SOME CLOSED-FORM EXPRESSIONS FOR ABSORPTIVE PLUME DETECTION

James Theiler[*]

Los Alamos National Laboratory
Los Alamos, NM 87545, USA
jt@lanl.gov

Alan Schaum[†]

U. S. Naval Research Laboratory
Washington DC 20375, USA
alan.schaum@nrl.navy.mil

ABSTRACT

For additive signals on Gaussian clutter, the optimal detector is a linear matched filter that is adapted to the known signal and the covariance of the background. This adaptive matched filter (AMF) is widely used for gas-phase plume detection, even though the effect of the plume on the background is not strictly additive. Here, a derivation of the matched filter for a strictly absorptive plume produces, even in the weak plume limit, a quadratic filter. Assuming a Gaussian background, we derive two expressions, one based on the locally most powerful (LMP) detector which corresponds to the weak plume limit, and one based on the generalized likelihood ratio test (GLRT). Numerical experiments indicate that, as long as the plume is strong enough to be detected at all, both the GLRT and the linear matched filter outperform the LMP detector.

1. INTRODUCTION

There is considerable interest in remote detection of gas-phase plumes, with particular interest in NO_2, which is often seen as a proxy for more general pollution as well as greenhouse gas emission, and SO_2, which is often diagnostic of volcanic activity. The physics of plume absorption is simple, but it is not linear, and our aim here is to develop closed-form expressions for detecting plumes in hyperspectral imagery.

This problem also provides an exercise in composite hypothesis testing. For such problems, a direct likelihood ratio test cannot be employed, because there is a nuisance parameter (in this case, plume concentration) whose value is not *a priori* known. The usual approach in this situation employs the Generalized Likelihood Ratio Test (GLRT), which estimates this unknown parameter. An alternative is to consider the weak plume limit with the Locally Most Powerful (LMP). We note that more general Bayesian methods [1] (of which LMP is a special case) or Clairvoyant Fusion [2–4] (for which GLRT is a special case) can also be considered.

[*]JT was partially supported by a 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.

[†]AS acknowledges support from the Office of Naval Research

1.1. Absorptive plume

For an absorptive plume, we have from Beer's Law that the radiance observed at some wavelength λ is given by $x_\lambda = z_\lambda \exp(-\epsilon t_\lambda)$, 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 ϵ is the plume strength. 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:

$$\mathbf{x} = \exp(-\epsilon T)\mathbf{z}, \tag{1}$$

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

1.2. Linear matched filter

The classic adaptive matched filter (AMF) [5] was originally applied to radar signal detection, but is widely used for plume detection. To adapt the AMF to the absorptive plume problem, two approximations must be made. The first assumption is that the plume is weak, so that the exponential is approximately linear; that is: $\exp(-\epsilon T)\mathbf{z} \approx \mathbf{z} - \epsilon T\mathbf{z}$.

The second approximation treats the additive term $-\epsilon T\mathbf{z}$ as if it were a constant. A naive yet popular approximation takes $T\mathbf{z} \propto \mathbf{t}$; this is based on the argument that the background $\mathbf{z}$ varies much more slowly with wavelength than does the gas spectrum $\mathbf{t}$, and so the approximation is that the background is flat. A better approximation treats $T\mathbf{z} \approx T\boldsymbol{\mu}$, where $\boldsymbol{\mu} = \langle\, \mathbf{z} \,\rangle$ is the mean background over the image. In this absorptive plume context, then, the AMF becomes

$$\mathcal{D}_{\text{AMF-}T\mu}(\mathbf{x}) = -(T\boldsymbol{\mu})' R^{-1}(\mathbf{x} - \boldsymbol{\mu}). \tag{2}$$

where $R = \langle\, (\mathbf{z}-\boldsymbol{\mu})(\mathbf{z}-\boldsymbol{\mu})' \,\rangle$ is the covariance matrix of the background variability. A variant of this expression

$$\widehat{\epsilon}_{\text{AMF}} = \frac{-(T\boldsymbol{\mu})' R^{-1}(\mathbf{x} - \boldsymbol{\mu})}{(T\boldsymbol{\mu})' R^{-1} T\boldsymbol{\mu}} \tag{3}$$

has the advantage that it provides a direct estimator of plume strength; since the denominator is a constant, its performance as a detector is identical to $\mathcal{D}_{\text{AMF-}T\mu}$ in Eq. (2).

2. CLAIRVOYANT DETECTOR

Let us begin with the assumption that the strength of the plume is known (even as its presence or absence remains an open question). In practice, this is usually not the case, but the detector we obtain (called the *clairvoyant* detector [6]) provides a useful reference point. With ϵ known, the problem reduces to a binary hypothesis test, and the optimal solution is likelihood ratio.

$$\mathcal{L}(\epsilon, \mathbf{x}) = \frac{P_{\text{plume}}(\mathbf{x})}{P_{\text{bkg}}(\mathbf{x})} \tag{4}$$

Since Eq. (1) expresses the effect of plume on a background pixel, we can use the usual formula for change of variables in probability distributions:

$$P_{\text{plume}}(\mathbf{x}) = P_{\text{bkg}}(\mathbf{z}) \left| \frac{d\mathbf{z}}{d\mathbf{x}} \right| = P_{\text{bkg}}(\exp(\epsilon T)\mathbf{x}) \left| \exp(\epsilon T) \right| \tag{5}$$

where $|\cdot|$ indicates the determinant. Note that

$$|\exp(\epsilon T)| = \prod_\lambda \exp(\epsilon t_\lambda) = \exp(\epsilon \sum_\lambda t_\lambda) = \exp(\epsilon\tau) \tag{6}$$

where $\tau = \sum_\lambda t_\lambda = \text{Trace}(T)$. So the likelihood ratio becomes

$$\mathcal{L}(\epsilon, \mathbf{x}) = \frac{P_{\text{plume}}(\mathbf{x})}{P_{\text{bkg}}(\mathbf{x})} = \frac{P_{\text{bkg}}(\exp(\epsilon T)\mathbf{x})\ \exp(\epsilon\tau)}{P_{\text{bkg}}(\mathbf{x})} \tag{7}$$

For a Gaussian background, we have

$$P_{\text{bkg}}(\mathbf{x}) = (2\pi)^{-\frac{d}{2}} |R|^{-\frac{1}{2}} \exp\left[-\tfrac{1}{2}(\mathbf{x}-\boldsymbol{\mu})' R^{-1} (\mathbf{x}-\boldsymbol{\mu})\right]. \tag{8}$$

Incorporating this expression in Eq. (7), and taking the logarithm, we obtain the clairvoyant detector:

$$\begin{aligned} \mathcal{D}(\epsilon, \mathbf{x}) &= \log \mathcal{L}(\epsilon, \mathbf{x}) \\ &= \log P_{\text{bkg}}(\exp(\epsilon T)\mathbf{x}) + \epsilon\tau - \log P_{\text{bkg}}(\mathbf{x}) \\ &= -\tfrac{1}{2}(\exp(\epsilon T)\mathbf{x} - \boldsymbol{\mu})' R^{-1} (\exp(\epsilon T)\mathbf{x} - \boldsymbol{\mu}) \\ &\quad + \epsilon\tau + \tfrac{1}{2}(\mathbf{x}-\boldsymbol{\mu})' R^{-1} (\mathbf{x}-\boldsymbol{\mu}) \end{aligned} \tag{9}$$

3. QUADRATIC MATCHED FILTER

Because we care about *weak* plumes, we will derive a locally most powerful (LMP) detector [6] that is optimal for $\epsilon \to 0$. In this small ϵ regime, we can write

$$\exp(\epsilon T)\mathbf{x} = \mathbf{x} + \epsilon T\mathbf{x} + O(\epsilon^2) \tag{10}$$

and substituting this into Eq. (9) leads to the "quadratic matched filter"

$$\mathcal{D}_{\text{QMF}}(\mathbf{x}) = \lim_{\epsilon\to 0} \frac{\mathcal{D}(\epsilon, \mathbf{x})}{\epsilon} = -(T\mathbf{x})' R^{-1} (\mathbf{x}-\boldsymbol{\mu}) + \tau. \tag{11}$$

Comparing this with Eq. (2), we see that this looks like a linear matched filter, but the match is to $T\mathbf{x}$ instead of $T\boldsymbol{\mu}$. The additive constant τ does not affect performance at all (its value can be subsumed into the threshold used for detection), but we have chosen to include it in the definition.

4. GLRT

The Generalized Likelihood Ratio Test (GLRT) formulation recognizes the dependence of the detector on plume strength ϵ. The above LMP formulation considered the small ϵ limit; by contrast, the GLRT formulation takes two steps: first an estimated plume strength $\widehat{\epsilon}$ is computed and then the likelihood is evaluated at that estimate.

The maximum likelihood estimate for ϵ is the value that maximizes $P_{\text{plume}}(\mathbf{x})$; that is:

$$\widehat{\epsilon} = \text{argmax}_\epsilon\ P_{\text{bkg}}\left(\exp(\epsilon T)\mathbf{x}\right) \exp(\epsilon\tau) \tag{12}$$

$$= \text{argmin}_\epsilon\ \left(\exp(\epsilon T)\mathbf{x} - \boldsymbol{\mu}\right)' R^{-1} \left(\exp(\epsilon T)\mathbf{x} - \boldsymbol{\mu}\right) - 2\epsilon\tau \tag{13}$$

Eq. (13) is a transcendental equation, and an exact closed-form solution is beyond the humble algebraic skills of these authors. However, since we are interested in small ϵ, we can approximate the solution using a Taylor series expansion up to quadratic terms in ϵ. In this formulation, Eq. (13) becomes:

$$\begin{aligned} \widehat{\epsilon} = \text{argmin}_\epsilon\ \big[&(\mathbf{x}-\boldsymbol{\mu})' R^{-1} (\mathbf{x}-\boldsymbol{\mu}) \\ &+ 2\epsilon (T\mathbf{x})' R^{-1} (\mathbf{x}-\boldsymbol{\mu}) - 2\epsilon\tau \\ &+ \epsilon^2 \Big((T\mathbf{x})' R^{-1} T\mathbf{x} + (T\mathbf{x})' T R^{-1} (\mathbf{x}-\boldsymbol{\mu}) \Big) \\ &+ O(\epsilon^3) \big] \end{aligned} \tag{14}$$

And by neglecting the $O(\epsilon^3)$ term, we can solve to obtain

$$\widehat{\epsilon} = \frac{-(T\mathbf{x})' R^{-1} (\mathbf{x}-\boldsymbol{\mu}) + \tau}{(T\mathbf{x})' R^{-1} T\mathbf{x} + (T\mathbf{x})' T R^{-1} (\mathbf{x}-\boldsymbol{\mu})}. \tag{15}$$

As written, this quantity can be positive or negative. We may choose to restrict $\widehat{\epsilon} \geq 0$ for purely absorptive plumes. In that case, negative quantities just get reset to zero: $\widehat{\epsilon} \leftarrow \max[0, \widehat{\epsilon}]$.

It has been suggested [7] that an *ad hoc* albedo correction be applied to the AMF estimator of plume strength; here,

$$\widehat{\epsilon}_{\text{albedo-corrected}} = \frac{1}{r}\, \widehat{\epsilon}_{\text{AMF}} = \frac{-(T\boldsymbol{\mu})' R^{-1} (\mathbf{x}-\boldsymbol{\mu})}{r (T\boldsymbol{\mu})' R^{-1} T\boldsymbol{\mu}} \tag{16}$$

where the scalar factor r is given by $r = \mathbf{x}'\boldsymbol{\mu}/\boldsymbol{\mu}'\boldsymbol{\mu}$. We observe that Eq. (15) already includes a kind of albedo correction (with equal powers of $\mathbf{x}$ in the numerator and denominator, it is less "sensitive" to the magnitude of $\mathbf{x}$), even though it was derived without explicitly imposing this property.

The estimate of plume strength and the likelihood of plume presence are not the same thing, though it is reasonable to presume that they are relatively well correlated, and they in fact agree for the linear matched filter. So one *can* use $\widehat{\epsilon}$ as a plume detector. But the GLRT detector is obtained by substituting $\widehat{\epsilon}$ into the clairvoyant formula in Eq. (9). This can be done directly (and the result is a closed-form expression), but since $\widehat{\epsilon}$ was estimated by neglecting $O(\epsilon^3)$ terms, we make that same approximation here, and obtain

$$\mathcal{D}_{\text{GLRT}}(\mathbf{x}) = \frac{-(T\mathbf{x})' R^{-1}(\mathbf{x}-\boldsymbol{\mu}) + \tau}{\sqrt{(T\mathbf{x})' R^{-1} T\mathbf{x} + (T\mathbf{x})' T R^{-1}(\mathbf{x}-\boldsymbol{\mu})}}. \quad (17)$$

5. MATCHED PAIR EVALUATION SCHEME

We will measure the quality of detection algorithms by implanting a plume into a hyperspectral image. But we will do this in a formalized way that makes two copies of the hyperspectral data [8]. The first copy is untouched, but the second copy has plume added to every pixel. Thus, if $\mathbf{z}$ is the pixel in the first (plume-free) copy, then the corresponding pixel in the on-plume copy is $\mathbf{x} = \exp(-\epsilon T)\mathbf{z}$. Mean and covariance will be estimated just from the off-plume data (in other words, we are neglecting contamination effects, arguing that in an operational scenario, the on-plume pixels will be rare). Each detector is applied to both on-plume and off-plume pixels and from these a ROC curve can be derived. Three statistics of interest to us are: FAR@DR=0.5, the false alarm rate at threshold with detection rate of 0.5; DR@FAR=0.5, the detection rate at threshold with false alarm rate of 0.5; and AUC, the area under the ROC curve.

The FAR@DR=0.5 is more appropriate for most detection scenarios (where low false alarm rates are crucial). The DR@FAR=0.5 provides a kind of counterpoint that might be relevant for the odd scenario in which detections are crucial and false alarms can be tolerated. The AUC is widely employed, and provides a kind of compromise between the first two. All of these statistics are scalar values between 0 and 1: for false alarms, smaller values are better; while for detection rates and AUC, larger values are better.

Table 2(a) corresponds to a weak but detectable[1] NO_2 plume implanted into a scene from the Ozone Monitoring Instrument (OMI) [9], while Table 2(b) implants an SO_2 plume in a different OMI image [10]. For the additive detectors, we find AMF-$T\boldsymbol{\mu}$ is more effective than AMF-$\mathbf{t}$ for all the statistical measures. For minimizing false alarm rate, AMF-$T\boldsymbol{\mu}$ is best in both cases, even (slightly) better than the Clairvoyant detector for SO_2. More consistent with the theory, the GLRT is the best non-Clairvoyant detector in terms of AUC. And it is interesting that the plume strength estimator ϵ, which was never designed as a plume detector *per se*, gets the highest non-Clairvoyant scores in terms of high detection rate at FAR=0.5.

Table 3 provides results of an experiment similar to that of Table 2, but the plumes are implanted on Gaussian data (with the same mean and covariance as the OMI data). Here the results favor GLRT for both SO_2 and NO_2 plumes, based on both the FAR@DR=0.5 and AUC statistics. And in fact the GLRT is very nearly the best for DR@FAR=0.5 as well, with the plume strength estimator $\widehat{\epsilon}$ only slightly surpassing it. For both plumes, and all three statistics, we see that the Clairvoyant detector is (as theoretically predicted) optimal.

Since the plumes are artificially implanted into the scenes in exact accord with the Beer's Law formula in Eq. (1), the only reason for deviation of theory and practice in Table 2 is the distribution of the background data. This speaks in favor of more sophisticated models for background distribution [11], based for example on parametric distributions (such as the multivariate t-distribution [12]) or nonparametric [13] or even machine learning approaches [14, 15].

6. CONCLUSION

It is often, albeit informally, asserted that a linear filter works for plume detection because the exponential in Beer's law becomes linear in the weak plume limit. But a more careful derivation shows that the linear AMF is not strictly appropriate even in the limit as plume strength goes to zero.

Furthermore, this weak plume limit is itself not quite appropriate if it implies that the plume is so weak that it is undetectable. The locally most powerful (LMP) solution that corresponds to the weak plume limit ($\epsilon \to 0$) is empirically found to be less effective than the GLRT (indeed, often less effective than the linear matched filter) when the plume is strong enough to be detected with a reasonable false alarm rate.

7. REFERENCES

[1] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*. New York: Springer, 2005.

[2] A. Schaum, "Continuum fusion: a theory of inference, with applications to hyperspectral detection," *Optics Express*, vol. 18, pp. 8171–8181, 2010.

[3] J. Theiler, "Confusion and clairvoyance: some remarks on the composite hypothesis testing problem," *Proc. SPIE*, vol. 8390, p. 839003, 2012.

[4] A. Schaum, "Clairvoyant fusion: a new methodology for designing robust detection algorithms," *Proc. SPIE*, vol. 10004, p. 100040C, 2016.

[5] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerospace and Electronic Systems*, vol. 28, pp. 208–216, 1992.

[6] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. New Jersey: Prentice Hall, 1998, vol. II.

[1] We chose ϵ so that the AMF in Eq. (2) would see a 2.5 sigma effect.

Table 1. Expressions for detecting absorptive plumes on Gaussian backgrounds

Additive	AMF-$\mathbf{t}$	$\mathcal{D}_{\text{AMF-}\mathbf{t}}(\mathbf{x}) = -\mathbf{t}'R^{-1}(\mathbf{x}-\boldsymbol{\mu})$
	AMF-$T\boldsymbol{\mu}$	$\mathcal{D}_{\text{AMF-}T\mu}(\mathbf{x}) = -(T\boldsymbol{\mu})'R^{-1}(\mathbf{x}-\boldsymbol{\mu})$
LMP	QMF	$\mathcal{D}_{\text{QMF}}(\mathbf{x}) = -(T\mathbf{x})'R^{-1}(\mathbf{x}-\boldsymbol{\mu}) + \tau$
GLRT	$\widehat{\epsilon}$	$\widehat{\epsilon} = \mathcal{D}_{\text{QMF}}(\mathbf{x}) / \left[(T\mathbf{x})'R^{-1}T\mathbf{x} + (T\mathbf{x})'TR^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]$
	GLRT	$\mathcal{D}_{\text{GLRT}}(\mathbf{x}) = \mathcal{D}_{\text{QMF}}(\mathbf{x}) / \sqrt{(T\mathbf{x})'R^{-1}T\mathbf{x} + (T\mathbf{x})'TR^{-1}(\mathbf{x}-\boldsymbol{\mu})}$
Known ϵ	Clairvoyant	$\mathcal{D}(\epsilon,\mathbf{x}) = (\exp(\epsilon T)\mathbf{x}-\boldsymbol{\mu})' R^{-1} (\exp(\epsilon T)\mathbf{x}-\boldsymbol{\mu}) - 2\epsilon\tau - (\mathbf{x}-\boldsymbol{\mu})'R^{-1}(\mathbf{x}-\boldsymbol{\mu})$

Table 2. Implanted plumes on real OMI data.

(a) SO_2

Detector	FAR@DR=0.5	AUC	DR@FAR=0.5
AMF-$\mathbf{t}$	0.03823	0.8415	0.8856
AMF-$T\boldsymbol{\mu}$	0.00920	0.9036	0.9412
QMF	0.02832	0.8790	0.9589
$\widehat{\epsilon}$	0.15119	0.8380	0.9837
GLRT	0.04628	0.9088	0.9795
Clairvoyant	0.01092	0.9406	0.9914

(b) NO_2

Detector	FAR@DR=0.5	AUC	DR@FAR=0.5
AMF-$\mathbf{t}$	0.07225	0.8357	0.9092
AMF-$T\boldsymbol{\mu}$	0.01619	0.9057	0.9481
QMF	0.05404	0.8855	0.9772
$\widehat{\epsilon}$	0.07104	0.9030	0.9857
GLRT	0.02581	0.9211	0.9826
Clairvoyant	0.01104	0.9445	0.9932

Table 3. Implanted plume on a Gaussian background

(a) SO_2

Detector	FAR@DR=0.5	AUC	DR@FAR=0.5
AMF-$\mathbf{t}$	0.05076	0.8206	0.8659
AMF-$T\boldsymbol{\mu}$	0.00630	0.8838	0.9131
QMF	0.01750	0.9120	0.9759
$\widehat{\epsilon}$	0.02573	0.9386	0.9766
GLRT	0.00241	0.9440	0.9763
Clairvoyant	0.00200	0.9658	0.9983

(b) NO_2

Detector	FAR@DR=0.5	AUC	DR@FAR=0.5
AMF-$\mathbf{t}$	0.05359	0.8242	0.8730
AMF-$T\boldsymbol{\mu}$	0.00585	0.8812	0.9087
QMF	0.02695	0.8771	0.9446
$\widehat{\epsilon}$	0.05907	0.8864	0.9477
GLRT	0.00521	0.9111	0.9467
Clairvoyant	0.00388	0.9481	0.9964

[7] M. D. Foote, P. E. Dennison, A. K. Thorpe, D. R. Thompson, S. Jongaramrungruang, C. Frankenberg, and S. C. Joshi, "Fast and accurate retrieval of methane concentration from imaging spectrometer data using sparsity prior," *IEEE Trans. Geoscience and Remote Sensing*, 2020, doi: 10.1109/TGRS.2020.2976888.

[8] M. Bar-Tal and S. R. Rotman, "Performance measurement in point target detection," *Infrared Physics & Technology*, vol. 37, pp. 231–238, 1996.

[9] A. K. Mebust, A. R. Russell, R. C. Hudman, L. C. Valin, , and R. C. Cohen, "Characterization of wildfire NO_X emissions using MODIS fire radiative power and OMI tropospheric NO_2 columns," *Atmospheric Chemistry and Physics*, vol. 11, pp. 5839–5851, 2011.

[10] K. Yang, N. A. Krotkov, A. J. Krueger, S. A. Carn, P. K. Bhartia, and P. F. Levelt, "Retrieval of large volcanic SO_2 columns from the Aura Ozone Monitoring Instrument: Comparison and limitations," *J. Geophysical Research*, vol. 112, p. D24S43, 2007.

[11] S. Matteoli, M. Diani, and J. Theiler, "An overview background modeling for detection of targets and anomalies in hyperspectral remotely sensed imagery," *IEEE J. Sel. Topics in Applied Earth Observations and Remote Sensing*, vol. 7, pp. 2317–2336, 2014.

[12] D. Manolakis, D. Marden, J. Kerekes, and G. Shaw, "On the statistics of hyperspectral imaging data," *Proc. SPIE*, vol. 4381, pp. 308–316, 2001.

[13] S. Matteoli, M. Diani, and G. Corsini, "Closed-form nonparametric GLRT detector for subpixel targets in hyperspectral images," *IEEE Trans. Aerospace and Electronic Systems*, vol. 56, no. 2, pp. 1568–1581, 2020.

[14] J. Theiler, "Matched-pair machine learning," *Technometrics*, vol. 55, pp. 536–547, 2013.

[15] A. Ziemann, M. Kucer, and J. Theiler, "A machine learning approach to hyperspectral detection of solid targets," *Proc. SPIE*, vol. 10644, p. 1064404, 2018.