**, p. 1064404, 2018.



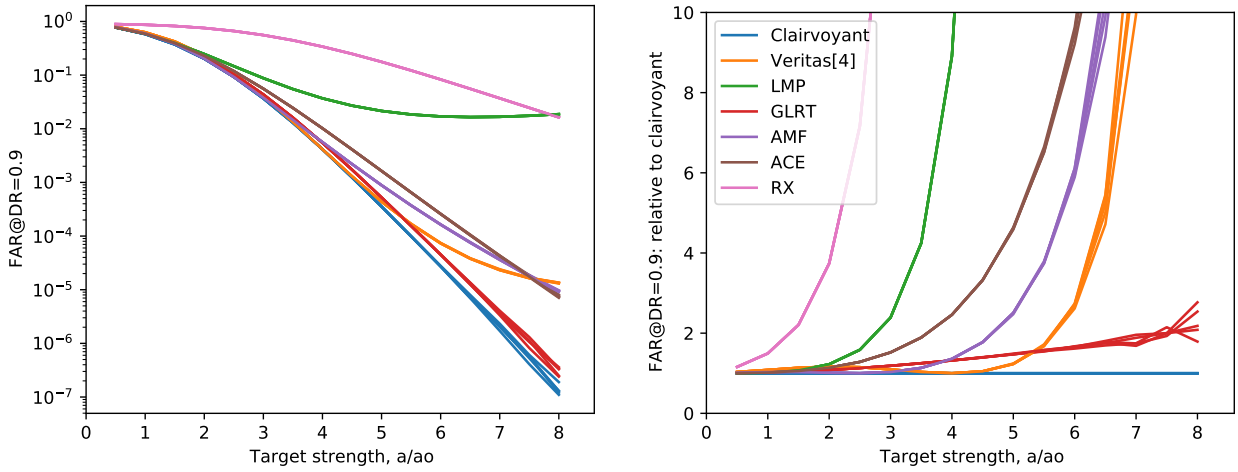
(a) AUC (b) AUC relative to clairvoyant

Figure 1. Performance for seven detectors is plotted as a function of target strength using the area under the ROC curve (AUC) as a performance measure. The left panel is the AUC statistic directly (actually it is 1-AUC so that AUC values near 1 can be seen in a log plot), and the right panel is the ratio of the AUC for each detector divided by the AUC for the clairvoyant detector. We cannot actually use the clairvoyant detector when the target strength is unknown, but it provides a useful lower bound on the 1-AUC performance. The *veritas* detector is best when $a/a_o = 4$; that is by design. But it is also best over the range from about $2.5 \leq a/a_o \leq 7.0$. One might argue that targets with $a/a_o < 2.5$ are effectively un-detectable anyway, so performance in that regime is not important; similarly, the targets with $a/a_o > 7$ are easily detectable and all of the detectors (except RX and LMP) are effective in detecting those targets.



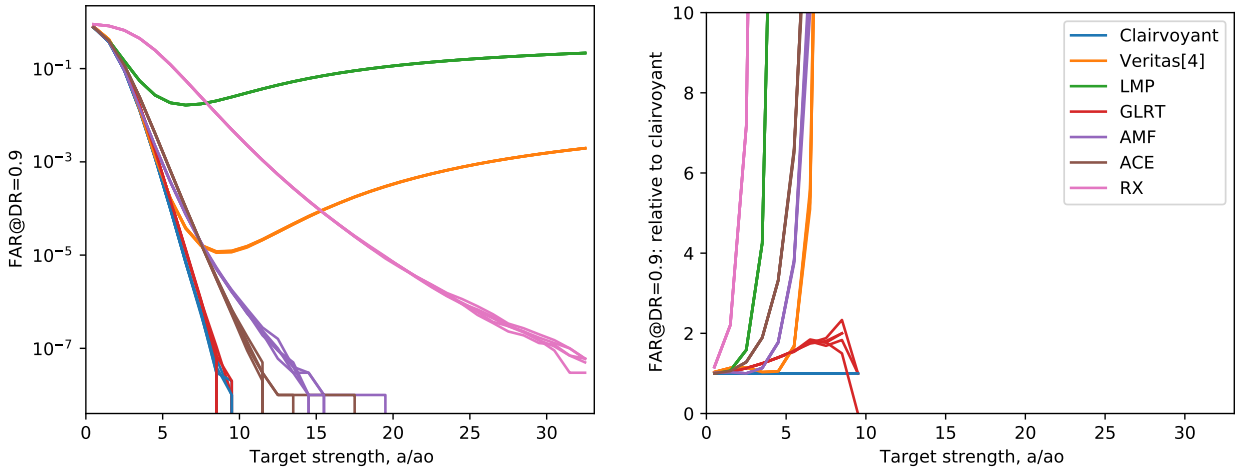
(a) DR (b) DR relative to clairvoyant

Figure 2. Similar to Fig. 1 but instead of AUC, we compare using the detection rate at the threhold for which false alarm rate is 0.0001. We actually plot 1-DR, so smaller values are better. Although the *veritas* detector is, as theory predicts, optimal at $a/a_o = 4$, we see that on the whole, the GLRT detector is doing a better job. *veritas* outperforms ACE essentially over the whole range of target strengths, and is better than AMF for $a/a_o < 6$. Interestingly, the AMF behavior changes course for larger values of a and for $a/a_o > 7$, it is (by this metric) the best detector.



(a) FAR (b) FAR relative to clairvoyant

Figure 3. Similar to Fig. 1 and Fig. 2, but this time using as a performance metric the false alarm rate at the threhold for which detection rate is 0.9. Again, *veritas* is optimal at $a/a_o = 4$, but already by $a/a_o > 5$, the GLRT is better. And by $a/a_o > 8$, we see that ACE and AMF have caught up to *veritas*.



(a) FAR (b) FAR

Figure 4. Same as Fig. 3 but over a much wider range of target strengths. We see that as $a/a_o > 8$, the *veritas* performance literally turns around. Not only is the *veritas* detector outperformed by ACE, AMF, and GLRT; but the slope of the false alarm rate curve actually becomes positive. The stronger the targets, the harder they are to detect! Meanwhile, as the targets approach 30+ sigmas, the RX can detect targets with very low false alarm rates. (Note that with only $N = 10^8$ points in the simulation, we cannot estimate false alarm rates below 10^{-8}; thus the clairvoyant detection performance cannot be used as a denominator for $a/a_o > 8$.)