Noise Contaminated Transmittance

Andrew Zardecki, Brian D. McVey, Douglas H. Nelson, and Mark J. Schmitt
Los Alamos National Laboratory, MS E541, Los Alamos, NM 87545

ABSTRACT

We compare the efficiency of a classifier based on probabilistic neural networks and the general least squares method. Both methods must accommodate noise due to uncertainty in the measured spectrum at each wavelength. The evaluation of both methods is based on a simulated transmittance spectrum, in which the received signal is supplemented by an additive admixture of noise. To obtain a realistic description of the noise model, we generate several hundred laser pulses for each wavelength under consideration. These pulses have a predetermined correlation matrix for different wavelengths; furthermore, they are composed of three components accounting for the randomness of the observed spectrum. The first component is the correlated $1/f$ noise; the second component is due to uncorrelated $1/f$ noise; the third one is the uncorrelated white noise. The probabilistic neural network fails to retrieve the species concentration correctly for large noise levels; on the other hand, its predictions being confined to a fixed number of concentration bins, the network produces relatively small variances. To a large extent, the general least square method avoids the false alarms. It reproduces the average concentrations correctly; however, the concentration variances can be large.

Keywords: Species detection, Species identification, LIDAR simulation

1. INTRODUCTION

Molecular absorption determines radiative heating and cooling in the infrared, thus affecting the climate modeling. Transmission of laser beams and the determination of trace gas concentrations are exacting requirements,¹ only partially solved with high-resolution spectroscopy.

We assume that Beer's law describes the transmittance through either the atmosphere or effluent region. In simplified form, Beer's law is

$$\tau_{\kappa} = \exp\left(-\sum_{j=1}^J \sigma_{\kappa j} c_j L_j\right), \quad (1)$$

where τ_{κ} is the transmission at wavenumber κ , $\sigma_{\kappa j}$ is the absorption cross section for species j at wavelength κ , c_j is the concentration of species j , $j = 1, \dots, J$, and L_j is the path length for species j . Equation. (1) does not account for multiple scattering effects; however, it is a good approximation to the Beer Law when applied to ratios of pairs of wavelengths.

To determine the concentration of chemicals under controlled release conditions, a combined experimental and computational approach relying on tunable CO₂ lasers and chemometric analysis is often employed.² In typical situations, the transmittance data, used by chemometric analysis, are encumbered by uncertainties in the measurement of the return signal caused by speckle, detector noise, and laser scintillations. A simple model represents the transmittance as a sum of the noiseless term and the term given in terms of the white noise. Thus, if $\tau_0(\kappa)$ denotes the theoretical transmittance at wavenumber κ , the observed transmittance is

$$\tau(\kappa) = \tau_0(\kappa)[1 + W(\kappa)] . \quad (2)$$

In Eq. (1), $W(\kappa)$ is a white noise random process; it is uncorrelated with respect to κ .

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

MASTER
kg

DISCLAIMER

**Portions of this document may be illegible
in electronic image products. Images are
produced from the best available original
document.**

In a typical lidar experiment, we deal with a series of pulses, which, in general, are correlated in time. The resulting random process is not reducible to the white noise; rather, it requires a more sensible model. In the following we retain the additive form of Eq. (1), in which the noise term will be prescribed by more general conditions. Specifically, we include both the correlated and uncorrelated $1/f$ -noise components.

2. NOISY TRANSMITTANCE MODEL

The computational algorithm yielding the prescribed spectral density follows the methodology used in modeling the propagation of laser beams through the atmosphere with a given turbulence profile.³ In our case the problem is slightly more general for we consider a vector random process $V_p(t)$ with zero mean, in which the subscript p refers to a discrete value of the wavenumber. The correlation function of $V_p(t)$ is

$$\Gamma_{pq}(t_1 - t_2) = \langle V_p(t_1) V_q(t_2) \rangle, \quad (3)$$

where the angular brackets denote an ensemble average and where, for stationary processes, the correlation function depends only on the time difference of the temporal arguments.

For a finite number of laser pulses the time variable is discretized too, leading to a two-dimensional array representation of $V_p(t_k) = V_k^{(p)}$, $k = 0, 1, \dots, N-1$. Let $v_n^{(p)}$ denote a discrete Fourier transform of $V_k^{(p)}$ defined as

$$v_n^{(p)} = \sum_{k=0}^{N-1} V_k^{(p)} e^{2\pi i k n / N}, \quad (4)$$

with the inverse transform given by

$$V_k^{(p)} = \frac{1}{N} \sum_{n=0}^{N-1} v_n^{(p)} e^{-2\pi i k n / N}. \quad (5)$$

In general, the discrete Fourier transform converts a set of N complex numbers into another set of N complex numbers. If the original time series is real, the redundancy in the Fourier transform translates through the condition: $v_n^{(p)} = v_{N-n}^{(p)*}$, for $n = 1, 2, \dots, N/2-1$; furthermore, the components $v_0^{(p)}$ and $v_{N/2}^{(p)}$ need to be real.

In the case of a continuous-time, stationary random process, the Wiener-Khintchine theorem provides a relationship between the correlation function and the spectral density of the random process. For a finite number of pulses, the discrete representation of the random process automatically provides the temporal truncation of the random signal needed to derive the Wiener-Khintchine theorem.⁴ In fact, a discrete time version of Eq. (3) is

$$\Gamma_{k-q}^{(pq)} = \langle V_k^{(p)} V_q^{(q)} \rangle. \quad (6)$$

If we use the inverse Fourier transform representation, given by Eq. (5), we obtain

$$\Gamma_{k-q}^{(pq)} = \frac{1}{N} \sum_{n=0}^{N-1} C_n^{(pq)} e^{-2\pi i n l / N}, \quad (7)$$

provided that

$$\langle v_m^{(p)*} v_n^{(q)} \rangle = N C_n^{(pq)} \delta_{mn} . \quad (8)$$

Equation (7), derived under the assumption that different Fourier components are uncorrelated, is a manifestation of the Wiener-Khinchine relation in the discrete case.⁵ The complex numbers $C_n^{(pq)}$ satisfy the same reality conditions as the original Fourier amplitudes. In terms of the real $\text{Re}(v_n^{(p)})$ and imaginary parts $\text{Im}(v_n^{(p)})$ of $v_n^{(p)}$, we obtain the following relations

$$\langle \text{Re}(v_m^{(p)}) \text{Re}(v_n^{(q)}) \rangle = \langle \text{Im}(v_m^{(p)}) \text{Im}(v_n^{(q)}) \rangle = \frac{1}{2} N C_n^{(pq)} \delta_{mn} , \quad (9)$$

$$\langle \text{Re}(v_m^{(p)}) \text{Im}(v_n^{(q)}) \rangle = 0 , \quad (10)$$

which are equivalent to Eq. (8). Equations (9) and (10) imply that, to generate the complex spectrum whose spectral density satisfies the required reality conditions, one has to generate two sets of real amplitudes, mutually uncorrelated, with identical variance. Numerically, this procedure is readily implemented by replacing a complex array of length N by a real array of length $2N$.⁶

The expression for the correlation function simplifies when $C_n^{(pq)}$ factorizes into a product of two factors G and f , in which G accounts for the correlations of laser wavenumbers, whereas f accounts for the correlations of Fourier amplitudes:

$$C_n^{(pq)} = G^{(pq)} f(n) . \quad (11)$$

By analogy to cross-spectrally pure light beams in the theory of optical coherence,^{4,7} the decorrelation of spectral properties of laser wavenumbers and those of the Fourier amplitudes will be called the cross-spectral purity of the Fourier spectrum.

We now require that the factor $C_n^{(pq)}$ in the spectral density be represented as a sum of three terms:

$$C_n^{(pq)} = \frac{A^2}{n^\alpha} + \frac{B_{pq}}{n^\alpha} + W^2 \delta_{kl} , \quad (12)$$

corresponding to the correlated, partially correlated, and white noise parts, respectively. Written in this form, Eq. (12) implies that the spectral purity condition applies to each of the three parts separately.

The interpretation of the three terms in Eq. (12) can be obtained by inspection of the corresponding spectral densities. The first term, in which the correlation strength A^2 is independent of the laser wavenumber p , describes the $1/f$ noise with exponent α . The $1/f$ noise has the same realization for all the wavenumbers; it is the correlated component of the spectrum. In the second term, the positive-definite square matrix B_{pq} accounts for inter-wavenumber correlations. When we write B_{pq} in the form

$$B_{pq} = \sqrt{B_p} \sqrt{B_q} b_{pq} , \quad (13)$$

we factor out the covariance matrix b_{pq} (with diagonal elements equal to unity) and the strengths B_p and B_q of the contribution due to wavenumbers p and q . Finally, the third term in Eq. (12) describes the effect of uncorrelated (with respect to wavenumbers) white noise whose strength is given by W^2 . Unlike the first term, we do not assume the same realization of white noise for different wavenumbers.

The random process with the cross-spectral density given by Eqs. (8) and (12) can be obtained by considering three complex, zero-mean, Gaussian random processes $\xi_n, \eta_n^{(p)}, \zeta_n^{(p)}$ intervening as three components of the Fourier amplitude $v_n^{(p)}$:

$$\frac{v_n^{(p)}}{\sqrt{N}} = \frac{A}{n^{\alpha/2}} \xi_n + \frac{\sqrt{B_p}}{n^{\alpha/2}} \eta_n^{(p)} + W \zeta_n^{(p)} . \quad (14)$$

In fact, the spectral density as specified by Eq. (8) with $C_n^{(pq)}$ given by Eq. (12) is retrieved if the following conditions are obeyed:

$$\langle \xi_m^* \xi_n \rangle = \delta_{mn} , \quad (15)$$

$$\langle \eta_m^{(p)*} \eta_n^{(q)} \rangle = b_{pq} \delta_{mn} , \quad (16)$$

$$\langle \zeta_m^{(p)*} \zeta_n^{(q)} \rangle = \delta_{pq} \delta_{mn} , \quad (17)$$

while all other correlations vanish.

3. NOISE MODELS

The inverse Fourier transform formula, given by Eq. (5), converts the random series of Fourier amplitudes described by Eqs. (14) - (17) into a random process. The goal of this section is to investigate the effect of the α -exponent of the $1/f$ noise, as well as the effect of wavenumber correlations on the resulting random process.

3.1. Fractal Noise

The $1/f$ noise is a random process defined in terms of the shape of its power spectrum density. The power or the square of some variable associated with the random process, measured in a narrow bandwidth, is proportional to a power of reciprocal frequency.⁸ The exponent α , introduced in Eq. (12) is usually close to 1. When $\alpha = 1$, the $1/f$ noise can be identified with the Brownian motion; otherwise, one speaks about a fractional Brownian motion, used extensively in the study of random fractals.⁹

To study the effect of α on the $1/f$ noise, we select one laser line subject to the uncorrelated part of the random noise. In the accompanying plots, Figs. 1-3, we show the $1/f$ random process for α varies between 0.3, 1.1, and 1.9, respectively. As α increases, the $1/f$ noise changes its shape, essentially from white noise (low values of α) to a highly correlated random process (α close to 2). Reference 9 provides examples of stochastically composed fractal music based on different types of noise.

FIGURE 1. Fractal noise pattern; $\alpha = 0.3$.

FIGURE 2. Fractal noise pattern; $\alpha = 1.1$.

FIGURE 3. Fractal noise pattern; $\alpha = 1.9$.

3.2. Wavenumber correlations

To illustrate the effect of wavenumber correlations, we select two laser lines: 9R30 and 9R08, corresponding to frequencies of 1084.635 cm^{-1} and 1070.462 cm^{-1} , respectively. In Figs. 4–6, we show the noise pattern consisting of a mixture of the uncorrelated $1/f$ noise and partially correlated $1/f$ noise in equal proportions, with $\alpha = 1$. This mixture is supplemented by a 10% component of white noise. In the first case, Fig. 6, there is no correlation between the noise patterns at two wavelengths. The second case, Fig. 7, corresponds to small correlation, in which the partially correlated component has the covariance coefficient $b_{12} = 0.1$. Finally, in the third case, Fig. 8, the partially correlated part is nearly completely correlated, $b_{12} = 0.98$, leading to almost indistinguishable patterns.

FIGURE 4. Mixture of correlated and totally uncorrelated noise in equal proportions, supplemented by 10% white noise. The two lines represent two different wavenumbers.

FIGURE 5. Mixture of correlated and weakly correlated noise in equal proportions, supplemented by 10% white noise. The two lines represent two different wavenumbers.

FIGURE 6. Mixture of correlated and strongly correlated noise in equal proportions, supplemented by 10% white noise. The two lines represent two different wavenumbers.

4. CONCENTRATION RETRIEVAL

We use two algorithms for species concentration retrieval from transmittance measurements. The general least squares (GLS) algorithm minimizes a merit function; the probabilistic neural network (PNN) approach is essentially a classifier.

4.1. General least squares

In the context of laser transmittance, the least squares problem is defined as the minimization of the merit function

$$\chi^2 = \sum_{\kappa=1}^K \left[\frac{y_{\kappa} - \sum_{j=1}^J \sigma_{\kappa j} c_j L_j}{s_{\kappa}} \right]^2, \quad (18)$$

where y_{κ} is the negative natural logarithm of the measured transmittance at wavenumber κ , and where s_{κ} is the measurement error. Defining matrix elements

$$A_{\kappa j} = \frac{\sigma_{\kappa j} L_j}{s_{\kappa}} \quad (19)$$

and vector components

$$b_{\kappa} = \frac{y_{\kappa}}{s_{\kappa}}, \quad (20)$$

the minimization of χ^2 is recast in the matrix form: Find a vector $\vec{c}$ that minimizes $\chi^2 = |\mathbf{A}\vec{c} - \mathbf{b}|^2$. The solution to this problem can be obtained by singular value decomposition of the matrix $\mathbf{A}$, which is well adapted to handle singular and ill-conditioned matrices.⁶

4.2. Probabilistic neural network

Probabilistic neural networks are related to the Bayesian decision theory, in which one estimates a probability density function based on training patterns. Borrowing from the approach of Parzen¹⁰ to the probability density function estimation, Specht¹¹ showed that by organizing the flow of operations into layers, and assigning primitive operations to individual neurons in each layer, the algorithm can be made to resemble a four-layer feedforward network with exponential activation function.

The gist of Specht's algorithm is quite simple. We sample an m -component multivariate random vector $\vec{x} = [x_1, \dots, x_m]$. The K populations from which the samples are drawn will be indexed by $k, k = 1, \dots, K$. Let the prior probability of an unknown sample being drawn from population k be p_k , and the associated cost be c_k (the costs allow one to weight the populations according to the prior information). Given the probability density function $f_k(\vec{x})$ for population k , the Bayes optimal decision rule is to classify $\vec{x}$ into population i if

$$p_i c_i f_i(\vec{x}) > p_j c_j f_j(\vec{x}) \quad (21)$$

for all populations j not equal to i . In practice, the prior probabilities and costs are often treated as being equal, and hence can be ignored.

In PNN, the probability density function estimator is given as a weighted sum of Parzen's windows

$$f_k(\vec{x}) = \frac{1}{N} \sum_{n=1}^N W\left(\frac{\vec{x} - \vec{x}_n}{h}\right), \quad (22)$$

where h describes the width of the window function $W(\vec{x})$, and where N is the number of observations of the vector $\vec{x}$ in the k th population. Actually, the width parameter can be different for different components of the vector $\vec{x}$; more generally, it becomes an array accounting for correlations between different components of $\vec{x}$.

Equation (22) can be interpreted as an mean distribution, obtained as an arithmetic average of distributions centered at sample points. For this reason, the window function $\phi(\vec{x})$ can be chosen as any positive definite normalized function. It can be readily shown that, as the window size parameter h tends to zero and the number of sample points N tends to infinity, the Parzen density estimator, Eq. (22), tends to the true density of $\vec{x}$.

5. NUMERICAL RESULTS

To illustrate the former considerations, we assume four different noise models and analyze their impact on concentration retrieval. In the numerical simulation reported in this section, we generate 256 samples, each containing the values of transmittance given at 40 wavelengths in the infrared, ranging from 9.30 to 10.65 μm . The two gases that produce the transmittance spectrum are triethyl phosphate (TEP) with concentration of 80 ppm, and ethylene (ETH) whose concentration is 20 ppm. For a given noise level, the concentration retrieval results are then averaged over the 256 samples, resulting in a mean transmittance and the retrieval error expressed in terms of the standard deviation.

We start with the correlated noise model, in which the random perturbation of the transmittance is the same, independent of the wavenumber. This is equivalent to retaining only the first term in Eq. (14). Figure (7) gives the retrieval results for this case.

FIGURE 7. Results of concentration retrieval for correlated noise model as a function of noise strength.

The same amount of noise added for each wavenumber is equivalent to a shift of the transmittance curve with respect to the transmittance value. The GLS algorithm, which uses the relative values with respect to the mean, is capable of nearly perfect retrieval even for high noise amplitudes. On the other hand, the PNN method, which relies on the trained pattern, fails as the noise strength increases.

Figure 8 shows the retrieval results corresponding to uncorrelated noise model. In this scenario the noisy perturbations corresponding to different wavenumbers are uncorrelated. On the average, the GLS method still retrieves the

concentration correctly. The results of retrieval for individual samples, however, deteriorate as the noise strength increases, resulting in much higher GLS variance than the PNN one.

FIGURE 8. Results of concentration retrieval for uncorrelated noise model as a function of noise strength.

The case of white noise, illustrated in Fig. 9, is the worst-case scenario. Here the PNN method fails even for small noise strengths, whereas the GLS algorithm produces errors close to 100%.

FIGURE 9. Results of concentration retrieval for white noise model as a function of noise strength.

Finally, in Fig. 10, we show the retrieval results for mixed case, in which a sum of equal contributions from the three noise terms is taken.

FIGURE 10. Results of concentration retrieval for a mixed noise model as a function of noise strength.

This case can be viewed as an intermediate between the best case scenario of Fig. 7 and the worst case scenario of Fig. 9.

6. CONCLUSION

The random vector process, defined in terms of the power spectrum, was investigated as a function of the $1/f$ exponent, as well as in terms of the correlation between two different wavenumbers. Our objective is to use the model to study the effect of the noise pattern on the concentration retrieval from lidar transmittance measurements. Two different techniques for concentration retrieval were tested. The general least square algorithm provides, on average, good concentration results, though at the cost of high variance. The probabilist neural network classifier fails for large noise levels; the variance, however, is smaller than the GLS variance.

REFERENCES

1. R. M. Measures, *Laser Remote Sensing* (Krieger Publishing, Malabar, FL, 1992)
2. J. R. Quagliano, P. O. Stoutland, R. R. Petrin, R. K. Sander, R. J. Romero, M. C. Whitehead, C. R. Quick, J. J. Tiee, and L. J. Jolin, "Quantitative chemical identification of four gases in remote infrared (9–11 μm) differential absorption lidar experiments," *Appl. Opt.* 36, 1915–1927 (1997).
3. J. A. Fleck, J. R. Morris, and M. D. Feit, "Time-dependent propagation of high-energy laser beams through the atmosphere," *Appl. Phys.* 10, 129–160 (1976).
4. L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, 1995).
5. R. G. Brown and P. Y. C. Hwang, *Introduction to Random Signals and Applied Kalman Filtering* (Wiley, New York, 1992).

6. W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, *Numerical Recipes in C* (Cambridge University Press, Cambridge, 1992).
7. A. Zardecki and C. Delisle, "Distribution of the sum of photocounts in Gaussian light of any state of coherence and polarization," *Can. J. Phys.* **51**, 1017-1023 (1973).
8. M. S. Kershner, "1/f noise," *Proc. IEEE*, **70**, 212-218 (1982).
9. H. O. Peitgen and D. Saupe, editors, *The Science of Fractals* (Springer, New York, 1988).
10. E. Parzen, "On estimation of a probability density function and mode," *Ann. Math. Stat.* **33**, 1065-1076 (1962).
11. D. F. Specht, "Probabilistic Neural Networks," *Neural Networks* **3**, 109-118 (1990).

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.