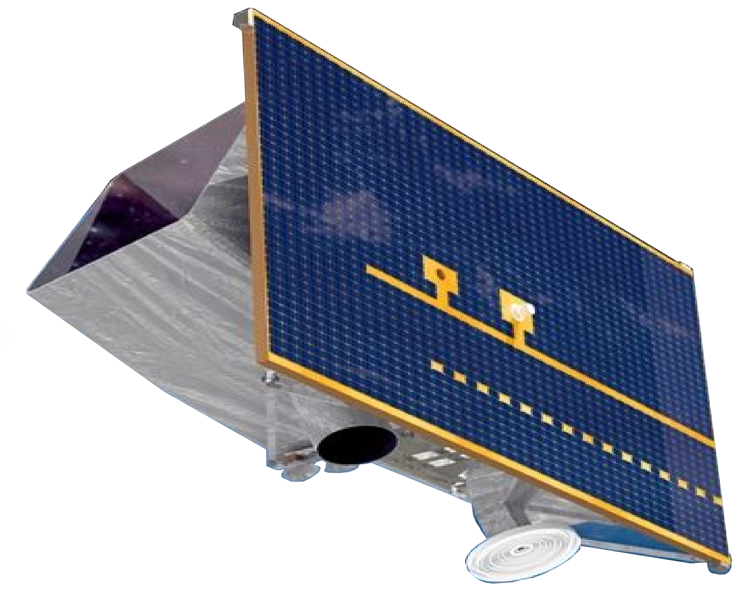
Assessment of Methane Retrieval Algorithms for EnMAP and PRISMA Shortwave Infrared Observations

Juan Bettinelli, Philipp Hochstaffl
juan.bettinelli@dlr.de
Deutsches Zentrum für Luft- und Raumfahrt
(MF–ATP)

DLR



EnMAP

Kamyshlydzha, Turkmenistan
G&O Facility, Homogeneous Albedo

PRISMA

Aix-En-Provence, France
Cattle Farm, Inhomogeneous Albedo

Albedo-Reweighted Matched Filter (AR-MF)

The Albedo-Reweighted Matched Filter (AR-MF) is a linear detection method for identifying CH_4 enhancements by projecting each observed spectrum onto a predefined target signature. To suppress background variability, the method maximizes the signal-to-noise ratio assuming Gaussian noise.
The standard MF enhancement factor α_i for pixel i is computed as:

$$\alpha_i = \frac{\mathbf{t}^{\mathrm{T}}\mathbf{C}^{-1}(\mathbf{y}_i - \boldsymbol{\mu})}{\sqrt{\mathbf{t}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{t}}}$$

α_i: CH_4 enhancement estimate for pixel i
$\mathbf{y}_i$: observed spectrum
$\boldsymbol{\mu}$: mean background spectrum
$\mathbf{t}$: target spectrum
$\mathbf{C}$: covariance matrix of background spectra
r_i: reflectance normalization factor

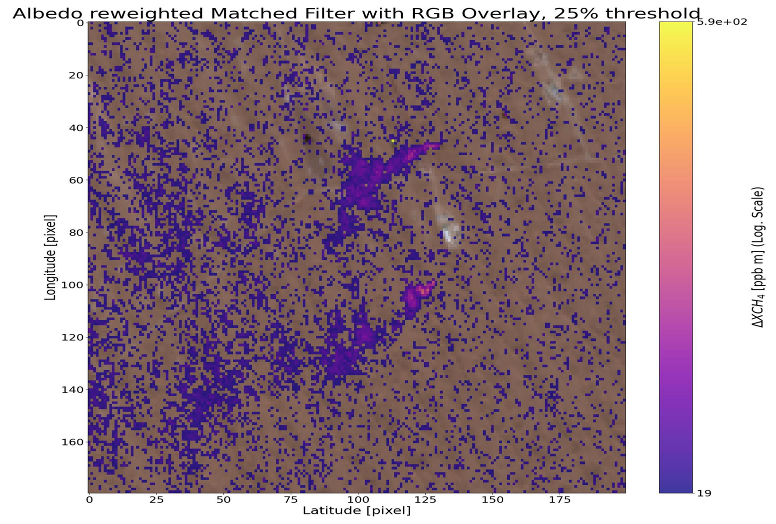
A key enhancement over the standard matched filter is albedo normalization: observed spectra are scaled by their overall reflectance to reduce false positives over bright surfaces (e.g., sand, concrete, snow). An iterative reweighting scheme improves robustness by down-weighting pixels with high variance in background components:

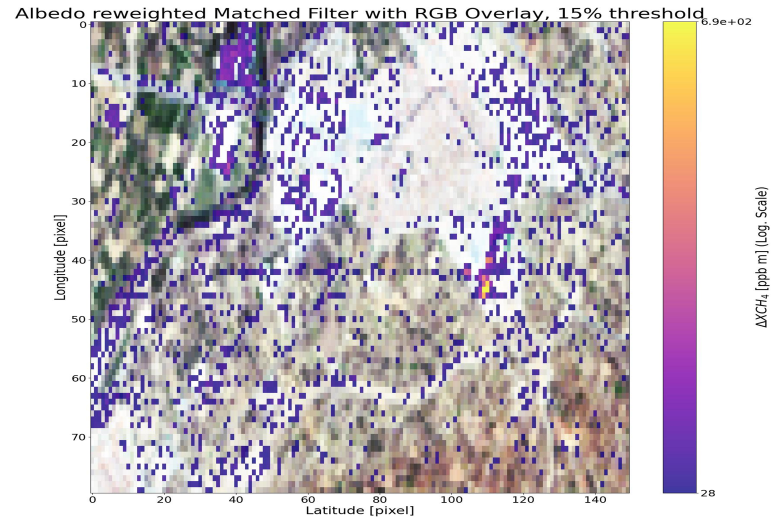
$$\alpha_i = \frac{\mathbf{t}^{\mathrm{T}}\mathbf{C}^{-1}(\mathbf{y}_i - r_i \cdot \mathbf{t})}{\sqrt{\mathbf{t}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{t}}} \qquad r_i = \frac{\mathbf{y}_i^{\mathrm{T}}\boldsymbol{\mu}}{\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\mu}}$$

The target spectrum is derived from synthetic radiative transfer models (Py4CATs & HITRAN).

Pros
- Fast and efficient: Suitable for processing millions of pixels quickly.
- Well-suited for large datasets: Enables regional-scale plume surveys.
- Compatible with EnMAP and PRISMA: Works well with moderate spectral resolution and high spatial detail.
- Improved robustness: Reweighting and albedo correction reduce background bias.

Cons
- False positives: Can produce spurious signals over high-albedo surfaces.
- Water surfaces: May misinterpret specular reflections as methane enhancements.
- No absolute quantification: Outputs relative enhancement, not calibrated column values.





Linear Least Squares (LLS)

The Linear Least Squares (LLS) method fits a linear model to the observed spectrum by combining background and target basis functions. The CH_4 enhancement factor $\mathbf{y}$ is estimated by minimizing the squared residual between model and measurement:

$$\mathbf{y} = \sum_{j=1}^{N} \beta_j \mathbf{x}_j + \alpha \cdot \mathbf{s}$$

$\mathbf{x}_j$: Basis vectors representing the background signal
β_j: Coefficients for the background
$\mathbf{S}$: CH_4 absorption signature
α : Enhancement coefficient to be retrieved

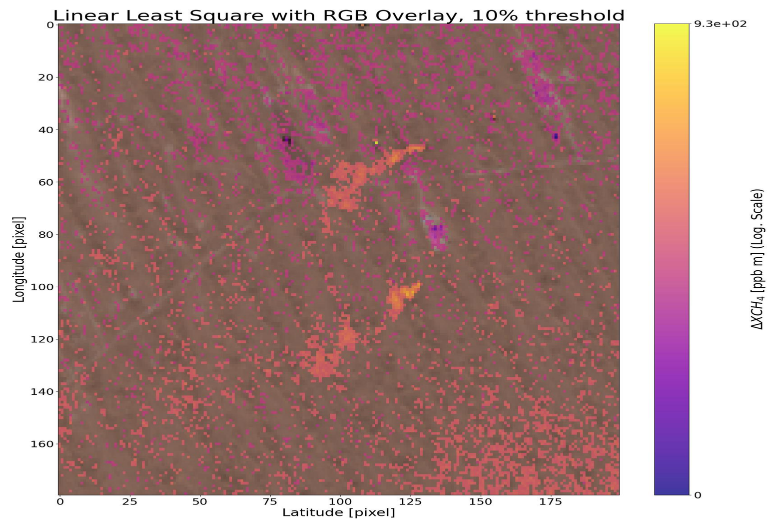
The system is solved in matrix form using ordinary least squares. The matrix $\mathbf{A} = [\mathbf{X}, , \mathbf{s}]$ combines the background and target spectra, and the optimal coefficients are computed by minimizing the squared residual:

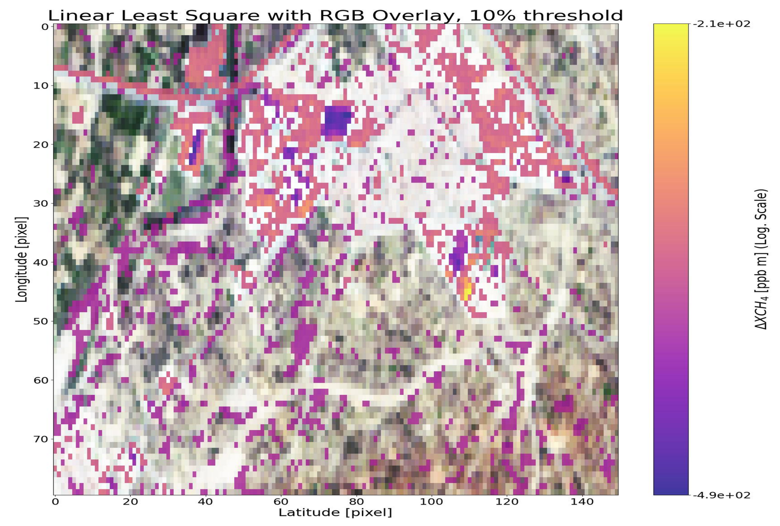
$$\min_{\boldsymbol{\beta},\alpha} \left\| \mathbf{y} - \mathbf{A} \begin{bmatrix} \boldsymbol{\beta} \\ \alpha \end{bmatrix} \right\|^2 \quad \text{with} \quad \mathbf{A} = [\mathbf{X} \quad \mathbf{s}]$$

LLS is straightforward, fast, and does not require prior atmospheric knowledge. However, it assumes the CH_4 signal is orthogonal to the background, which can lead to underestimation or leakage in spectrally complex scenes.

Pros
- Fast and easy to implement: Computationally efficient, ideal for quick analyses.
- Simple linear model: Requires minimal parameter tuning or prior knowledge.
- Good baseline method: Useful for comparative performance benchmarking.

Cons
- Underestimates CH_4 enhancements: Especially when CH_4 signals overlap with background spectral features.
- Sensitive to low signal-to-noise: Performance degrades in noisy measurements.
- False positives: Susceptible to retrieval errors in regions with inhomogeneous albedo (e.g., complex terrain, urban surfaces).
- No physical correction: Lacks atmospheric or surface reflectance corrections.





Generalized Least Square (GLS)

The Generalized Least Squares method estimates methane (CH_4) enhancements by fitting the observed spectrum while explicitly accounting for spectral correlations in the background. It models the observation $\mathbf{y}$ as:

$$\mathbf{y} = \sum_{j=1}^{N} \beta_j \mathbf{x}_j + \alpha \cdot \mathbf{s}$$

$\mathbf{x}_j$: Basis vectors representing the background signal
β_j: Coefficients for the background
$\mathbf{S}$: CH_4 absorption signature
α : Enhancement coefficient to be retrieved

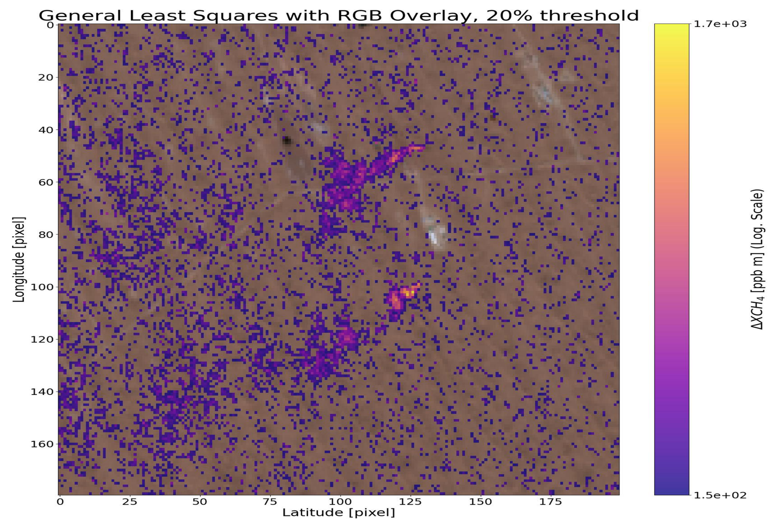
What distinguishes GLS from standard LLS is that the fit is weighted by the inverse of the background covariance matrix C^{-1}:

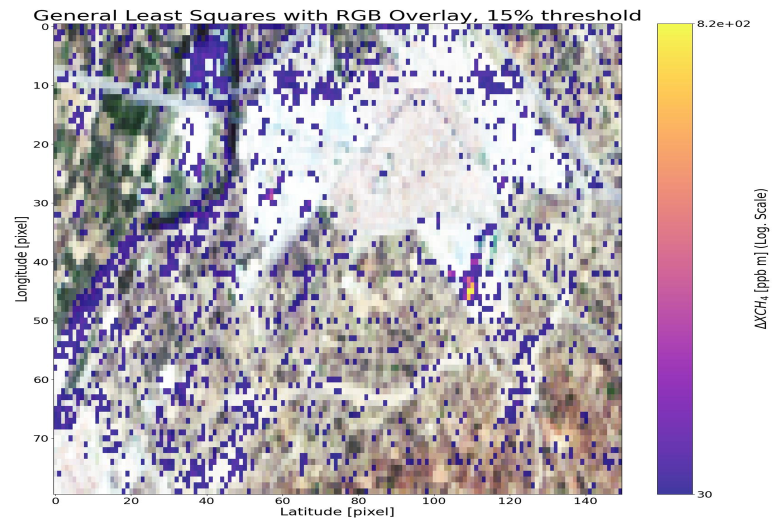
$$\min_{\boldsymbol{\beta},\alpha} \left\| \mathbf{C}^{-1/2} \left(\mathbf{y} - \sum_j \beta_j \mathbf{x}_j - \alpha \cdot \mathbf{s} \right) \right\|^2$$

This allows GLS to downweight spectral components that exhibit strong natural variability (e.g., surface reflectance features), leading to a more reliable separation of methane absorption from background structures.

Pros
- Better signal-to-noise handling: Weights the fit using background variability, improving detection in noisy data.
- Deals well with inhomogeneous albedo: Suppresses false positives in spectrally complex regions.
- Reliable source detection: Accurately separates CH_4 signals from background structures.
- Computationally efficient: Still relatively fast despite using a covariance-weighted fit.
- Physically grounded: Makes use of the statistical structure of the background spectra.

Cons
- Requires background statistics: Needs a representative background dataset to estimate the covariance matrix.
- Sensitivity to covariance quality: Poor or unrepresentative background models can degrade performance.
- More complex than LLS: Implementation and tuning are less straightforward.





Cluster Tuned Singular Value Decomposition (SVD)

The SVD method decomposes the observed spectral data into orthogonal components, separating background features from potential CH_4 signals. It assumes the measured spectrum $\mathbf{y}$ can be expressed as:

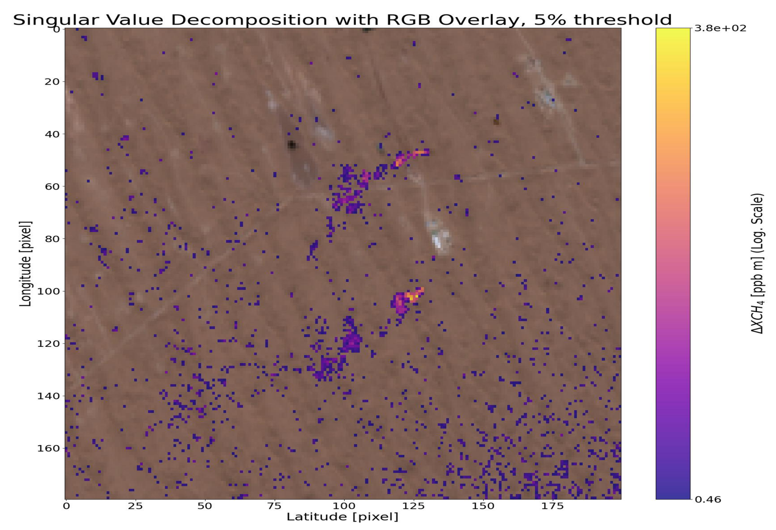
$$\mathbf{y} = \sum_{j=1}^{K} \beta_j \mathbf{u}_j + \alpha \cdot \mathbf{s}$$

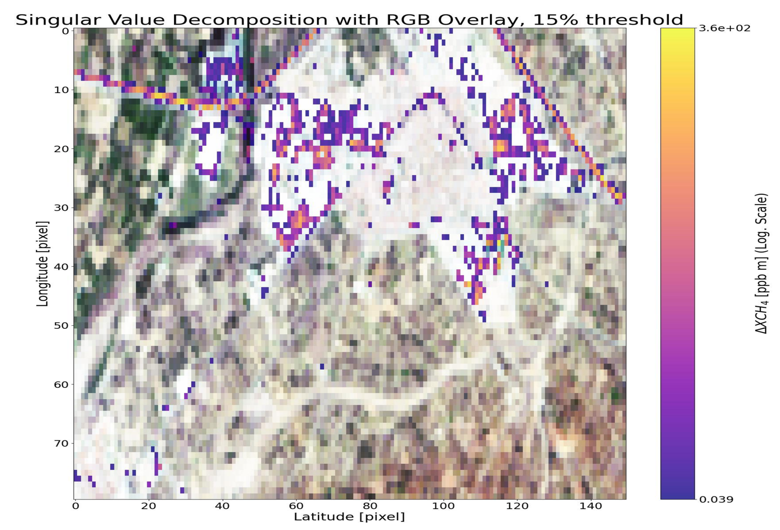
$\mathbf{u}_j$: first K singular vectors (principal components) derived from the background spectra
β_j: Coefficients describing the background contribution
$\mathbf{S}$: CH_4 target spectrum
α: CH_4 enhancement factor

The "cluster-tuned" variant improves performance by applying SVD locally within background clusters (e.g., similar surface types), which reduces spectral variability within each group and leads to more accurate separation of CH_4 signals from the background.
SVD fits the model using ordinary least squares, solving for α after subtracting the reconstructed background signal from the measurement.

Pros
- Separates methane enhancements from background: Effectively distinguishes CH_4 signals from spectral clutter.
- Adaptable to surface variability: Clustering allows tailored background modeling for different surface types.
- No need for external atmospheric data: Entirely data-driven using the observed spectra.
- Moderately fast: Computationally efficient once clusters and basis vectors are defined.

Cons
- Parameter tuning required: Performance depends on selecting a suitable number of K-means clusters and SVD basis vectors per cluster.
- Background leakage possible: If clusters are poorly defined, background features may be misinterpreted as CH_4.
- Assumes local linearity: May struggle in highly nonlinear or mixed-pixel environments.





Beer InfraRed Retrieval Algorithm (BIRRA)

BIRRA is a nonlinear inversion algorithm designed to retrieve trace gas concentrations from infrared spectra using line-by-line radiative transfer modeling. It is based on the Beer–Lambert law for optical absorption:

$$I(\lambda) = I_0(\lambda) \cdot \exp\left(- \sum_k \alpha_k(\lambda) \cdot c_k \cdot l_k \right)$$

$I(\lambda)$: measured radiance
$I_0(\lambda)$: background radiance (without gas absorption)
$\alpha_k(\lambda)$: absorption cross section of species k
c_k : concentration of gas k
l_k : effective path length

In practice, BIRRA retrieves the gas concentration $cCH4$ by fitting synthetic spectra (generated from HITRAN-based optical depths) to the observed spectrum, using Levenberg–Marquardt optimization to minimize the residual between model and measurement.
BIRRA explicitly models:
- Instrumental spectral response
- Surface reflectance
- Atmospheric scattering and absorption

This physically grounded approach provides quantitative trace gas estimates rather than relative enhancements.

Pros
- Physically accurate: Uses full radiative transfer modeling based on the Beer–Lambert law.
- Quantitative results: Retrieves absolute CH_4 concentrations, not just relative enhancements.
- High precision: Effective at separating CH_4 absorption from background variability.
- Accounts for multiple effects: Models surface reflectance, atmospheric scattering, and instrument response.

Cons
- Computationally expensive: Significantly slower than linear or semi-empirical methods.
- Complex implementation: Requires detailed spectroscopic data and careful setup of radiative transfer parameters.
- Sensitive to model inputs: Accuracy depends on reliable prior information (e.g., pressure, temperature, surface properties).
- Parameter tuning required: Regularization, convergence thresholds, and spectral fitting windows must be carefully selected.