ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume V-3-2022
XXIV ISPRS Congress (2022 edition), 6–11 June 2022, Nice, France

AUTOMATIC METHANE PLUMES DETECTION IN TIME SERIES OF SENTINEL-5P L1B IMAGES

E. Ouerghi[⋆,*] T. Ehret[⋆] C. de Franchis[⋆,†] G. Facciolo[⋆] T. Lauvaux[‡] E. Meinhardt[⋆] J.-M. Morel[⋆]

[⋆]Université Paris-Saclay, CNRS, ENS Paris-Saclay, Centre Borelli, 91190, Gif-sur-Yvette, France
[†] Kayrros SAS
[‡] Laboratoire des Sciences du Climat et de l'Environnement, CEA, CNRS, UVSQ/IPSL

KEY WORDS: Methane, Hyperspectral, Time series, Atmospheric modeling, Anomaly detection.

ABSTRACT:

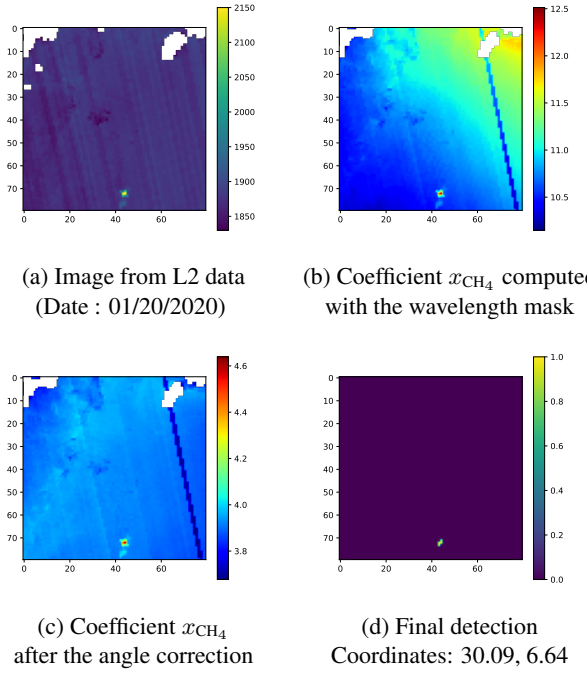
Reducing methane emissions is essential to tackle climate change. Here, we address the problem of detecting automatically large methane leaks using hyperspectral data from the Level 1B product of the Sentinel-5P satellite. To do this, two features of TROPOMI (TROPOspheric Monitoring Instrument), the Sentinel-5P satellite sensor, are exploited. The first one is the fine spectral sampling of the data which allows to isolate features of the methane absorption spectrum in the shortwave infrared wavelength range (SWIR). The second one is the daily coverage of the whole Earth which allows to perform time series analysis. Our method involves three main steps: i) a pixel reconstruction, ii) an angle correction and iii) a plume detection with a time series. In the first step, a simplified absorption model is inverted to recover, for each pixel, a coefficient representative of the presence of methane which we call the methane coefficient. In the second step, a correction is made to the methane coefficient to take into account the viewing angle of the satellite. In the third step, the obtained coefficient is normalized spatially and then the detection is carried out pixel by pixel, by looking for anomalies in a time series. We validate our method by comparing the detected plumes against a recently published dataset of plumes manually detected in the Sentinel-5P L2 methane product. We then show how our method can complement the Sentinel-5P L2 methane product for the detection of methane plumes.

1. INTRODUCTION

The detection of large methane (CH_4) leaks linked to the oil and gas industry is a major stake in order to reduce greenhouse gas (GHG) emissions. In a time lapse of 20 years, a CH_4 molecule has a global warming potential 80 times larger than carbon dioxide (CO_2) (Anthropogenic and Natural Radiative Forcing, 2013). A significant part of human CH_4 emissions could be controlled or avoided, as about 33% come from oil rigs and other oil and gas infrastructures.

In order to detect GHG fossil fuel emissions produced by human activities, several satellites have been placed in orbit around the Earth over the past ten years. Here, we focus on the data provided by TROPOMI, the main instrument of the satellite Sentinel-5P, launched in 2017 by ESA (European Space Agency). TROPOMI provides hyperspectral images in wavebands for which CH_4 has a significant absorption coefficient.

The TROPOMI data are publicly available and already being used by ESA to quantify CH_4 emissions and other greenhouse gases (Pandey et al., 2019). The approach promoted by ESA involves "exhaustive" physical modeling of the phenomenon of absorption, reflection and back-scattering by atmosphere and ground of the radiation emitted by the sun. This approach is based on the work presented in (Butz et al., 2012) and (Landgraf et al., 2016). Its modeling is extremely precise: the atmosphere is divided into several layers to take into account the variations in pressure and temperature that modify the gas absorption coefficients. This model considers not only the absorption of gases but also the absorption and diffusion of aerosols (Butz et al., 2009) (solid or liquid particles in a gaseous medium) and the back-scattering of the atmosphere. Methane quantification is then performed by inverting the model. The



(a) Image from L2 data (Date : 01/20/2020)

(b) Coefficient x_{CH_4} computed with the wavelength mask

(c) Coefficient x_{CH_4} after the angle correction

(d) Final detection Coordinates: 30.09, 6.64

Figure 1. On all images, pixels in white are discarded pixels (cloudy for example). Figure (a) shows the L2 data with a methane plume at the bottom. Figures (b) and (c) show the evolution of the coefficient x_{CH_4} estimated from L1B data at each step of the algorithm. Figure (d) shows the final detection, performed automatically.

* Corresponding author, elyes.ouerghi@gmail.com

inverse problem being ill-posed, standard regularization techniques are applied such as those presented in (Tikhonov, 1963).

In addition to the quantification aspect, the problem of detecting methane emissions belongs to the traditional field of anomaly detection in hyperspectral imagery. Usually, detection is performed on infrared wavelengths (Scafutto and Filho, 2018, Crevoisier et al., 2009). Although the absorption features of CH_4 and water vapor (H_2O) sometimes overlap in the infrared spectrum, their separation can be addressed by a wise selection of wavelengths (Crevoisier et al., 2009).

A classic anomaly detection strategy consists in performing a background subtraction – for example with a principal component analysis (PCA) (Reitberger and Sauer, 2019, Ouerghi et al., 2021) – then locally modeling the remaining residual image as stationary Gaussian. This probabilistic model then enables Neyman-Pearson tests to be carried out to detect anomalous pixels (Manolakis and Shaw, 2002, Matteoli et al., 2010).

In (Scafutto and Filho, 2018), a CH_4-specific technique is developed where the background is removed by precise atmosphere modeling. Then, CH_4 is detected by a matched filter using the CH_4 absorption spectrum. Another matched filter is used for CH_4 detection in the Cluster-Tuned Matched Filter (CTMF), originally dedicated to sulfur dioxide (Funk et al., 2001), and then applied to CH_4 detection in (Thorpe et al., 2013). The idea proposed by (Funk et al., 2001) is to use a modified version of a k-means clustering prior to the matched filter. Spatial clustering allows to obtain results over very large areas without the risk of confusing anomalies due to CH_4 with those due to other gases.

Another well-established technique in remote sensing is the band ratio technique. This method is quite usual for glacier monitoring (Citterio et al., 2010) but is also used for hydrocarbon detection (Garain et al., 2019), and is highly-relevant to our goals (Bradley et al., 2011, Varon et al., 2020, Ehret et al., 2021). Band ratios are used to enhance the spectral signature of hydrocarbons. In (Garain et al., 2019), those band ratios are used jointly with a PCA. PCA of several spectral bands is computed in order to highlight the presence of hydrocarbons in the resulting principal component image. Then, all the computed images are superimposed to perform the detection.

Many methane monitoring approaches aim at quantifying methane emissions. Our objective is different because we are interested in the detection of methane plumes on a daily basis. By covering the entire earth once a day, Sentinel-5P allows to do that. Thus, we want to introduce a flexible CH_4 emissions detection method using Sentinel-5P level 1B (L1B) data.

The level 2 (L2) product provides a quantification of methane and is already being used by private and public actors to identify and track incidents leading to large methane emissions (Lauvaux et al., 2021). However, these detections are manually identified. Moreover, the results provided by ESA on methane quantification are rather incomplete. In most L2 methane images, large regions are discarded by the quantification algorithm. Indeed, this complex algorithm relies on data from other sources and satellites which are not always available.

Our aim here is to provide an automatic methane emissions detection algorithm that will give results with far fewer missing areas. To perform this detection we start by inverting a simplified atmosphere absorption model. This allows us to obtain a coefficient representative of the presence of methane, called hereafter the methane coefficient. We then make a correction on the methane coefficient to take into account the viewing angle of the satellite. The detection of methane plumes is done automatically using a time series. The methane coefficient obtained is therefore normalized spatially to be able to compare it between different dates. For each pixel, a statistical model is then established for its time series, which allows for anomaly detection. Figure 1 illustrates the main steps of the proposed method. Statistical modeling allows us to control the number of false positives. To validate the method, we compare our detections with a recently proposed dataset (Lauvaux et al., 2021) of manually annotated plumes on the Sentinel-5P L2 methane product. We also compare our detection rate to the detection rate of the method proposed by (Ouerghi et al., 2021) which is also an automatic detection method. We then show that our method can complement the Sentinel-5P L2 methane product for the detection of methane plumes by detecting plumes that are not visible in the L2 product.

2. MATERIALS

We use hyperspectral images from Sentinel-5P. This satellite provides a dense spectrum (nearly 4,000 wavelengths) for each pixel and covers the entire earth once a day. All the available wavelengths are organized in eight wavelength bands. Each band is composed of a given number of channels. A channel corresponds to an image at a precise wavelength. Thus, each band can be seen as a hyperspectral image. Here, we use the SWIR bands 7 and 8 which cover the 2,300-2,389 nm range where the main absorption features of CH_4 are located. Together, these two bands contain nearly 1000 wavelengths. These images are part of the level 1B (L1B) product.

We also use the level 2 (L2) fully-processed data including cloud maps and X_{CH4} column mixing ratios. Cloud maps are necessary because clouds preclude any CH_4 detection. X_{CH4} images will be used to identify plumes of CH_4 only for validation purposes.

It should be noted that the L2 product provides a quantification of methane but not a plume detection. Therefore it is not a proper ground truth for plume detection. To get a methane plume mask from the L2 product, a detection must be performed on the L2 product. This detection, performed manually or automatically, can introduce its own bias in the results, with false alarms or under-detections. For example, the L2 product can yield high methane concentrations on some areas, which are unlikely to be real methane sources. However, a detection algorithm based on the L2 product might detect those areas as methane plumes.

We also use the data of the viewing zenith angle and the solar viewing angle which is provided in the L2 data. These angles will allow us to correct the methane coefficient in Section 5.

Lastly, we use detailed CH_4 and H_2O absorption spectra taken from the HITRAN spectral database (Gordon et al., 2017). Those spectra vary depending on pressure and temperature. However, here we are only interested in the CH_4 and H_2O spectra within the bottom layers of the atmosphere (below 1,500 meters above ground), because we want to detect CH_4 shortly after being emitted, before it rises and dilutes.

Hence, in practice, spectrum variations are small in the conditions of the near-surface atmospheric layers. We can see in Figure 2 the CH_4 absorption spectrum under several pressure and

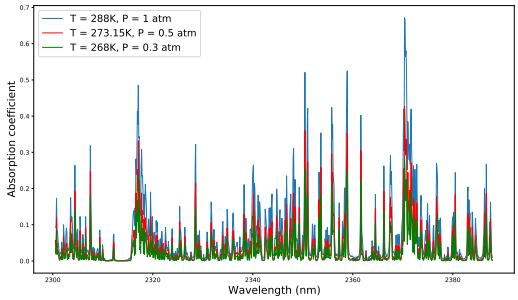


Figure 2. Comparison of methane absorption spectra under several pressure/temperature conditions.

temperature conditions. Its shape practically does not change for the near-surface conditions. Furthermore, since in our algorithm we can normalize all the spectra, we do not need to worry about changes in peak heights. Here, we selected both spectra at 15°C and 1 atm to represent near-surface atmospheric conditions.

3. MODELING

We shall use a simplified atmospheric absorption model to explain the value of each pixel of an image. Let us consider a pixel $\mathbf{P}$ in a Sentinel-5P image. Such a pixel is a vector in $\mathbb{R}^d$ where d is the number of channels in the image (for our case 960 channels from bands 7 and 8). Each pixel component corresponds to a wavelength λ.

The main assumption here is that the gases that predominate in terms of absorption in the bands 7 and 8 are methane and water vapor. Therefore, our model takes into account: the effect of sun irradiance $F_I(\lambda)$, the albedo A and the absorption coefficients of water vapor $K_{\mathrm{H_2O}}(\lambda)$ and methane $K_{\mathrm{CH_4}}(\lambda)$. We denote by e_{gas} the thickness of gas crossed by the radiation before reaching the sensor. Making implicit the dependence on λ, thanks to the Beer–Lambert law, we can write the absorption model for the whole vector $\mathbf{P}$ as

$$\mathbf{P} = A\exp(F_I - K_{\mathrm{H_2O}}e_{\mathrm{H_2O}} - K_{\mathrm{CH_4}}e_{\mathrm{CH_4}}). \quad (1)$$

From now on, we shall work with $-\log(\mathbf{P})$ instead of $\mathbf{P}$. This allows us to have a linear model where methane comes out positively. We denote by P_0 the new pixel value.

$$P_0 = -\log(A) - F_I + K_{\mathrm{H_2O}}e_{\mathrm{H_2O}} + K_{\mathrm{CH_4}}e_{\mathrm{CH_4}}. \quad (2)$$

Similarly to other works (Coakley, 2002), here we assume that the albedo is roughly constant over the part of the infrared spectrum we use, as it is extremely regular near 2000nm (Montpetit et al., 2012). The same assumption can be made for the irradiance. Indeed, the spectrum of irradiance is extremely regular and its variations are very low on the range of wavelengths that interests us (Thuillier et al., 2003), which covers barely 100nm. Therefore, we denote by A_0 the term depending on A and F_I i.e :

$$A_0 = -\log(A) - F_I. \quad (3)$$

which leads to the following formula for our model :

$$P_0 = A_0 + K_{\mathrm{H_2O}}e_{\mathrm{H_2O}} + K_{\mathrm{CH_4}}e_{\mathrm{CH_4}}. \quad (4)$$

Thus, we can see our model as the linear combination of a constant vector and the vectors $K_{\mathrm{H_2O}}$ and $K_{\mathrm{CH_4}}$, their respective coefficients being A_0, $e_{\mathrm{H_2O}}$ and $e_{\mathrm{CH_4}}$.

4. PIXEL RECONSTRUCTION

The absorption spectra of $K_{\mathrm{H_2O}}$ and $K_{\mathrm{CH_4}}$ are known. We can therefore try to invert the model we have just established by solving the following optimization problem:

$$\min_{x_{\mathrm{H_2O}},x_{\mathrm{CH_4}},x_A} \|P_0 - x_{\mathrm{H_2O}}\mathbf{K}_{\mathrm{H_2O}} - x_{\mathrm{CH_4}}\mathbf{K}_{\mathrm{CH_4}} - x_A\mathbf{1}\|, \quad (5)$$

where P_0 denotes the negative logarithm of a pixel and $\mathbf{1}$ denotes the vector for which all components are 1. This last vector is used to model the term A_0 which is still assumed to be constant over the infrared spectrum. The norm used here is the Euclidean norm. Here, we do not intend to retrieve the actual thickness of gas but a proxy which will be representative of the presence of each gas. Thus, we can normalize the water and methane spectra by their Euclidean norms and denote them by $\mathbf{K}_{\mathrm{H_2O}}$ and $\mathbf{K}_{\mathrm{CH_4}}$. This optimization problem is solved by least squares.

We can see in Figure 3 the result of this approximation. Overall the approximated pixel seems to have the same features as the original pixel. However, there are some differences in height on some of the peaks.

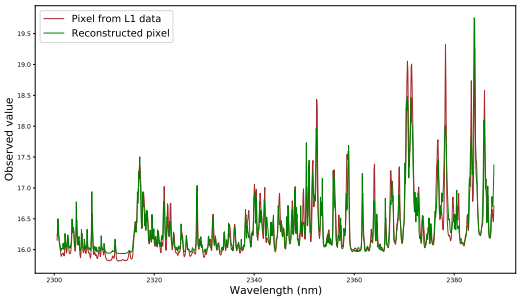


Figure 3. Comparison between the original pixel and the approximated pixel.

The largest values of the spectrum of a pixel are generally at a peak around 2370nm or 2390nm. This area corresponds to absorption peaks of water and methane. When reconstructing a pixel from these two spectra, we naturally tend to minimize the error around these peaks, at the expense of other regions of the spectrum which end up with a relatively large error of approximation. Moreover, the accuracy of the approximation at the 2370nm peak provides very little information because it is difficult to differentiate water from methane with this peak since both spectra have a peak around this wavelength. We can see how these peaks of the absorption spectra appear in the L1B data by observing the Figure 4. Therefore, in order to prevent this ill-posed situation we propose to use a sub-part of bands 7 and 8 of Sentinel-5P to perform the approximation. We isolate areas where the methane absorption peaks are disjoint from those of water vapor. In particular, we look for areas with methane absorption peaks but where the water vapor absorption coefficient is almost zero. Indeed, for a given wavelength, even if methane has a higher absorption coefficient than water vapor, water vapor can dominate in the spectrum acquired by

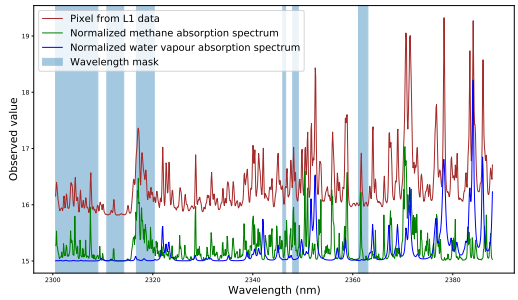


Figure 4. The highlighted areas indicate the wavelengths belonging to the wavelength mask. The absorption spectra are shifted along the y-axis for display purposes.

Sentinel-5P because usually there is much more water vapor than methane in the atmosphere. However, this phenomenon cannot occur if the absorption coefficient of water vapor is zero or almost zero. From now on, we will refer at this sub-part of bands 7 and 8 as the wavelength mask. This wavelength mask is shown in Figure 4.

In Figure 5 we can compare the original pixel and the approximated pixel when the approximation is performed using the wavelength mask as mentioned above. In this image, the approximated pixel is again very close to the original pixel. The approximation problem on the peaks is still present but its magnitude is reduced. This residual difference may be due to the fact that the spectrum of methane changes between the different layers of the atmosphere. Here we perform the approximation with only one spectrum. We can therefore miss details such as higher peaks in some methane spectra.

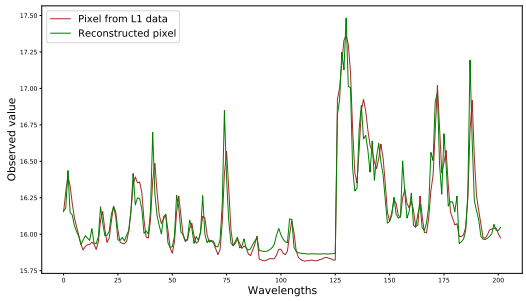


Figure 5. Comparison between the original pixel and the approximated pixel when using a wavelength mask.

The results can be compared quantitatively by calculating the relative approximation error (RAE) over different areas, as shown in Table 1. Let us denote by $P_0[mask]$ the pixel P_0 restricted to the wavelength mask and P_{rec} the reconstructed pixel, then the RAE is defined as :

$$RAE = 100 \times \frac{\|P_0[mask] - P_{rec}[mask]\|}{\|P_0[mask]\|}. \quad (6)$$

We restrict ourselves to the mask because it is the part of the spectrum which interests us most here. Moreover, this allows us to compare the approximation errors without being biased by the difference in dimension between the pixel approximated with and without the mask. We separate the different geographical areas because they contain different backgrounds. The approximation error can therefore be higher on a complex background.

Area	Mean RAE no mask	Std no mask	Mean RAE with mask	Std with mask
North Africa	0.84	0.09	0.74	0.09
South America	0.89	0.23	0.57	0.15
Middle East	0.81	0.17	0.63	0.12
Latin America	0.62	0.14	0.50	0.11

Table 1. Relative approximation error (in percentage) computed for several locations. The Mean RAE column gives the mean RAE calculated on several images of each area. The Std column gives the standard deviation of the relative error of approximation calculated on these same images.

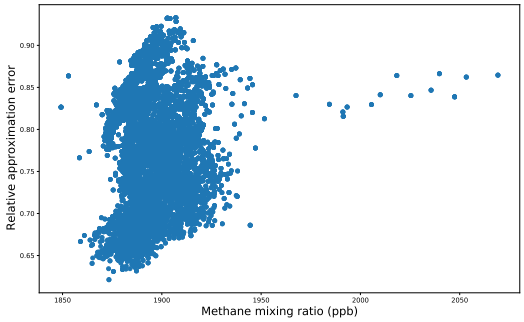


Figure 6. Relative approximation error as a function of the methane mixing ratio in particles per billion (ppb). Each dot in the figure corresponds to a pixel.

Table 1 clearly shows the effectiveness of the mask. The approximation error is always larger when not using the mask. It should be noted that the effectiveness of the mask depends on the area we consider. Indeed, we can notice for example that the gain in precision, due to the mask, is greater in South America than in North Africa.

To ensure the validity of our model in the context of methane detection, we can also verify that the approximation error is not affected by the methane rate. This is done in Figure 6 where we display the approximation error as a function of the methane mixing ratio in particles per billion (ppb). We can see that there does not seem to be any correlation between these two quantities. In particular, the points belonging to a methane plume, which are furthest to the right on the figure, do not have an abnormal error of approximation compared to the others.

5. ANGLE CORRECTION

After solving the optimization problem, the information of the presence of methane is contained in the image I_{CH_4} giving for each pixel the coefficient x_{CH_4}. In this image, we can often see a gradient effect from one side of the image to the other. This effect is due to the fact that the viewing zenith angle and solar zenith angle are different for each pixel. The solar ray that

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-V-3-2022-147-2022 | © Author(s) 2022. CC BY 4.0 License.

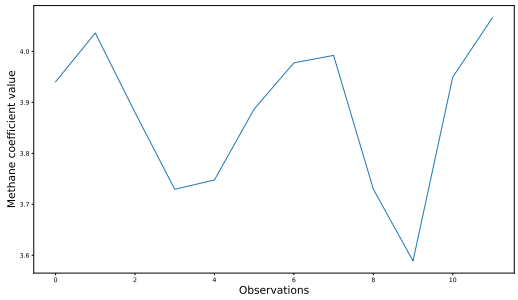


Figure 7. Value of the coefficient x'_{CH_4} on several observations of the same location. The last observation contains a methane plume.

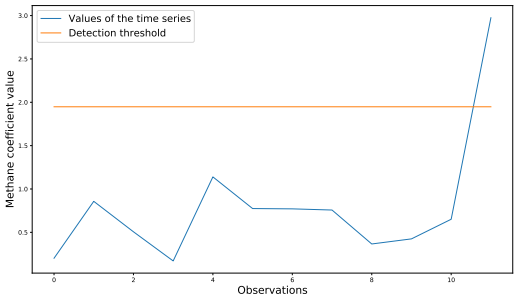


Figure 8. Value of the coefficient y_{CH_4} on several observations of the same location. The last observation contains a methane plume.

arrives to the satellite travels a longer or shorter path in the atmosphere depending on the viewing zenith angle and therefore the ray can cross a greater or lesser thickness of methane. The same effect occurs with the variations of the solar zenith angle. We correct the value of each pixel thanks to the data of these two angles which is present in the L2 data. From now on we will refer to this step as the angle correction.

Let us denote by θ_s the solar zenith angle, by θ_v the viewing zenith angle and by H the thickness of the atmosphere at nadir. Let us note

$$\begin{cases} u_s = \cos(\theta_s), \\ u_v = \cos(\theta_v). \end{cases} \tag{7}$$

When a ray from the sun reaches the ground it crosses a thickness $d_s = H/u_s$ of atmosphere and when it bounces back towards the satellite it crosses a thickness $d_v = H/u_v$ of atmosphere. In the end, the thickness of atmosphere crossed is equal to

$$d_{tot} = H\left(\frac{1}{u_s} + \frac{1}{u_v}\right) = H\left(\frac{u_s + u_v}{u_s u_v}\right). \tag{8}$$

Therefore, to remove the effect of the angles we need to compute :

$$x'_{\mathrm{CH}_4} = x_{\mathrm{CH}_4} \frac{u_s u_v}{u_s + u_v}. \tag{9}$$

6. PLUME DETECTION

To perform the plume detection, we look for anomalies in a time series. For each pixel, we build a time series of observations. For each observation we compute the coefficient x'_{CH_4} which gives us a one dimensional signal. An example of such a signal can be seen in Figure 7.

A methane plume is present on the last observation. However, in this signal we cannot distinguish the observation containing a methane plume from the others. Indeed, the methane coefficient that we calculate can vary greatly from one observation to another. To compare observations, we normalize them by calculating the average methane coefficient for each observation and the corresponding standard deviation. If we have a time series of observations, let us denote by $I'_{\mathrm{CH}_4,t}$ the image containing the coefficient $x'_{\mathrm{CH}_4,t}$ for each pixel at the instant t. This image covers an area of 400 km $\times$ 400 km. This image is much smaller than the original L1B image to try to have a uniform background. We compute

$$y_{\mathrm{CH}_4,t} = \left(x'_{\mathrm{CH}_4,t} - \mathrm{mean}(I'_{\mathrm{CH}_4,t})\right) / \mathrm{std}(I'_{\mathrm{CH}_4,t}). \tag{10}$$

This normalization sets the mean methane level at zero for each observation which makes them comparable. This step is based on the assumption that a methane plume will induce a spatial local maximum in the image. Thus, we use the presence of the spatial local maximum to obtain the local maximum temporally. We can now observe the signal given by the coefficient y_{CH_4}, which is shown in Figure 8. As expected, in this last figure, the methane plume stands out clearly and we can use this last signal to proceed with the detection.

We assume that all points in the signal $(y_{\mathrm{CH}_4,t})$ follow the same normal distribution, we compute the parameters of this distribution and then we establish a detection threshold τ which guarantees a given false alarm rate p_{fa} :

$$p_{fa}(\tau) = \mathbb{P}(P_0 > \tau | P_0 \text{ without excess } \mathrm{CH}_4). \tag{11}$$

By "P without excess CH_4", we mean that there was no detection of a methane plume on this pixel in the dataset proposed by (Lauvaux et al., 2021), which is used for the validation of the method. The value of p_{fa} is set to 10^{-6} in our experiments, which amounts statistically to less than 0.01 false alarm per image. To deduce τ from p_{fa}, we just have to invert the function given in equation (11) and the inverse is known because we are in the case of a Gaussian modeling. At least 8 observations are required in order to properly estimate the Gaussian parameters needed to perform the computation for a given pixel.

It should be noted that we look at the false positives in relation to the L2 product, which means that we can only talk about false positives for pixels where the L2 product gives a value. For the other pixels we have no information. Moreover, it should be noted that the L2 product may contain quantification errors on some pixels, therefore it is not a ground truth.

7. RESULT ANALYSIS

7.1 Algorithm analysis

To fully understand the effect of each step of the method we can look at the resulting image at each step of the algorithm as shown in Figure 1. At the bottom of Figure 1 (a) we can see a methane plume. Figure 1 (b) gives the coefficient x_{CH_4} computed with the wavelength mask. The plume stands out with a high coefficient. However, there is also an area in the top-right part of the image which presents a high coefficient x_{CH_4} and could prevent the detection of the plume (by enhancing the

mean methane level of the background) or create a false positive. This high coefficient on the top right is the manifestation of the gradient effect mentioned in Section 5. Indeed, in this image we can observe a gradient effect from the bottom left to the top right. This effect is removed on Figure 1 (c) where the angle correction has been performed. Now, the area in the top-right no longer presents high values with respect to the rest of the image. In this case the detection of the plume is possible without any false positive. In Figure 1 (d) we see the result of the final detection. It should be noted that only the part of the plume with the highest methane concentration is detected. This is explained by our high detection threshold, which guarantees a very low probability of false positives, and by the fact that the detection is performed pixel by pixel. Hence, we do not use the information that neighboring pixels are part of a plume to perform the detection.

In Figure 1, modifications affect mainly areas in the background but in practice they do not change the result of the detection. Indeed, if we had made the detection on the Figure 1 (b) we would have still detected the plume. This is not always the case as we can see in Figure 9. Figure 9 (a) shows the L2 product, with a plume visible at the center. In Figures 9 (b) and (c) the gradient effect is strong and prevents the detection. We can nevertheless note that in (c) the plume presents higher values (with respect to the background) than in (b). In Figure 9 (d) the gradient effect is removed and the contrast of the plume is strongly enhanced. Figure 9 (e) shows that the whole plume is detected, whereas in (b) and (c) the plume seemed to have holes and only the right part of the plume had values high enough to be detected.

In the example presented in Figure 10 the plume seems to be mixed up with the gradient effect. As we can see in Figure 10 (d), the angle correction still works well in this case.

7.2 Validation

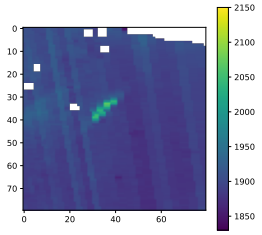
To validate our method we compare our detections to a plume dataset proposed by (Lauvaux et al., 2021). We assume that this dataset does not contain false positives because the detections were validated manually. The number of detections performed by our algorithm on this dataset is summarized in Table 2.

Location	Number of tested plumes	Number of detected plumes
North Africa	23	21
Eastern Europe	12	11
Middle East	10	8
South America	5	4
Total	50	44

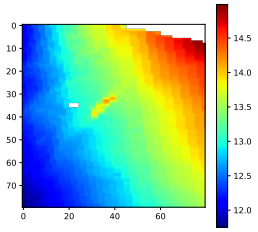
Table 2. Comparison of the number of detected plumes versus the number of tested plumes in the dataset proposed by (Lauvaux et al., 2021)

We can see that for all locations, we have a detection rate of at least 80%. In particular, the average detection rate is 88% which exceeds the detection rate of the automatic detection method proposed by (Ouerghi et al., 2021), which is 74%.

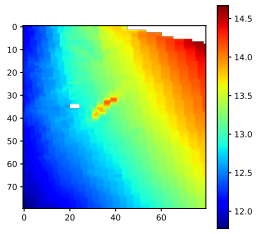
The non-detections can be due to two possible factors. The detection model can be incorrect When not enough observations are available in the time series. Another explanation is that the



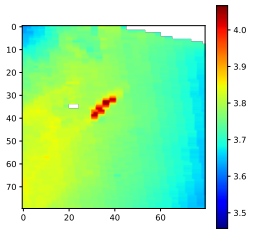
(a) Image from L2 data (Date : 01/08/2020)



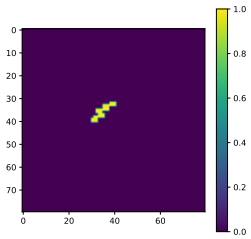
(b) Coefficient x_{CH_4} computed with all wavelengths



(c) Coefficient x_{CH_4} computed with a subset of all wavelengths



(d) Coefficient x_{CH_4} after the angle correction



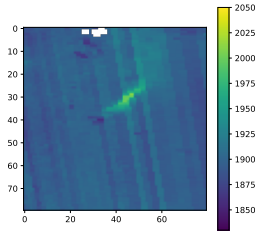
(e) Final detection
Coordinates : 31.89, 6.19

Figure 9. On all images, pixels in white are discarded pixels (cloudy for example). In (a), the L2 data with a methane plume in the middle of the image. In the other figures, the evolution of the coefficient x_{CH_4} at each step of the algorithm.

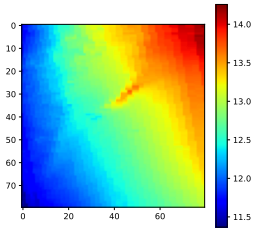
plume stood out on the time series but not enough to exceed the detection threshold. Indeed, we impose here a false positive probability of 10^{-6} per pixel, which gives a high detection threshold. Thus, it is possible that some plumes do not exceed this detection threshold. They could be detected with a lower threshold but this would lead to more false positives. Unfortunately, in the absence of sufficient ground truth data, we cannot produce a ROC curve. In fact we only dispose of a number of plumes detected on the L2 product, but another expertise will be necessary to decide if our additional detections are valid. Indeed, as explained in the next section, it is possible to confirm some of our additional detections by using, for example, wind data, but it is not always the case. In short, we evaluate here a false negative rate, but so far we do not have enough tools to evaluate an empirical false positive rate for our new detections.

7.3 Detection of additional plumes

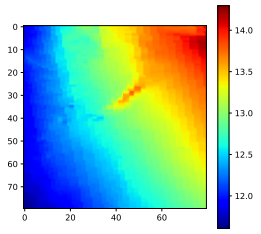
Our gain over the L2 data, besides the fact that the plume detection is automatic, is that our algorithm has fewer missing pixels and still allows to detect the same plumes as if we performed the detection on the L2 data. The fact that we have fewer missing pixels than the L2 data allows us to detect even more plumes



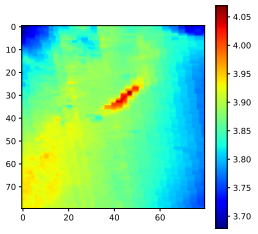
(a) Image from L2 data
(Date : 01/03/2020)



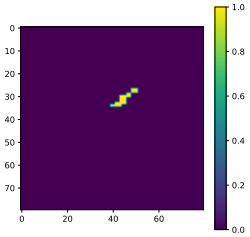
(b) Coefficient x_{CH_4} computed with all wavelengths



(c) Coefficient x_{CH_4} computed with a subset of all wavelengths

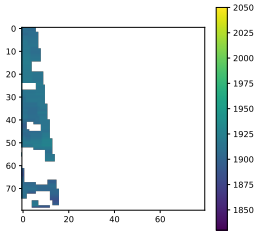


(d) Coefficient x_{CH_4} after the angle correction

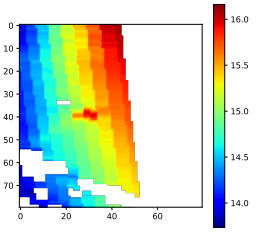


(e) Final detection
Coordinates: 31.89, 6.33

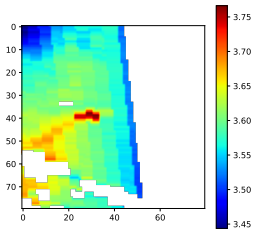
Figure 10. On all images, pixels in white are discarded pixels (cloudy for example). In (a), the L2 data with a methane plume in the middle of the image. On the other figures, the evolution of the coefficient x_{CH_4} at each step of the algorithm.



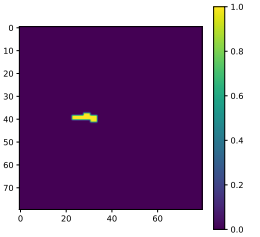
(a) Image from L2 data
(Date : 01/07/2020)



(b) Coefficient x_{CH_4} computed with the wavelength mask



(c) Coefficient x_{CH_4} after the angle correction



(d) Final detection
Coordinates: 31.78, 6.12

Figure 11. There is nearly no valid pixel available in the L2 data, as shown in Figure (a). However, by using the L1B data we are able to detect a plume, as shown in Figure (d).

than if we used the L2 data. We can perform a detection where there are no L2 data as shown in Figure 11.

To confirm the detection of additional plumes, several pieces of information are used. If the wind is strong we check that the plume is aligned with the wind. In addition, we can verify that the area where the plume is emitted is an area where methane emissions are expected. For example, in the case of Figure 11, the area observed is around a hydrocarbon extraction installation where methane emissions are regularly detected. Moreover, the high detection rate and the low false alarm probability of the method bring confidence about the existence of the plume.

However, this information is not always sufficient to draw conclusions about the existence of a plume. For example when the wind is weak or the plume is concentrated on very few pixels, it is hard to use the alignment of the plume with the wind. Similarly, we do not always have information on the presence of oil extraction installations. In all these situations, we do not have the means to determine if we observe a plume or a false positive.

8. CONCLUSION

We have shown that our method based on the L1B product of Sentinel-5P can detect methane plumes automatically with an accuracy higher than 80%. Moreover, the statistical modeling leading to the detection keeps the false positive rate very low with a false positive probability of 10^{-6} per pixel with respect to the L2 product. Our method does not detect all the plumes detected by (Lauvaux et al., 2021); however, unlike the method proposed by (Lauvaux et al., 2021), our method detects plumes automatically. Moreover, our method allows to detect plumes that are not visible in the L2 product. It thus complements a detection algorithm based on the L2 product alone.

However, there is still room for improvement. For example, more objective and quantitative criteria can be found to determine the wavelength mask which is currently obtained manually. Moreover, our algorithm, like the L2 product, does not allow to perform detections over water because the albedo is extremely low.

ACKNOWLEDGEMENTS

Work partly financed by Office of Naval research grant N00014-17-1-2552, MENRT, and a PhD scholarship financed by MESRI (Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation). T. Lauvaux was supported by the French research program «Make Our Planet Great Again » (CNRS, project CIUDAD).

REFERENCES

Anthropogenic and Natural Radiative Forcing, 2013. Intergovernmental Panel on Climate Change (ed.), *Climate Change 2013 - The Physical Science Basis*, Cambridge University Press, Cambridge, 659–740.

Bradley, E. S., Leifer, I., Roberts, D. A., Dennison, P. E., Washburn, L., 2011. Detection of marine methane emissions with AVIRIS band ratios. *Geophysical Research Letters*, 38(10).

Butz, A., Galli, A., Hasekamp, O., Landgraf, J., Tol, P., Aben, E., 2012. TROPOMI aboard Sentinel-5 Precursor: Prospective performance of CH4 retrievals for aerosol and cirrus loaded atmospheres. *Remote Sens Environ*, 120(SI), 267–276.

Butz, A., Hasekamp, O., Frankenberg, C., Aben, I., 2009. Retrievals of atmospheric CO2 from simulated space-borne measurements of backscattered near-infrared sunlight: accounting for aerosol effects. *Applied optics*, 48, 3322-36.

Citterio, M., Paul, F. et al., 2010. Remote sensing of glacier change in West Greenland: Accounting for the occurrence of surge-type glaciers. *Ann Glaciol*, 50, 70-80.

Coakley, J., 2002. Reflectance and albedo. *Encyclopedia of the Atmosphere*, Academic Press, 1914–1923.

Crevoisier, C., Nobileau, D., Fiore, A. M., Armante, R., Chédin, A., Scott, N., 2009. Tropospheric methane in the tropics–first year from IASI hyperspectral infrared observations. *Atmospheric Chemistry and Physics*, 9(17), 6337–6350.

Ehret, T., Truchis, A. D., Mazzolini, M., Morel, J.-M., d'Aspremont, A., Lauvaux, T., Duren, R., Cusworth, D., Facciolo, G., 2021. Global Tracking and Quantification of Oil and Gas Methane Emissions from Recurrent Sentinel-2 Imagery. *CoRR*, abs/2110.11832. http://arxiv.org/abs/2110.11832.

Funk, C., James, T., Roberts, D., Borel, C., 2001. Clustering to Improve Matched Filter Detection of Weak Gas Plumes in Hyperspectral Thermal Imagery. *Geoscience and Remote Sensing, IEEE Transactions on*, 39, 1410 - 1420.

Garain, S., Mitra, D., Das, P., 2019. Detection of hydrocarbon microseepage-induced anomalies by spectral enhancements of Landsat 7 ETM+ images in part of Assam–Arakan Fold Belt, India. *J Pet Explor Prod Technol*, 9.

Gordon, I. E., Rothman, L. S., Hill, C., Kochanov, R. V., Tan, Y. et al., 2017. The HITRAN2016 molecular spectroscopic database. *J Quant Spectrosc Ra*, 203, 3 - 69. https://hal.archives-ouvertes.fr/hal-01765945.

Landgraf, J., Brugh, J. et al., 2016. Carbon monoxide total column retrievals from TROPOMI shortwave infrared measurements. *Atmos Meas Tech*, 9, 4955-4975.

Lauvaux, T., Giron, C., Mazzolini, M., D'Aspremont, A., Duren, R., Cusworth, D., Shindell, D., Ciais, P., 2021. Global Assessment of Oil and Gas Methane Ultra-Emitters. *CoRR*, abs/2105.06387.

Manolakis, D., Shaw, G., 2002. Detection algorithms for hyperspectral Imaging applications. *Signal Processing Magazine, IEEE*, 19, 29 - 43.

Matteoli, S., Diani, M., Corsini, G., 2010. A tutorial overview of anomaly detection in hyperspectral images. *IEEE Aero El Sys Mag*, 25, 5 - 28.

Montpetit, B., Royer, A. et al., 2012. New shortwave infrared albedo measurements for snow specific surface area retrieval. *Journal of Glaciology*, 58, 941-952.

Ouerghi, E., Ehret, T., de Franchis, C., Facciolo, G., Lauvaux, T., Meinhardt, E., Morel, J.-M., 2021. Detection of methane plumes in hyperspectral images from Sentinel-5P by coupling anomaly detection and pattern recognition. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, V-3-2021, 81–87. https://www.isprs-ann-photogramm-remote-sens-spatial-inf-sci.net/V-3-2021/81/2021/.

Pandey, S. et al., 2019. Satellite observations reveal extreme methane leakage from a natural gas well blowout. *PNAS*, 116(52), 26376–26381.

Reitberger, G., Sauer, T., 2019. Background Subtraction using Adaptive Singular Value Decomposition. *CoRR*, abs/1906.12064. http://arxiv.org/abs/1906.12064.

Scafutto, R., Filho, C. D. S., 2018. Detection of Methane Plumes Using Airborne Midwave Infrared Hyperspectral Data. *Remote Sensing*, 10(8). https://www.mdpi.com/2072-4292/10/8/1237.

Thorpe, A., Roberts, D., Bradley, E., Funk, C., Dennison, P., Leifer, I., 2013. High resolution mapping of methane emissions from marine and terrestrial sources using a Cluster-Tuned Matched Filter technique and imaging spectrometry. *Remote Sens Environ*, 134, 305 - 318.

Thuillier, G., Hersé, M., Labs, D., Foujols, T., Peetermans, W., Gillotay, D., Simon, P. C., Mandel, H., 2003. The Solar Spectral Irradiance from 200 to 2400 nm as Measured by the SOLSPEC Spectrometer from the Atlas and Eureca Missions. *Solar Physics*, 214(1), 1-22.

Tikhonov, N., 1963. A., 1943, Solution of incorrectly formulated problems and the regularization method. *Doklady akademii nauk sssr*, 501–504.

Varon, D. J., Jervis, D., McKeever, J., Spence, I., Gains, D., Jacob, D. J., 2020. High-frequency monitoring of anomalous methane point sources with multispectral Sentinel-2 satellite observations. *Atmospheric Measurement Techniques Discussions*, 2020, 1–21. https://amt.copernicus.org/preprints/amt-2020-477/.