Supplementary Information for:

Automatic Methane Plume Masking Based on Wavelet Transform Image Processing: Application to MethaneAIR and MethaneSAT Data

Contents

1 Supplemental Methods

1.1 Concentration filtering

In the concentration filter, we define a concentration "hotspot" as a XCH_4 clump with elevated pixel values within the plume mask. This requires us to set a pixel value threshold and a size threshold. The way we determine the pixel value threshold is described below.

We first calculate a base value from a percentile threshold across the entire scene. Then for each mask, we compute two standard deviations: the overall scene standard deviation (σ_{bg}) and the standard deviation of a subregion covering the mask (σ_{sub}). If these two numbers are close, the pixel value threshold for this mask is the base value. Otherwise, the subregion is likely to be noisier than the entire scene, thus the pixel value threshold is raised by an increment to avoid including noise pixels. The increment is proportional to the difference between two standard deviations.

For the percentile threshold that decides the base value, we choose between two candidates: 99.7% and 99.97%. the former is used when the scene is clean (low mean and standard deviation) or when strong emissions exist (high concentration value at 99.9% percentile). When the scene is noisier (high mean and standard deviation) we raise the threshold to the latter to avoid including false detections from background noise.

For the size threshold, we apply a constant value for all masks across the scene ((10 pixels for MethaneAIR, 20 for MethaneSAT). Note that one mask can contain multiple hotspots.

1.2 Sensitivity analysis

For the parameters that need to be empirically determined, we provide a detailed sensitivity analysis to show how different values affect detection results. For MethaneAIR, these parameters include: the scaling factor to determine the pre-processing pixel value threshold (Table S1), the decomposition level factor S2), the scaling factor to determine the plume masking pixel value threshold (Table S3), the plume mask size threshold (Table S4), the plume hotspot size threshold (Table S5), the angle buffer in wind direction filtering (Table S6), and the fiber length ratio in shape filtering (Table S7). For MethaneSAT, we include the same seven parameters above, plus the local background width to determine the pre-processing pixel value threshold (Table S11), and the local background width to determine the plume masking pixel value threshold (Table S13).

Here the scaling factor is used to determine thresholds for denoising or masking, where the threshold used is the mean XCH_4 value of the scene plus the scaling factor times the standard deviation of XCH_4 for the scene. The decomposition level factor is used to determine how many levels of decomposition (L) is conducted in Equation (1), where L is the logarithm (base 2) of the smaller image dimension multiplied the decompostion level factor.

For each parameter, we independently choose three example scenes that have numerous plumes and represent varying surface features and local wind characteristics for MethaneAIR and MethaneSAT, respectively: Scene 1 represents the Permian basin (high albedo) with high wind speed, Scene 2 represents the Appalachian basin (low albedo), and Scene 3 represents the Permian basin with low wind speed.

1.3 Computation efficiency comparison

The computation efficiency comparison is used to evaluate computational efficiency across the wavelet method, the DI thresholding method, and a representative machine learning (ML)–based approach. The comparison focuses on two metrics: total runtime and peak memory usage for processing a single scene. All methods were applied to the same set of MethaneAIR and

MethaneSAT observations to ensure consistency. For each method, total runtime is defined as the wall-clock time required to process a single scene from input data loading to final plume detection output. Peak memory usage is defined as the maximum resident memory consumed during execution. This was tracked using system-level profiling tools that record memory allocation throughout the processing pipeline.

The ML-based method used for comparison is an instance segmentation framework designed to detect methane plumes from XCH_4 concentration maps. The model is based on Mask R-CNN with a ResNet-50 backbone, which has been widely used for object detection and segmentation tasks. To address the limited availability of labeled MethaneSAT data, the model is initialized using weights pre-trained on MethaneAIR observations and then fine-tuned for plume detection. Plume labels are from the wavelet method.

Note that the reported runtime and memory usage reflect the application of fully configured methods and do not include parameter tuning or model development. For the wavelet and DI methods, parameter tuning typically involves testing a limited number of parameter combinations and results in relatively low computational cost. In contrast, many ML approaches require extensive hyperparameter optimization and training iterations, which can significantly increase total computational demand.

2 Supplemental Figures

3 Supplemental Tables

References

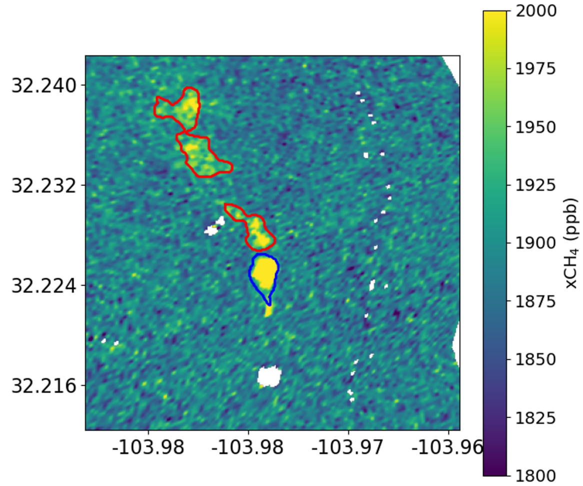

Figure S1: A MethaneAIR example of a plume dissected into multiple detached clumps by eddies. In this case, the plume mask closest to the source (blue) is considered true and the others downwind (red) are considered false.

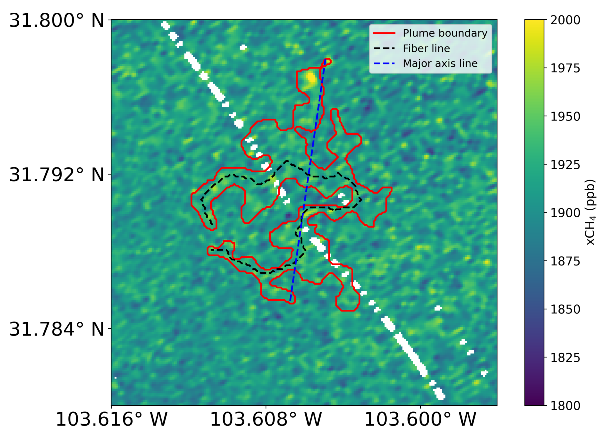

Figure S2: A MethaneAIR false plume example with the “spider” shape plume mask. Such a mask has its fiber line (black) much longer than its major axis line (blue).

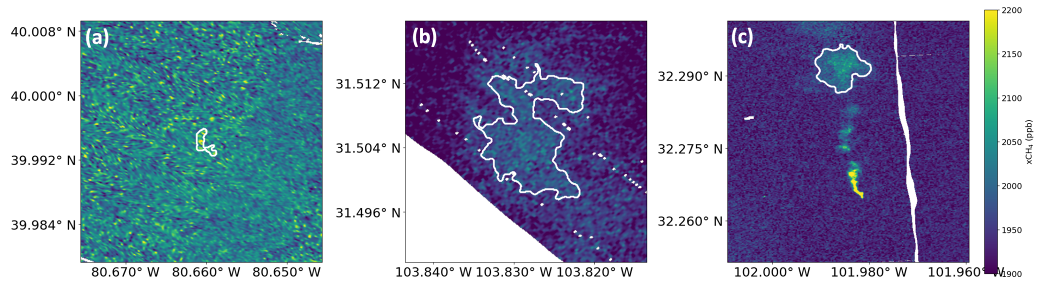

Figure S3: Examples of false positive detections from MethaneAIR. (a) Background interference, where the standard deviation of the subscene is 6 ppb higher than that of the full scene. (b) Diffuse enhancements likely arising from dispersed small emissions below the detection limit. (c) Diffuse enhancements that may represent fragmented plumes caused by atmospheric eddies, or emissions from intermittent sources.

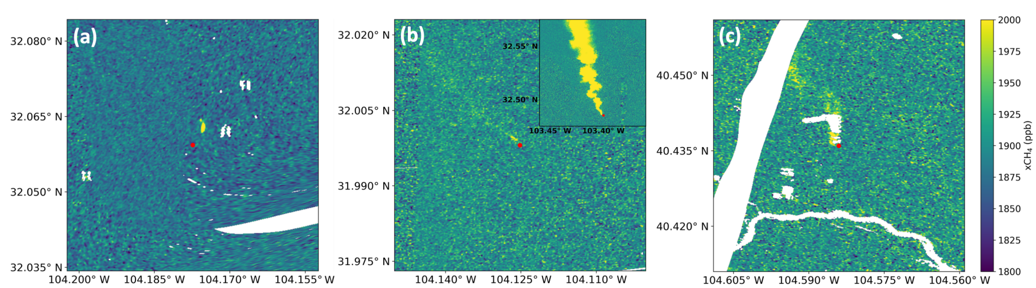

Figure S4: Examples of false negative detections from MethaneAIR. Red dots indicate plume origins identified by the DI thresholding method. (a) Small plume size, with a limited number of elevated enahcnement pixels. (b) Low-flux emissions coexisting with a larger plume. (c) Additional screening to the concentration map removes plume pixels, increasing the difficulty of detection.

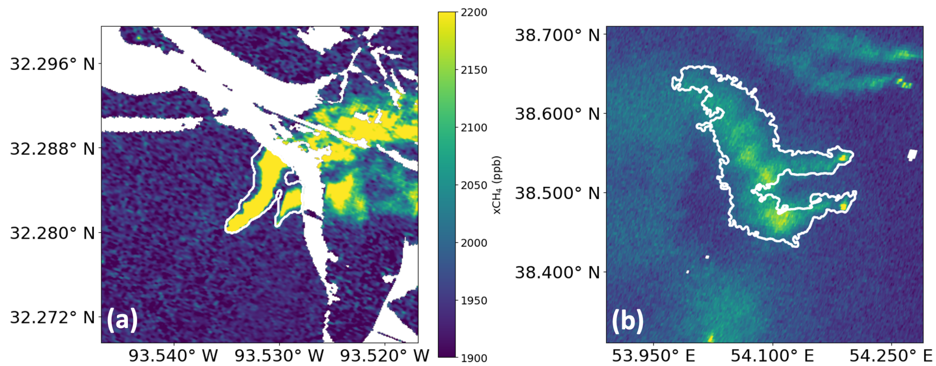

Figure S5: Examples of plumes merging, where multiple plumes are grouped within a single mask, shown for (a) MethaneAIR and (b) MethaneSAT.

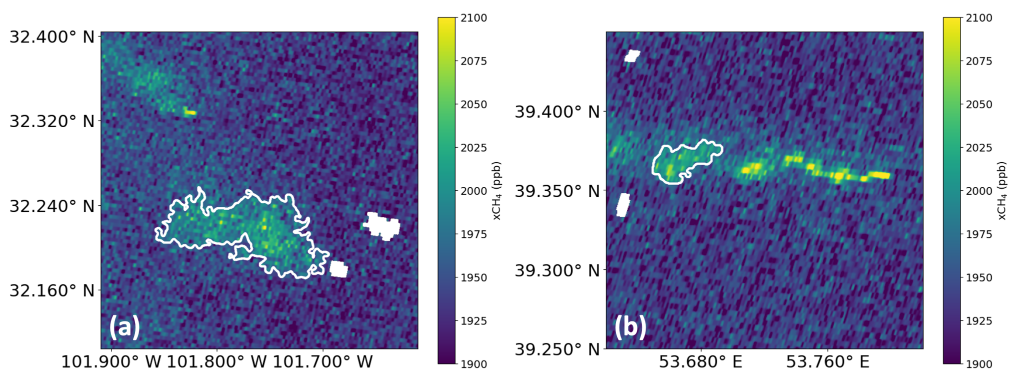

Figure S6: Examples of false positive detections from MethaneSAT. (a) Diffuse enhancements likely arising from dispersed small emissions below the detection limit. (b) Diffuse enhancements that may represent fragmented plumes caused by atmospheric eddies.

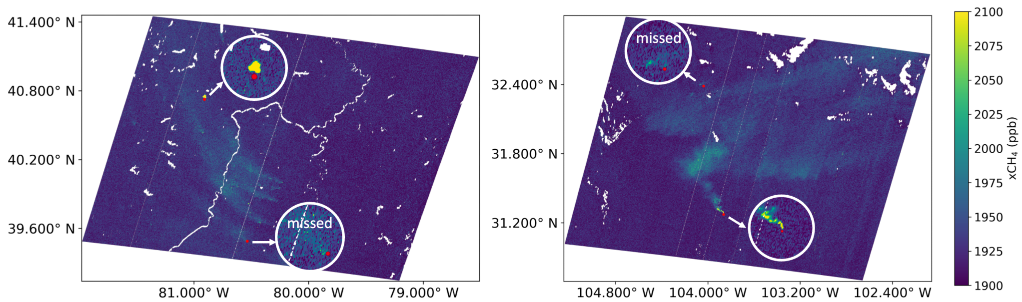

Figure S7: Two false negative detections from MethaneSAT. Both missed plumes coexist with a large plume in the same scene, where the elevated concentration threshold caused by the larger plume filters out weaker enhancements.

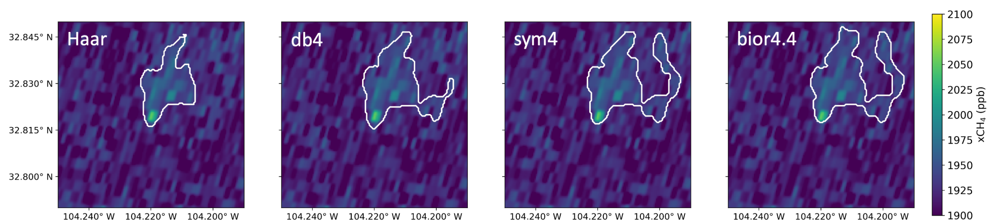

Figure S8: Plume masks generated from a plume example using different types of wavelets.

	Pre-processing scaling factor – MethaneAIR					
	Scene 1		Scene 2		Scene 3	
Value	# True	# False	# True	# False	# True	# False
1.0	54	47	41	20	15	5
1.5	53	33	40	6	13	1
2.0	56	21	37	4	13	1
2.5	49	18	37	3	7	1

Table S1: Sensitivity analysis for the pre-processing scaling factor in three MethaneAIR flights with varying surface features and wind speeds. The value 2.0 (bold) is selected.

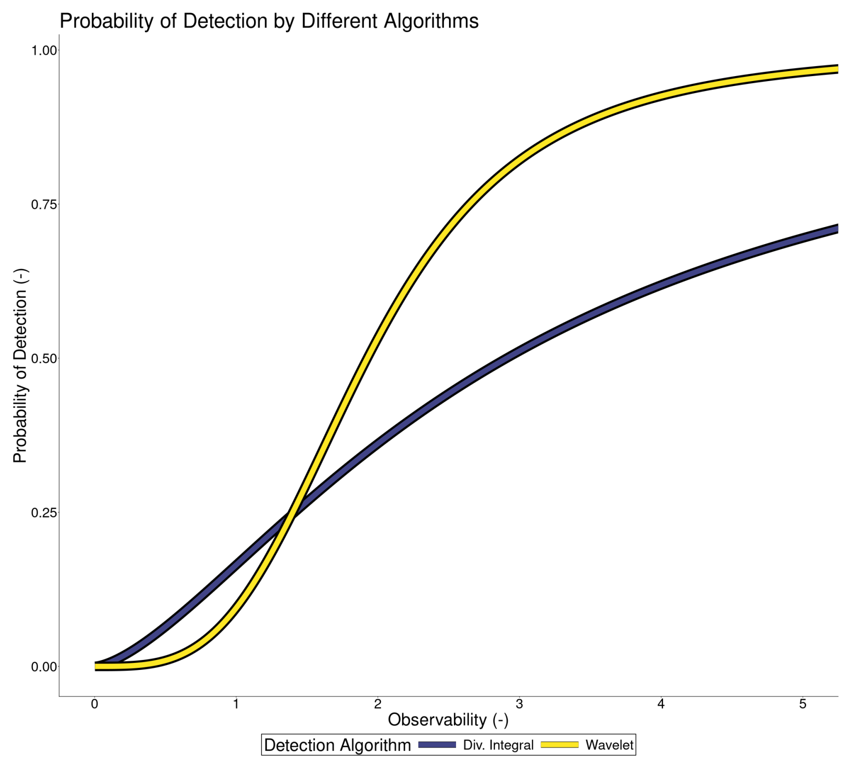

Figure S9: Predicted probability of plume detection (y-axis) as a function of observability (x-axis) by the Divergence Integral (blue curve) and Wavelet (yellow curve) methods, calculated for MethaneAIR and MethaneSAT. Here observability = $log(\frac{q}{\sqrt{a}nu})$, q = emission rate per unit area, a = pixel area, n = noise amplitude, u = windspeed. The wavelet method initially shows lower P_d than the DI thresholding at low observability (0 to 1.3). This is likely because the wavelet method relies on multi-scale feature extraction and thresholding, which can suppress weak plumes near the detection limit. However, the wavelet method surpasses DI starting at observability of ∼1.3, with its advantage growing as observability increases and peaking at observability of ∼2.5. This reflects the wavelet method's stronger ability to enhance plume structures and reject noise, therefore leads to more reliable detections. At higher observability levels, the difference in P_d gradually decreases but remains positive, as both methods approach near-perfect detection.

	Decomposition level factor – MethaneAIR					
	Scene 1		Scene 2		Scene 3	
Value	# True	# False	# True	# False	# True	# False
1/8	52	36	35	26	13	10
3/8	52	29	35	16	13	1
1/2	56	21	37	4	13	1
5/8	47	11	31	3	13	3
3/4	40	11	32	2	11	3
1	36	7	25	3	8	1

Table S2: Sensitivity analysis for the decomposition level factor in three MethaneAIR flights with varying surface features and wind speeds. The value 1/2 (bold) is selected.

	Plume masking scaling factor – MethaneAIR					
	Scene 1		Scene 2		Scene 3	
Value	# True	# False	# True	# False	# True	# False
1.0	53	50	40	10	11	10
1.5	56	21	37	4	13	1
2.0	50	19	36	2	12	0
2.5	43	9	33	2	12	0

Table S3: Sensitivity analysis for the plume masking scaling factor in three MethaneAIR flights with varying surface features and wind speeds. The value 1.5 (bold) is selected.

	Plume mask size – MethaneAIR					
	Scene 1		Scene 2		Scene 3	
Value	# True	# False	# True	# False	# True	# False
50	58	32	41	17	15	10
100	56	21	37	4	13	1
150	54	20	37	4	13	1
200	54	17	37	4	13	1

Table S4: Sensitivity analysis for plume mask size in three MethaneAIR flights with varying surface features and wind speeds. The value 100 (bold) is selected.

	Plume hotspot size – MethaneAIR					
	Scene 1		Scene 2		Scene 3	
Value	# True	# False	# True	# False	# True	# False
5	56	35	37	20	13	8
10	56	21	37	4	13	1
15	51	18	37	4	13	1
20	49	13	34	4	13	1

Table S5: Sensitivity analysis for plume hotspot size in three MethaneAIR flights with varying surface features and wind speeds. The value 10 (bold) is selected.

	Wind direction filter angle buffer – MethaneAIR					
	Scene 1		Scene 2		Scene 3	
Value (°)	# True	# False	# True	# False	# True	# False
45	51	18	36	4	13	0
50	55	21	37	4	13	0
55	56	21	37	4	13	1
60	56	29	39	9	13	2

Table S6: Sensitivity analysis for wind-direction filter angle buffer in three MethaneAIR flights with varying surface features and wind speeds. The value 55° (bold) is selected.

	Fiber length ratio – MethaneAIR					
	Scene 1		Scene 2		Scene 3	
Value	# True	# False	# True	# False	# True	# False
1.00	41	17	26	4	11	1
1.25	56	21	37	4	13	1
1.50	56	26	37	8	13	4
1.75	46	28	37	11	13	4

Table S7: Sensitivity analysis for the fiber length ratio in three MethaneAIR flights with varying surface features and wind speeds. The value 1.25 (bold) is selected.

	Comparison with alternative methods (MethaneAIR)					
	Scene 1		Scene 2		Scene 3	
Method	# True	# False	# True	# False	# True	# False
Wavelet	56	21	37	4	13	1
DI thresholding	37	28	21	4	10	8
Gaussian filtering	38	29	24	14	11	8
Median filtering	37	30	24	14	8	9

Table S8: Comparison of the wavelet method with alternative approaches across three MethaneAIR scenes. The wavelet method yields higher numbers of true detections and comparable or lower numbers of false detections.

	Comparison with alternative wavelets (MethaneAIR)					
	Scene 1		Scene 2		Scene 3	
Wavelet	# True	# False	# True	# False	# True	# False
Haar	56	21	37	4	13	1
db4	57	23	37	7	13	1
sym4	55	24	37	6	12	1
bior4.4	57	26	36	7	11	0

Table S9: Comparison of Haar wavelet with alternative wavelets across three MethaneAIR scenes. No significant differences are observed between Haar and the other wavelets in terms of true and false detections.

	Pre-processing scaling factor (MethaneSAT)					
	Scene 1		Scene 2		Scene 3	
Value	# True	# False	# True	# False	# True	# False
1.25	8	7	12	16	1	4
1.50	8	3	12	10	1	2
1.75	7	2	12	8	1	0
2.00	2	0	7	3	1	0

Table S10: Sensitivity analysis of the pre-processing scaling factor in three MethaneSAT scenes with varying surface features and wind speeds. The value 1.75 (bold) is selected.

	Pre-processing local background width (MethaneSAT)					
	Scene 1		Scene 2		Scene 3	
Value (m)	# True	# False	# True	# False	# True	# False
2250	4	2	12	8	1	0
4500	7	2	12	8	1	0
6750	7	3	10	12	1	0
9000	5	1	11	7	1	0

Table S11: Sensitivity analysis of the pre-processing local background width in three MethaneSAT scenes with varying surface features and wind speeds. The value 4500 m (bold) is selected.

	Decomposition level factor (MethaneSAT)					
	Scene 1		Scene 2		Scene 3	
Value	# True	# False	# True	# False	# True	# False
1/8	7	105	12	23	1	10
3/8	7	54	12	15	1	4
1/2	7	2	12	8	1	0
5/8	3	0	11	6	1	0
3/4	0	0	11	6	0	0
1	0	0	10	2	0	0

Table S12: Sensitivity analysis of the decomposition level factor in three MethaneSAT scenes with varying surface features and wind speeds. The value 1/2 (bold) is selected.

	Plume masking scaling factor (MethaneSAT)					
	Scene 1		Scene 2		Scene 3	
Value	# True	# False	# True	# False	# True	# False
1.25	8	6	12	12	1	5
1.50	7	3	12	11	1	0
1.75	7	2	12	8	1	0
2.00	4	0	7	3	1	0

Table S13: Sensitivity analysis of the plume masking scaling factor in three MethaneSAT scenes with varying surface features and wind speeds. The value 1.75 (bold) is selected.

	Plume masking local background width (MethaneSAT)					
	Scene 1		Scene 2		Scene 3	
Value (m)	# True	# False	# True	# False	# True	# False
2250	6	2	12	8	1	0
4500	7	2	12	8	1	0
6750	7	2	12	8	1	0
9000	5	0	10	5	1	0

Table S14: Sensitivity analysis of the plume masking local background width in three MethaneSAT scenes with varying surface features and wind speeds. The value 4500 m (bold) is selected.

	Plume mask size (MethaneSAT)					
	Scene 1		Scene 2		Scene 3	
Value	# True	# False	# True	# False	# True	# False
250	8	15	12	14	1	3
500	7	2	12	8	1	0
750	5	2	11	6	1	0
1000	3	2	11	4	1	0

Table S15: Sensitivity analysis of plume mask size in three MethaneSAT scenes with varying surface features and wind speeds. The value 500 (bold) is selected.

	Plume hotspot size (MethaneSAT)					
	Scene 1		Scene 2		Scene 3	
Value	# True	# False	# True	# False	# True	# False
10	7	2	12	11	1	0
20	7	2	12	8	1	0
30	7	2	11	8	1	0
40	7	2	11	6	1	0

Table S16: Sensitivity analysis of plume hotspot size in three MethaneSAT scenes with varying surface features and wind speeds. The value 20 (bold) is selected.

	Wind direction filter angle buffer (MethaneSAT)					
	Scene 1		Scene 2		Scene 3	
Value (°)	# True	# False	# True	# False	# True	# False
45	7	2	12	8	1	0
50	7	2	12	8	1	0
55	7	2	12	8	1	0
60	7	2	12	9	1	0

Table S17: Sensitivity analysis of the wind direction filter angle buffer in three MethaneSAT scenes with varying surface features and wind speeds. The value 55° (bold) is selected.

	Fiber length ratio (MethaneSAT)					
	Scene 1		Scene 2		Scene 3	
Value	# True	# False	# True	# False	# True	# False
1.00	4	2	3	4	0	0
1.25	7	2	12	8	1	0
1.50	7	4	12	8	1	0
1.75	7	4	12	10	1	0

Table S18: Sensitivity analysis of the fiber length ratio in three MethaneSAT scenes with varying surface features and wind speeds. The value 1.25 (bold) is selected.

	Comparison with alternative methods (MethaneSAT)					
	Scene 1		Scene 2		Scene 3	
Method	# True	# False	# True	# False	# True	# False
Wavelet	7	2	12	8	1	0
DI thresholding	2	1	7	13	0	0
Gaussian filtering	5	12	7	3	0	0
Median filtering	5	15	6	3	0	0

Table S19: Comparison of the wavelet method with alternative approaches across three MethaneSAT scenes. The wavelet method yields higher numbers of true detections and comparable or lower numbers of false detections.

	Comparison with alternative wavelets (MethaneSAT)					
	Scene 1		Scene 2		Scene 3	
Wavelet	# True	# False	# True	# False	# True	# False
Haar	7	2	12	8	1	0
db4	7	6	12	8	1	0
sym4	7	3	12	7	1	0
bior4.4	6	2	12	7	1	0

Table S20: Comparison of Haar wavelet with alternative wavelets across three MethaneSAT scenes. No significant differences are observed between Haar and the other wavelets in terms of true and false detections.

	Wavelet		DI thresholding		ML		
	Time (hr)	Memory (GB)	Time (hr)	Memory (GB)	Test time (hr)	Test memory (GB)	Train time (hr)
MethaneAIR	1.06	25	0.18	15	<0.1	1	29
MethaneSAT	0.16	23	0.14	29	<0.1	1	15

Table S21: Computation efficiency comparison between the wavelet method, DI thresholding method and a representative ML approach.