Gas plume species identification by regression analyses

Dave Pogorzala
May 12, 2004
Introduction – the RIT Gas Problem

• Detection
  – does this pixel contain a plume?

• Identification
  – if so, which gases are in the plume?

• Quantification
  – what is the mixing ratio or column density of these gases?
Overview

- Data
- Stepwise regression
- Radiance model
- Results
- Conclusions
Data Sources

Test Image

- DIRSIG image
- two plumes
- one gas per plume
- Freon-114 and NH$_3$
- plumes do not overlap
- complex background
- 128 bands
- 7.5 - 13.6 µm
- SEBASS band centers
- 10.73 µm
Data Sources

Test Image

- DIRSIG image
- two plumes
- one gas per plume
- Freon-114 and NH$_3$
- plumes do not overlap
- complex background
- 128 bands
- 7.5 - 13.6 µm
- SEBASS band centers
- 10.73 µm
Data Sources

PNNL Gas Spectra
- used to create DIRSIG image
- used to populate basis vector library
- our subset contains 30 gases
- most measured at 3 temps: 5, 25 and 50°C
- basis vector library contains 88 gas spectra
Overview

• Data

• Stepwise regression

• Radiance model

• Results

• Conclusions
Stepwise Regression

- Stepwise regression is an iterative approach that selects only those basis vectors that contribute to the fit.
- This eliminates needless vectors from the model that have insignificant abundances.
- In applying this to the gas problem, we begin with a set of vectors that include all possible gases. The hope is that only those gases contained in the plume will be found by the regression.
- This allows us to identify gases without prior knowledge of which are present.
Stepwise Regression

Formulate regression as:

\[ x = Af + \varepsilon \]

\[ x = (J \times 1) \text{ pixel vector} \]
\[ A = (J \times N) \text{ matrix of basis vectors} \]
\[ f = (N \times 1) \text{ vector of abundances} \]
\[ \varepsilon = \text{error vector} \]
\[ J = 128 \text{ bands} \]
\[ N = \text{number of basis vectors currently in model} \]
\[ M = \text{total available basis vectors} \]
Stepwise Regression

- Determining whether or not to keep a vector in $A$ is done by an $F$-test (calculated at a probability of 0.99) based on the Analysis of Variance (ANOVA).

- ANOVA compares the Sum of Squares due to Regression (SSR) from the $N$ element model to the $N-1$ element model (primed quantities denote a transpose).

$$SSR_N = f' A' x$$

$$SSR_{N-1} = f'_{N-1} A'_{N-1} x$$
Stepwise Regression

- If the difference between these two SSRs is more significant than the Sum of Squares about Regression (SSE) normalized by its degrees of freedom (its Mean Squared Error, MSE), then the vector is added.

\[
SSE = \text{total SS} - SSR = x'x - f'A'x
\]

\[
MSE = \frac{SSE}{\text{dof}}
\]

- The difference is significant if it’s ratio over the MSE is greater than the F-statistic

\[
\frac{f'A'x - f'_{N-1}A'_{N-1}x}{MSE} > F - \text{stat}
\]
Overview

• Data
• Stepwise regression
• Radiance model
• Results
• Conclusions
Radiance Model

- Neglecting atmospheric effects, we can express a plume pixel as:

\[
L(\lambda) = \left[ \sum_{j=1}^{C} \beta_j \varepsilon_j(\lambda) B(\lambda, T_s) \right] \left[ 1 - \sum_{i=0}^{D} c_i k_i(\lambda) \right] + \sum_{i=0}^{D} c_i k_i(\lambda) B(\lambda, T_p)
\]

\[
L(\lambda) = \text{pixel spectrum}
\]

\[
\sum_{j=1}^{C} \beta_j \varepsilon_j(\lambda) B(\lambda, T_s) = \text{self - emitted surface radiance}
\]

\[
c_i = \text{column density of gas } i
\]

\[
k_i(\lambda) = \text{absorption spectrum of gas } i
\]

\[
\sum_{i=0}^{D} c_i k_i(\lambda) B(\lambda, T_p) = \text{self - emitted plume radiance}
\]
Radiance Model

- We would like to account for plume emission/absorption.
  - The plume is in emission when \( T_p > T_s \), absorption when \( T_p < T_s \).

1. Begin by inverting the background radiance term to solve for a spectral brightness temperature. This is done on a per-pixel basis.
2. The maximum of this is the surface temperature estimate, \( \hat{T}_{s,(x,y)} \)
3. Define five values of \( \Delta T \): -10, -5, 0, 5 and 10
4. Rewrite \( T_p \) as: \( T_p = \hat{T}_{s,(x,y)} \pm \Delta T \)
Radiance Model

• Substituting the temperature contrast term in for $T_p$, we arrive at the final radiance model:

$$L(\lambda) - L_{(x,y)}(\lambda) = \sum_{i=0}^{D} c_i k_i(\lambda) \left[ B(\lambda, \hat{T}_{s,(x,y)} \pm \Delta T) - L_{(x,y)}(\lambda) \right]$$

• This can be related to the regression model defined earlier:

$$\mathbf{x} \rightarrow L(\lambda) - L_{(x,y)}(\lambda)$$

$$\mathbf{A} \rightarrow k_i(\lambda) \left[ B(\lambda, \hat{T}_{s,(x,y)} \pm \Delta T) - L_{(x,y)}(\lambda) \right]$$

$$\mathbf{f} \rightarrow c_i$$

• $M$ is now (88 PNNL spectra)(5 $\Delta T$ values) = 440 basis vectors
Overview

• Data
• Stepwise regression
• Radiance model
• Results
• Conclusions
Results

• Stepwise regression was performed on each plume pixel

• A binary image cube 256 x 256 x 440 was created
  – one “map” per basis vector

• The basis vectors found by the regression are denoted by turning “on” the pixel (x,y) in each corresponding map.

• The 440 maps were then collapsed down to 30, one map per gas species.
**Results**

- Detection maps for Freon-114 along with the “top 11” false alarms

1. Freon-114
2. Phosgene (phg)
3. Trichloroethylene (tce)
4. Dibromoethane (edb)
5. Sulfur Dioxide (SO$_2$)
6. Hydrazine (hyd)
7. Vinyl Chloride (vcl)
8. Benzene (C$_6$H$_6$)
9. Formaldehyde (HCHO)
10. Dichloropropane (dclp-13)
11. Carbon Monoxide (CO)
12. Methane (CH$_4$)
Results

• Detection maps for Ammonia along with the “top 3” false alarms

1. Ammonia (NH$_3$)
2. Hydrazine (hyd)
3. Freon-12 (f12)
4. Acrolein (acrol)
Results
Results

NH₃ and its False Alarms (normalized)

absorption (offset for clarity)

acrol
f12
hyd
nh3

microns
Results

• False detections are attributed to spectral overlap of key absorption features.

• Magnitude of absorption features are not important since the stepwise regression was unconstrained.
  – Basis vector abundances are allowed to take on any value.
Overview

• Data
• Stepwise regression
• Radiance model
• Results
• Conclusions
Conclusions

- Stepwise regression is successful at identifying plume gases.
- The gases are being identified in emission and absorption.
- No prior knowledge is required to correctly identify the gases.
- False detections are attributed to spectral overlap of key absorption features.
Conclusions

Future work:

1. Account for spectral overlap.
   - mask out overlapping absorption features
   - perform regression over specific spectral windows

2. Derive an in-scene background approximation
   - linear combination of background endmembers
   - background material ID using VIS/NIR image

3. Atmospheric compensation
   - which scheme is friendliest to native gases

Questions?