



S2MetNet: A novel dataset and deep learning benchmark for methane point source quantification using Sentinel-2 satellite imagery

Ali Radman ^a, Masoud Mahdianpari ^{a,b,*}, Daniel J. Varon ^c, Fariba Mohammadimanesh ^b

^a Department of Electrical and Computer Engineering, Memorial University of Newfoundland, St. John's, NL A1B 3X5, Canada

^b C-CORE, St. John's, NL A1B 3X5, Canada

^c School of Engineering and Applied Sciences, Harvard University, Cambridge, United States

ARTICLE INFO

Edited by Menghua Wang

Keywords:

Methane
Emission monitoring
Machine learning
Remote sensing
Quantification
Sentinel-2
Large Eddy simulation (LES)

ABSTRACT

Methane, as a crucial greenhouse gas, plays a significant role in global warming, contributing to approximately one-quarter of the observed climate change since pre-industrial times. Consequently, the detection and quantification of major methane emitters are vital in addressing this issue effectively. Satellite sensors with shortwave infrared (SWIR) spectral bands provide valuable information for monitoring methane emissions. For example, Sentinel-2 multispectral data have the capability to detect methane plumes of large point sources. As such, a wide range of quantification approaches have been developed to quantify methane source rates based on this dataset. Most of the existing methods, however, require auxiliary data, such as wind speed, and have large uncertainties. In this study, we introduce a novel approach based on deep learning models to enhance the precision of methane quantification using Sentinel-2 data without the reliance on external data sources. To train the proposed deep learning model, a comprehensive benchmark dataset has been generated, using Sentinel-2 data. This dataset is created by integrating simulated plumes and background noise extracted from real Sentinel-2 images. This approach ensures the integration of realistic environmental conditions within the simulated data, enhancing the robustness and reliability of our proposed model. The generated benchmark dataset is utilized in different deep learning architectures, namely VGG-19, ResNet-50, Inception-v3, DenseNet-121, Swin-T, and EfficientNet-V2L, to estimate methane source rate. The performance of deep models has been evaluated in three learning strategies, namely from scratch, transfer-learning, and fine-tuning. The fine-tuned EfficientNet-V2L achieves the highest accuracy with root-mean-square error (RMSE), mean absolute percentage error (MAPE), and Pearson R of 2101 kg h⁻¹, 10.05%, and 95.70%, respectively. More importantly, the proposed model demonstrates superior performance compared to conventional physical-based quantification methods (e.g., integrated mass enhancement) and recently developed deep learning model techniques (e.g., MethaNet). In particular, the proposed model exhibits an improvement of approximately 1287 kg h⁻¹ in terms of RMSE, a 3.92% reduction in MAPE, and a 5.01% enhancement in R compared to the IME method. These results highlight the advancements achieved by the proposed approach in accurately quantifying methane emissions using Sentinel-2 imagery. The generated benchmark dataset and the developed deep learning model presented in this study serve as a fundamental resource and constructive framework for future research, promoting extensive implementation across various methane monitoring scenarios on different satellites and in distinct geographic regions, which delivering greater effectiveness to global methane emission monitoring initiatives.

1. Introduction

Methane (CH₄) is one of the most important anthropogenic greenhouse gases, accounting for 25% of global warming since pre-industrial times, second only to carbon dioxide (CO₂) (Intergovernmental Panel on Climate Change, 2014). Despite having a shorter lifespan (about 9 years)

than carbon dioxide (>100 years), methane has an 86 times greater impact on global warming over a 20-year time horizon (Saunois et al., 2020; Etminan et al., 2016). This shorter lifespan can be beneficial in mitigating climate consequences on a significantly shorter timescale by reducing methane emissions (Jongaramrungruang et al., 2022; Montzka et al., 2011; Prather et al., 2012; Shindell et al., 2012). Anthropogenic

* Corresponding author at: Department of Electrical and Computer Engineering, Memorial University of Newfoundland, St. John's, NL A1B 3X5, Canada.
E-mail address: m.mahdianpari@mun.ca (M. Mahdianpari).

methane emitters including livestock, oil and gas, waste management, and coal mining contribute a large fraction of total emissions, with wetlands being the dominant natural source. This offers an opportunity to reduce methane emissions and mitigate climate change by monitoring, identifying, and repairing damaged facilities that cause emissions (Varon et al., 2021).

Due to the characteristics of methane, it can only be detected with specific gas-sensitive sensors (Irakulis-Loitxate et al., 2021). Observations from satellites equipped with shortwave infrared (SWIR) bands have the potential to measure point sources (Jacob et al., 2022). Currently, satellites carrying multispectral sensors, including Sentinel-2 and Landsat-8, offer frequent methane monitoring. Although these multispectral sensors were not specifically designed for methane monitoring, having methane absorption bands (about 1600 nm and 2300 nm) enables them to identify point source plumes and measure their column enhancement (Varon et al., 2021). Importantly, the higher spatial and temporal resolution of Sentinel-2 (i.e., 20 m and about 5 days) compared to Landsat-8 and -9 (i.e., 30 m and about 8 days) mean the data collected is more valuable for the regular monitoring of potential methane point sources (Ehret et al., 2021; Varon et al., 2021). However, it is important to note that the capabilities of multispectral sensors, including Sentinel-2, are limited to detecting only large point sources and not diffuse sources or area methane emissions. Although useful in detecting superemitters emitting above a few tons of methane per hour, recent studies indicate that super emitters account for only a minority of total regional emissions (Omara et al., 2022; Cusworth et al., 2022). Therefore, to achieve effective quantification and tracking of total emissions, more systematic characterization of smaller emission sources, including the use of other satellite capabilities, is necessary.

Several methods have been developed for retrieving point source flux rates from plume observations, including gaussian plume inversion (Krings et al., 2011, 2013; Rayner et al., 2014; Fioletov et al., 2015; Schwandner et al., 2017; Varon et al., 2018), source pixel (Jacob et al., 2016; Buchwitz et al., 2017; Varon et al., 2018), cross-sectional flux (CSF) (Krings et al., 2011, 2013; Frankenberg et al., 2016; Conley et al., 2016; Varon et al., 2018), and integrated mass enhancement (IME) (Frankenberg et al., 2016; Thompson et al., 2016; Varon et al., 2018). Even though many methods for plume detection have been developed, the measured methane column enhancements and source flux rates are subjected to substantial uncertainties due to the varying boundary layer conditions, spectral interferences, and the flux inversion's sensitivity to complex plume structures (Duren et al., 2019). These current techniques require auxiliary data, such as wind speed, when determining flux rate estimations, and differences in spatial resolution and acquisition time of wind speed data and satellite images further restrict methane flux rate accuracy. To address these issues, methane column enhancement and flux rate can be determined in a more automated manner using machine and deep learning approaches.

Deep learning methods have shown promising results in several remote sensing applications, such as image classification (Memon et al., 2021; Hosseini et al., 2022; Huang et al., 2018a), change detection (Karim and van Zyl, 2021; Keshk and Yin, 2020), and regression (Zhang et al., 2021; Boulila et al., 2021; Radman et al., 2022). The convolutional neural network (CNN) is the most prominent of these methods and has achieved successful results in various remote sensing applications (Mahdianpari et al., 2018; Ansari et al., 2021). Despite the extensive use and success of CNNs in several remote sensing and computer science applications, their capability for methane quantification has been only recently explored (Jongaramrunguang et al., 2022; Kumar et al., 2022; Joyce et al., 2022); for example, Jongaramrunguang et al. (2022) used a deep CNN model to predict flux rates directly from aerial methane plume images, and Joyce et al. (2022) used CNNs to detect and quantify methane plumes in PRISMA satellite data. A large dataset of methane plume images is required to train machine learning-based algorithms. Large Eddy Simulation (LES) (Matheou and Chung, 2014) can be used to produce simulated methane plumes at varying flux rates and wind

speeds (Jongaramrunguang et al., 2022; Varon et al., 2018; Nottrott et al., 2014). These simulated plumes can be used in deep-learning methods to quantify methane source rates where extensive real world plume data are inaccessible.

Many deep architectures have been successfully used in remote sensing regression problems, such as Inception (Szegedy et al., 2015), VGG (visual geometry group) (Simonyan and Zisserman, 2015), ResNet (residual network) (He et al., 2016), DenseNet (densely connected convolutional networks) (Huang et al., 2018b), Swin (shifted window) (Liu et al., 2021), and EfficientNet (Tan and Le, 2021). These architectures are potential solutions to addressing methane source rate estimation problems. In this study, we aim to develop a comprehensive methane plume benchmark dataset of large point sources of Sentinel-2 for automatic methane quantification and propose the most efficient existing deep architecture for that task. First, the benchmark dataset is produced by combining plume data and Sentinel-2 background noise data. The plume data are generated using the LES approach with various model specifications. Meanwhile, background noise is gathered from real Sentinel-2 scenes with an absence of methane plumes. The generated methane plume benchmark provides a substantial dataset that can be useful for methane retrieval and quantification studies, particularly when a large dataset is required (e.g., for developing deep learning networks). Next, the proposed dataset is used to compare the capability of state-of-the-art deep architectures to automatically estimate point source quantification rates without using any external auxiliary data. The primary contributions of the current study are as follows: (1) generating benchmark datasets by simulating methane plumes of Sentinel-2 satellite images; (2) designing an end-to-end procedure for automatic methane quantification using deep learning-based methods; and (3) comparing the potential of existing deep learning methods (e.g., VGG, Inception, ResNet, DenseNet, Swin, and EfficientNet) for methane quantification to identify the optimal learning strategy.

2. Methane plume benchmark

A sufficient volume of data representing a wide range of situations, including various point sources and plume conditions, is needed to develop and investigate new techniques (i.e., deep learning-based methods) for plume monitoring.

In case of monitoring a certain platform over different timesteps (Varon et al., 2021), the real detected plumes may vary, while certain criteria associated with these plumes, such as background and albedo, remain consistent in most cases. Moreover, constructing a large enough benchmark dataset of real cases of plumes to train very deep models is challenging. The detection and labeling of real cases require significant manual effort, and this process can be costly and time-consuming, as expert knowledge is required to identify and label plumes and emitters correctly. Furthermore, the accuracy of the labeled data significantly impacts the performance of the models, making it essential to ensure that the data labeling is consistent and reliable. Additionally, the deep learning models trained with detected real cases might be biased to those specific cases and may not be suitable for new undetected plumes. As a result, using real data to train deep models might not provide adequately varying data and could lead to underperformance of deep learning models in real-world applications.

To produce a comprehensive and large methane plume dataset for the Sentinel-2 satellite sensor, two main procedures are followed: (1) simulating methane plume with characteristics similar to Sentinel-2 plumes, and (2) generating background noise similar to the real conditions of satellite images.

Sentinel-2 is a two-satellite multispectral imaging constellation developed by the European Space Agency (ESA). The Sentinel-2 mission consists of two satellites, Sentinel-2A and Sentinel-2B, launched in June 2015 and March 2017, respectively. These satellites have sensors with 13 spectral bands, including 4 visible and near infrared (NIR) bands with 10 m resolution, 6 red-edge and SWIR bands with 20 m resolution, and 3

atmospheric bands with 60 m resolution. Sentinel-2 has a 290 km swath width and a revisit time of 5 days at the equator, allowing for frequent data acquisition of a particular area.

2.1. Methane plume simulation

In order to generate methane plume data, Large Eddy Simulations (LES) were performed using Weather Research and Forecasting (WRF) software version 3.8 (WRF Users' Guide, 2023; Skamarock et al., 2008). This technique produces three-dimensional (3D) methane distributions, emanating from a point source in the boundary layer over time, with various parameters (e.g., wind speed, heat flux). Detailed procedures and settings of this method are explained in several articles, including Nottrott et al. (2014), Jongaramrungruang et al. (2019), and Mattheou and Bowman (2016).

Table 1 provides a summary of some of the key configuration settings utilized for WRF-LES simulation. The passive scalar was used for atmospheric boundary layer (ABL) dispersion simulations by setting the value of the tracer option ("tracer_opt") (Nottrott et al., 2014). The passive tracers are influenced by the model rather than actively participating in the simulation (Blaylock et al., 2017). The LES boundary layer is incorporated with a specified surface heat flux (by setting "issflx = 2"). The surface momentum fluxes are estimated using Monin-Obukhov similarity theory, by modifying the "sf_sfclay_physics" parameter in the WRF configuration. This incorporates full diffusion for vertical mixing, utilizing the "diff_opt = 2" and "km_opt = 3" settings. The "periodic_x" and "periodic_y" parameters enable periodic lateral boundary conditions (Nottrott et al., 2014).

In this study, an initially uniform wind in the x direction is considered over each simulation in the WRF-LES. Wind speed is the main parameter affecting vertically integrated plume conditions owing to its impact on atmospheric turbulence dynamics (Varon et al., 2018; Jongaramrungruang et al., 2019). Accordingly, we perform 10 simulations with varied wind speeds ranging from 1 to 10 m s⁻¹. While other factors such as surface sensible and latent heat flux can influence the dynamic and stability of methane plumes, only variations in wind speed are considered in the simulations. This is due to the fact that the simulated plumes are presented as column enhancement in pixels of images and various plume shapes can be simulated with constant heat fluxes. This approach is consistent with previous studies in the field, which also employed constant latent and sensible heat flux values to simulate methane plumes (Varon et al., 2018; Sánchez-García et al., 2022; Gorroño et al., 2023). Accordingly, uniform latent and sensible heat flux of 40 and 400 W m⁻² are considered normal conditions for all simulations, similar to Jongaramrungruang et al. (2022). Each simulation consists of 1060 scenes captured at 10 s intervals. The simulated scenes of the first hour (360 scenes) are considered as spin-up and the remaining 700 scenes are used for methane plume generation.

The geometrical parameters are set using Sentinel-2 data. A spatial resolution of 20 m over a 2 × 2 km domain is considered with one-way nesting from external simulation with 100 m resolution and 3 × 3 km domain. Next, the methane plume from a 2 × 2 pixel (40 × 40 m) point

Table 1
WRF-LES key configuration setting.

Option	Setting	Value
Tracer option	Add passive tracer	tracer_opt = 2
Surface heat and moisture fluxes	Specified surface heat flux	issflx = 2
Surface layer option	Revised Monin-Obukhov scheme	sf_sfclay_physics = 1
Turbulence and mixing	Full diffusion: mixing in physical space	diff_opt = 2
Eddy coefficient	1.5 order TKE closure (3D)	km_opt = 2
Lateral boundary condition	Periodic lateral boundary condition	periodic_x = true periodic_y = true

source is generated using the passive tracer transport capability in WRF-LES (Varon et al., 2018; Nottrott et al., 2014; Nunalee et al., 2014). Then, 7000 scenes of plumes are produced by integrating the vertical field columns for the ensemble of 10 simulations.

The generated plume scenes represent a wide range of methane point source plumes that can be captured by Sentinel-2. However, WRF-LES produces a constant source rate over all the ensemble of simulations which leads to constant quantification rate. To address this issue and obtain the desired source rates, the initial simulations are scaled and resampled. This scaling process does not impact the sensitivity of the simulations (Sánchez-García et al., 2022). Accordingly, the produced scenes are randomly scaled into the range of 5000–30,000 kg h⁻¹. This range is selected based on the previously documented point sources observed by Sentinel-2. Large point sources up to 50,000 kg h⁻¹ have previously been documented with Sentinel-2 (Varon et al., 2021), and also down to 1500 kg h⁻¹ (Sherwin et al., 2022), so we select 5000–30,000 kg h⁻¹ as a feasible detection range for Sentinel-2.

The last step to modify plume scenes regarding the Sentinel-2 capacity is to mask pixels with a methane column below Sentinel-2 measurement precision. Varon et al. (2021) obtained precision of 0.31 mol m⁻² for the most common methane retrieval approach, known as Multi-Band-Single-Pass (MBSP), while this amount was 0.13 mol m⁻² for Multi-band-multi-pass (MBMP) technique. We utilize MBMP technique for both the base threshold and generating background noise (Section 2.2). Samples of simulated methane plumes at different time steps for various wind speeds at constant source rate of 10,000 kg h⁻¹ are represented in Fig. 1.

2.2. Sentinel-2 background noise

To produce a realistic benchmark of Sentinel-2 methane plumes, reconstructing specifications of the real plume images including background noise is necessary, as methane quantification approaches are highly affected by this noise (Gorroño et al., 2023). One way to produce this noise is using gaussian random noise which simulates systematic sensor noise with normal distribution. The background noise, however, in addition to systematic noise, has a high correlation with surface objects that are not distributed normally. The other strategy to reconstruct background noise is obtaining it from real Sentinel-2 scenes with an absence of methane plumes.

The real Sentinel-2 scenes are derived from regions around a selection of known sites with high methane emissions over the past few years. This includes oil and gas production regions in Algeria, Turkmenistan, and the United States (see Fig. 2), which have been amply documented in previous studies (Ehret et al., 2021; Varon et al., 2021; Sánchez-García et al., 2022). The characteristics of the collected data over these regions are presented in Table 2.

2.2.1. Methane retrieval

Thus far, several approaches have been developed to retrieve methane concentration enhancement from multispectral sensors including Sentinel-2. The MBMP method provides more precise results compared to single pass techniques such as MBSP (Varon et al., 2021; Z. Zhang et al., 2022). In the current study, the MBMP technique is used to obtain methane enhancements from Sentinel-2 SWIR bands (bands 12 and 11). This technique subtracts column enhancement of a reference scene (non-active point source) from the main scene (active point source) for methane retrieval. Methane column concentration (ΔXCH_4) for each single scene is determined using the following relation (Sánchez-García et al., 2022):

$$\Delta XCH_4 = -\log(T)/(AMF \cdot \sigma_{CH_4}) \quad (1)$$

Where T (transmittance) is the band ratio of the strong methane absorption band (band 12) to the methane-free band (band 11) radiance. AMF is airmass factor that is determined by slant optical path and is a function of angular condition. σ_{CH_4} is methane absorption cross-section.

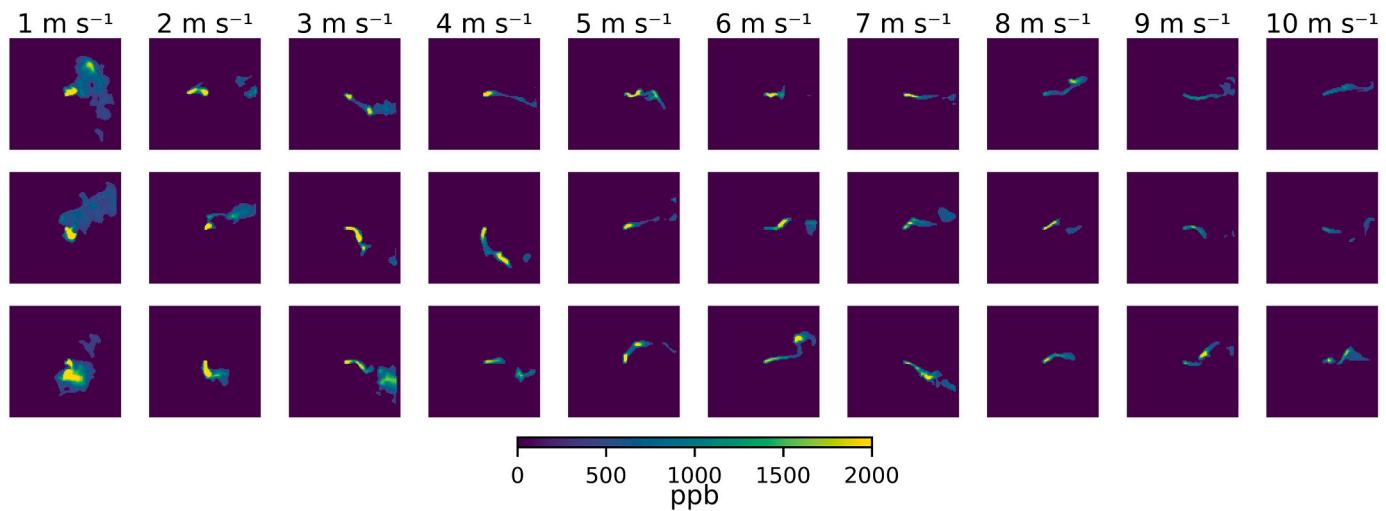


Fig. 1. Simulated methane plume samples at different time steps with varied wind speeds ranging from 1 to 10 m s⁻¹.

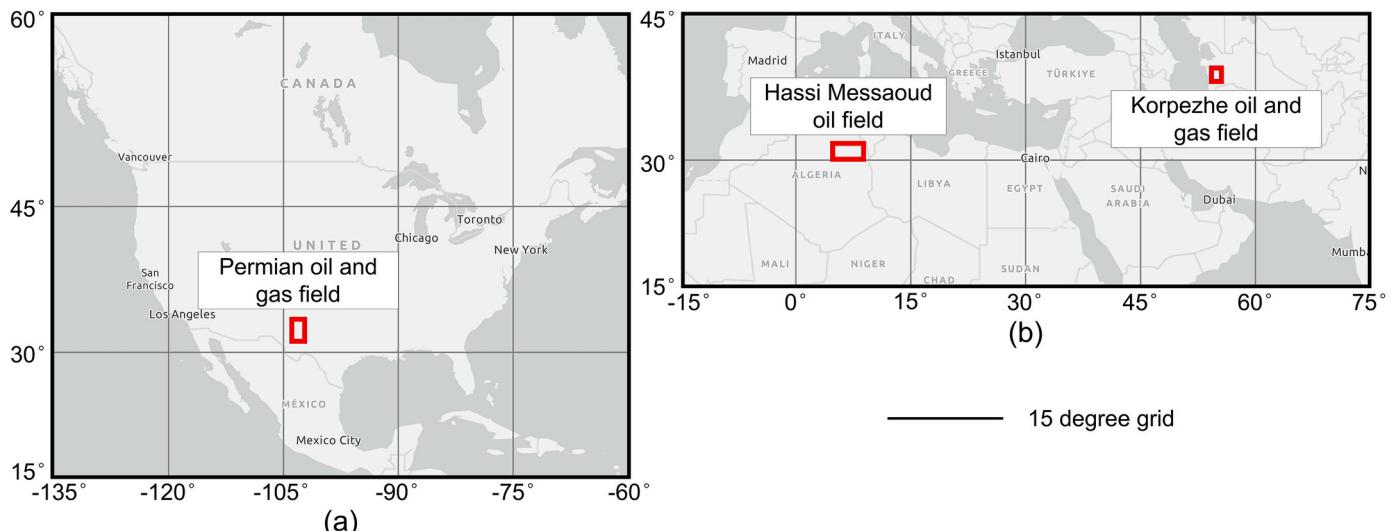


Fig. 2. Locations of large methane emission sites obtained for deriving background noise from the existing literature: (a) Permian oil and gas field in USA, (b) Hassi Messaoud oil field in Algeria, and Korpezhe oil and gas field in Turkmenistan.

Table 2

The region and date of acquired Sentinel-2 data used for generating background noise images.

Country	Site	Date	Sensor
Turkmenistan	Hassi Messaoud oil field	2022-07-13	Sentinel-2
		2021-07-18	
Algeria	Korpezhe oil and gas field	2022-07-12	
		2021-07-07	
USA	Permian oil and gas field	2022-06-14	
		2021-06-19	

2.2.2. Post-processing

The retrieved methane column concentration data are then clipped into 2 × 2 km scenes (similar to real plume images). These scenes are passed through filters to eliminate potential point sources and plumes from the background noise dataset. Accordingly, the top 5% scenes (with an average methane column exceeding 1.65 standard deviation above the mean of all background scene concentrations) are considered as potent plume images and excluded from the background noise dataset. Finally, 31,012 Sentinel-2 background methane scenes are extracted

from three of the largest methane point source sites. Examples of these background noise scenes are presented in Fig. 3.

2.3. Data augmentation

To imitate real plume images, we combine augmented WRF-LES plumes and Sentinel-2 background noise images. 10 augmentations, including random source rate scales (to reach quantification in range of 5000 to 30,000 kg h⁻¹), random rotations (in range of -180 to 180 degree), and random shifts in source location (between -5 and 5 pixels in each direction), are applied to each of the initial plume scenes. This augmentation produces 70,000 simulated plume images. The modified Sentinel-2 MBMP images, on the other hand, are randomly rotated (at multiples of 90 degrees) to establish background images for the plumes. As such, the benchmark dataset of Sentinel-2 methane plume images is produced by integrating the augmented plume scenes and real background data (Fig. 4).

2.4. Benchmark data analysis

Various sources of error can affect the detection of methane plumes

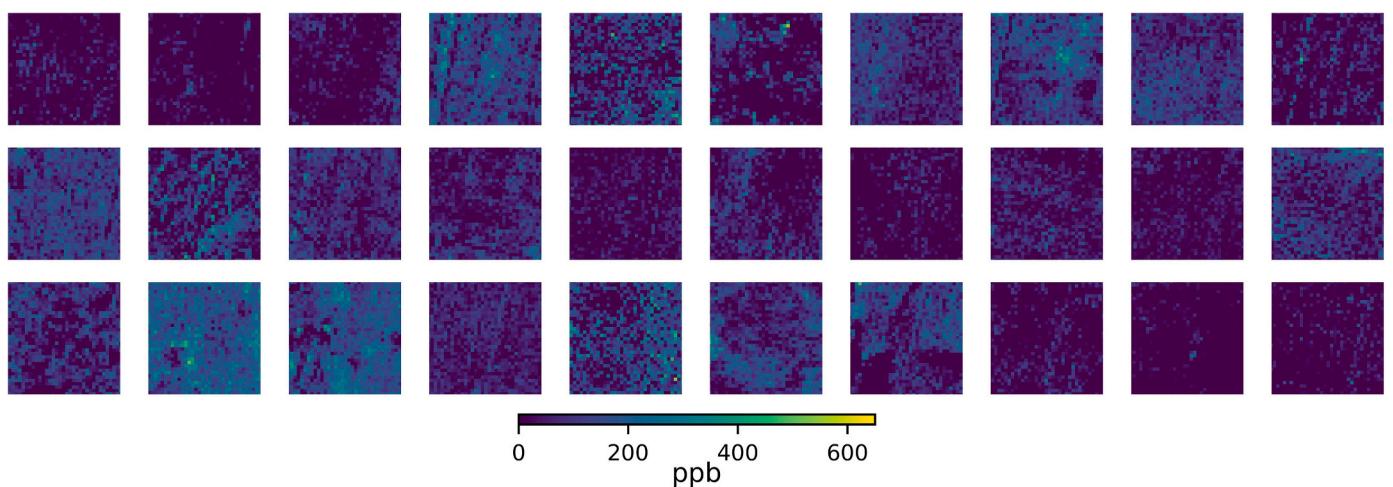


Fig. 3. Samples of extracted Sentinel-2 background noise images from Turkmenistan (first row), Algeria (second row), and USA (third row) sites.

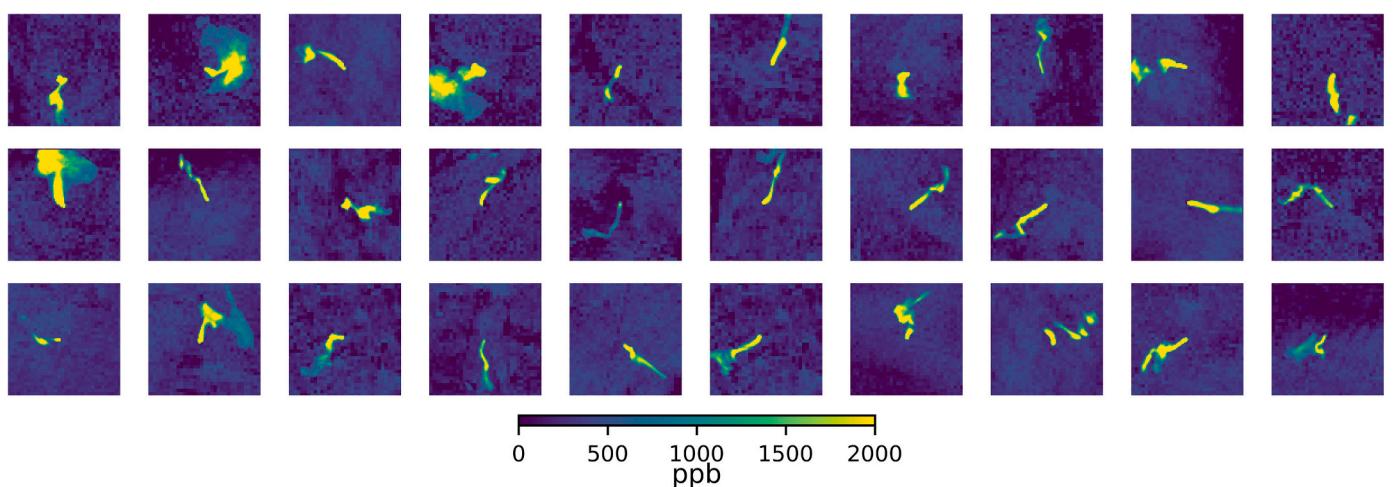


Fig. 4. Samples of produced methane point source scenes, generated by integrating the augmented plume scenes and real background data.

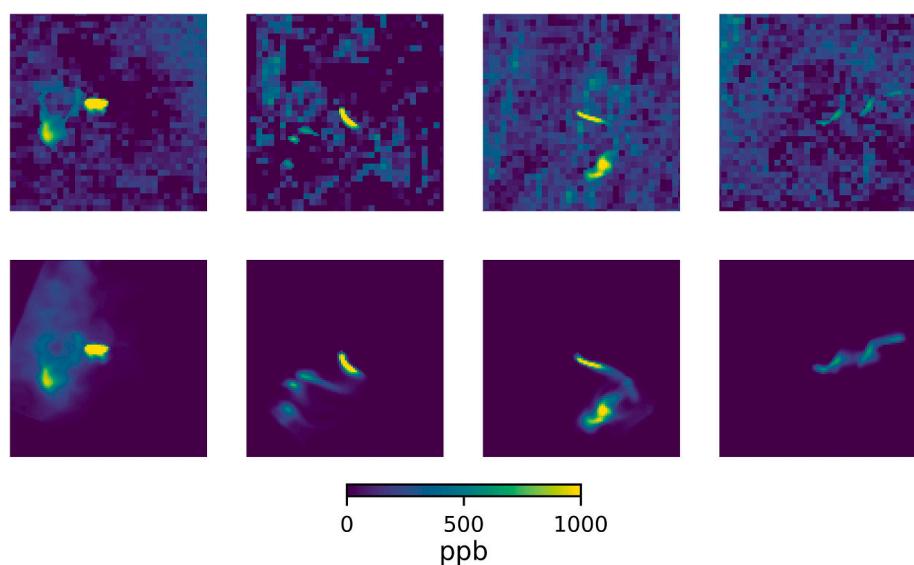


Fig. 5. Instances of produced Sentinel-2 methane concentration benchmark scenes with high background noises, broken and mixed plumes (top), and their corresponding simulated plume scenes (bottom).

and their backgrounds, including clouds, measurement precision errors, high-albedo surfaces, the presence of aerosols, and other retrieval errors.

The identification of methane plumes is significantly influenced by the presence of clouds in satellite imagery. In order to achieve accurate methane detection and monitoring, cloud-free images are favorable and provide valuable information (Pandey et al., 2019; Varon et al., 2021; Joyce et al., 2022). Accordingly, to generate accurate background images, we implemented a filtering process that excluded scenes containing clouds, and only utilized cloud-free observations.

To simulate real plume conditions and consider the capabilities of the Sentinel-2 sensor for methane detection, pixels with methane concentrations lower than the Sentinel-2 measurement precision were discarded. This low concentration removal resulted in the simulation of broken plume conditions.

Although aerosol impact is generally insignificant in Sentinel-2 satellite images of methane plumes (Varon et al., 2021), the slight impacts of aerosols on the background noise were considered by using real Sentinel-2 scenes for background images. Additionally, different background conditions, such as high surface albedo, were applied to the benchmark datasets using real background scenes.

Some instances of broken and mixed plumes with high background noises in the produced benchmark dataset are displayed in Fig. 5 and compared with the simulated full methane plume. The background noise may be attributed to various factors, including surface albedo artifacts and retrieval errors.

3. Deep learning

Deep neural networks (DNN) contain multiple processing layers with learnable weights which help them extract input data features and achieve high performance in many remote sensing and computer vision applications. The learnable weights are trained and optimized using the benchmark dataset in an iterative process namely backpropagation algorithm. This process typically requires a large set of data as the model parameters (weights) increase.

Several pioneering architectures have been developed and successfully utilized in computer vision tasks over the past few years. Most of these well-known models have been trained and evaluated using large famous benchmark datasets such as ImageNet (Russakovsky et al., 2015). These deep models are usually transferable to other applications in a new domain that uses similar visual features. Accordingly, the models can achieve an acceptable performance even with small training sets by using transferred weights. To achieve a more efficient performance in training stage, we obtained ImageNet weights in different learning strategies and compared them to the learning from scratch approach.

In the current study, some well-known deep learning architectures in computer vision as well as remote sensing, including VGG, Inception, ResNet, DenseNet, Swin, and EfficientNet are used. The VGG network (Simonyan and Zisserman, 2015) achieved significant performance in the ImageNet Large Scale Visual Recognition Competition (ILSVRC) 2014 for localization and classification tracks by using small 3×3 convolutional filters. VGG-16 was one of the most successful VGG structures that used 13 convolutional layers followed by 3 fully connected layers. The network reached a significant improvement by increasing the depth of convolutional layers to 16 in the VGG-19 network. The effective performance of this network indicated the potential of depth factor in deep structures.

Inception or GoogLeNet (Szegedy et al., 2015) incorporates Inception modules that are comprised of convolutional layers in a parallel manner. This results in a deep structure in both depth and width, while computational complexity does not increase. In comparison to conventional CNNs, such as the VGG network, the 42-layer Inception-v3 has fewer parameters. This increases the efficiency of the Inception model to learn complex features under minimal computing constraints.

Microsoft's Residual Networks (ResNet) (He et al., 2016) was the

winner of ILSVRC in 2015 for the classification task (Russakovsky et al., 2015). In this network, a residual learning module is implemented, which results in a very deep network. The residual module addresses problems associated with very deep models, including vanishing gradient problems, by adding identity connections. Thus, ResNet can be tuned with less complexity while using a deeper architecture compared to VGG.

In DenseNet architecture (Huang et al., 2018b), all layers are connected in the feed-forward fashion to address the vanishing gradient problem of very deep networks. Meanwhile, the number of network parameters is reduced, which improves efficiency of the network. Increasing the number of layers, even to more than a hundred, does not cause optimization difficulty. The 121-layer DenseNet performs successfully even with small training sets without overfitting owing to its all connected layers feature.

Recently, transformer-based models have demonstrated exceptional performance in natural language processing (NLP), and their applications have also extended to the computer vision domain. One such model, the shifted windows (Swin) transformer (Liu et al., 2021), utilizes a sliding window approach to extract both global and local features. This technique enables the extraction of long-distance and local information simultaneously. The Swin architecture comprises a Patch Partition module and four cascaded stages containing Swin transformer blocks. Each two Swin transformer blocks in the cascade consists of multi-head self-attention modules with both regular and shifted windowing configurations to enable global attention computation.

EfficientNets (Tan and Le, 2020, 2021) represent a new family of CNNs that offer improved training efficiency by reducing the number of trainable parameters and thereby minimizing training time. These models are designed using training-aware neural architecture search (NAS) and scaling to optimize the trade-off between training speed and parameter efficiency. EfficientNet incorporates ordinary convolutional layers and mobile inverted bottleneck convolutions (MBconv) modules. EfficientNetV2 (Tan and Le, 2021) utilizes fused mobile inverted bottleneck convolution (Fused-MBconv) and MBcov structures with smaller expansion ratios to reduce memory usage. In EfficientNetV2, the fused MBconv module is primarily utilized in the shallow layers of the network. Furthermore, the network width, depth, and input resolution are all simultaneously increased to enhance network performance.

3.1. Experiment setup

To estimate methane source rate, the deep learning architectures are utilized and trained with our methane plume benchmark dataset. Accordingly, VGG-19, Inception-v3, ResNet-50, DenseNet-121, the tiny version of Swin (Swin-T), and large version 2 of EfficientNet (EfficientNet-V2L) are deployed and modified for our purpose as a regression task. The deployed models are generally developed for classification task. To modify these architectures for the current study, the last fully connected layer and its following activation function are replaced by our case sensitive top layers (Table 3). This includes a flatten layer to pass base model output to a fully connected layer with 256 units. It is followed by a dropout with a rate of 0.5 to avoid overfitting. The last modified fully connected layer with its linear activation function estimates methane source rate.

We used 35,000 samples (50%) of the benchmark data for training, 14,000 for validation (20%), and the remaining 21,000 samples (30%) for testing the models. The models are trained with a learning rate of

Table 3
Modified top layers.

Layer 1	Flatten
Layer 2	Dense (256, activation function: ReLu)
Layer 3	Dropout (0.5)
Layer 4	Dense (1, activation function: linear)

10^{-4} using Adam optimizer over 50 epochs. For the learning stage, three different strategies are considered:

- Learning from scratch: This is the conventional learning process where the initial weights are set randomly, and the weights are adjusted during the training process to minimize the loss function.
- Transfer learning: This approach involves a pre-trained model in another domain or dataset. The pre-trained model's weights are transferred to the new domain, while they are fixed and not adjusted during the training process. The idea is to leverage the pre-trained model's learned features. This can be beneficial when the new dataset is small or when there are limited computing resources to train a model from scratch.
- Transfer learning with fine-tuning: This strategy is similar to transfer learning, when the transferred weights from another domain are used as initial weights, but they are tuned during the training stage by adjusting weights based on the new dataset features. Fine-tuning can lead to better performance compared to transfer learning when the new dataset is related to the pre-trained model's domain, and the new dataset is large enough to update the weights effectively.

In summary, learning from scratch involves random initialization of weights, transfer learning uses pre-trained model weights without adjusting weights, and transfer learning with fine-tuning initializes weights with pre-trained model weights and adjusts them during training. In the current study, the transfer learning strategies use weights obtained from the ImageNet dataset.

4. Experimental results

Performance of the deep learning architectures for methane source rate quantification are compared using the test dataset in terms of root mean square error (RMSE), mean absolute percentage error (MAPE), and Pearson correlation coefficient (R) (see Table 4). As mentioned earlier, three learning strategies, namely learning from-scratch, transfer learning, and transfer learning with fine-tuning, are examined for each network. The best result within each learning strategy in terms of evaluation metrics is highlighted in bold in Table 4.

Fig. 6 compares predicted methane source rates obtained from the fine-tuned models (for the testing dataset), as the best performing strategy, to the true source rates. EfficientNet-V2L (Fig. 6 (f)) has the tightest scatter plot, while Inception-v3 is sparse (Fig. 6 (c)). This agrees with the reported accuracies of the examined deep learning networks in Table 4.

4.1. Comparison with existing methane quantification techniques

In this section, the fine-tuned EfficientNet-V2L, the best model with the highest accuracy among all proposed approaches, is compared with state-of-the-art techniques for methane flux rate quantification. The most commonly used quantification method is IME, which has been extensively used in many studies and various satellite datasets including Sentinel-2 (Varon et al., 2021; Ehret et al., 2021). This technique is the

best adaptation for source rate estimation among the existing physical methane quantification approaches (Varon et al., 2018). This technique requires auxiliary data for wind speed. Here, we exploit the IME in a similar setting as presented in Varon et al., 2021. For further comparison, we also deploy the MethaNet architecture proposed by Jongaramrungruang et al., 2022. MethaNet provides an alternative approach to quantify methane flux rate from AVIRIS-NG (Airborne Visible Infrared Imaging Spectrometer - Next Generation) aerial observations without requiring auxiliary data (e.g., wind speed) using a DCNN that was developed specifically for methane quantification. The MethaNet contains four convolutional layers, max-pooling layers, a dropout, two fully-connected (FC) layers, and an output FC layer for flux rate estimation.

The Sentinel-2 methane plume benchmark dataset produced in this study is used for methane quantification using these state-of-the-art methods (i.e., IME and MethaNet). Their estimated flux rates are compared to the true values in Fig. 7.

Table 5 compares the accuracy of state-of-the-art methods to the best result reported in Table 4. The IME and MethaNet accuracies are very close, where IME has better MAPE and R, while MethaNet has a slightly more accurate RMSE. The proposed approach, based on fine-tuned EfficientNet-V2L, outperforms both existing methods in terms of all evaluation indices. In particular, the proposed model exhibits an improvement of approximately 1287 kg h^{-1} in terms of RMSE, a 3.92% reduction in MAPE, and a 5.01% enhancement in R compared to the IME method.

4.2. Effect of sampling ratio

To evaluate whether the amount of plume scenes in the generated benchmark dataset is enough to train deep learning models, we analyze the performance of fine-tuned EfficientNet-V2L (the best among utilized architectures) with varying sampling ratios ranging from 1% of the dataset (700 samples) to 90% (63,000 samples). Performance of the models are evaluated using the same unseen portion of the benchmark dataset (10% or 7000 samples) for validation. The effect of using different sampling ratios based on statistical factors of RMSE and R is compared in Fig. 8. By increasing the training ratio from 1% to 5%, the model's performance improves sharply, by $>400 \text{ kg h}^{-1}$ in RMSE and 2% in R. The experiment shows that using only 5% of the benchmark dataset for training of the proposed deep model (EfficientNet-V2L) results in higher accuracy when compared to IME and MethaNet with 50% of the dataset. When using $>10\%$ of the training dataset, the network performance is almost stable and the changes in accuracy parameters are insignificant (e.g., $<150 \text{ kg h}^{-1}$ RMSE and 1% R improvement by increasing sample ratio from 10% to 20%). The variation in evaluation factors are negligible for training ratios exceeding 50%. For example, a change in the sampling ratio from 50% to 60% results in a 0.2% increase in R and approximately a 36 kg h^{-1} drop in RMSE.

This experiment confirms that the methane plume benchmark dataset developed in this study is suitable for Sentinel-2-based methane monitoring studies.

Table 4
Performance of deep learning architectures for methane source rate estimation using the testing dataset.

Method	From-scratch			Transfer			Fine-tuning		
	RMSE	MAPE	R	RMSE	MAPE	R	RMSE	MAPE	R
	(kg h ⁻¹)	(%)	(%)	(kg h ⁻¹)	(%)	(%)	(kg h ⁻¹)	(%)	(%)
VGG-19	2639	12.81	93.14	3653	17.30	86.40	2206	10.51	95.24
ResNet-50	3086	14.42	90.44	3522	17.07	87.34	2659	12.69	93.08
Inception-v3	2829	13.38	92.05	4990	25.11	73.01	2784	13.50	92.41
DenseNet-121	2681	13.12	93.07	4148	20.16	81.94	2383	11.19	94.48
Swin-T	3137	14.96	90.03	4930	22.14	75.57	2678	12.56	92.86
EfficientNet-V2L	3027	14.47	90.96	3917	18.20	84.19	2101	10.05	95.70

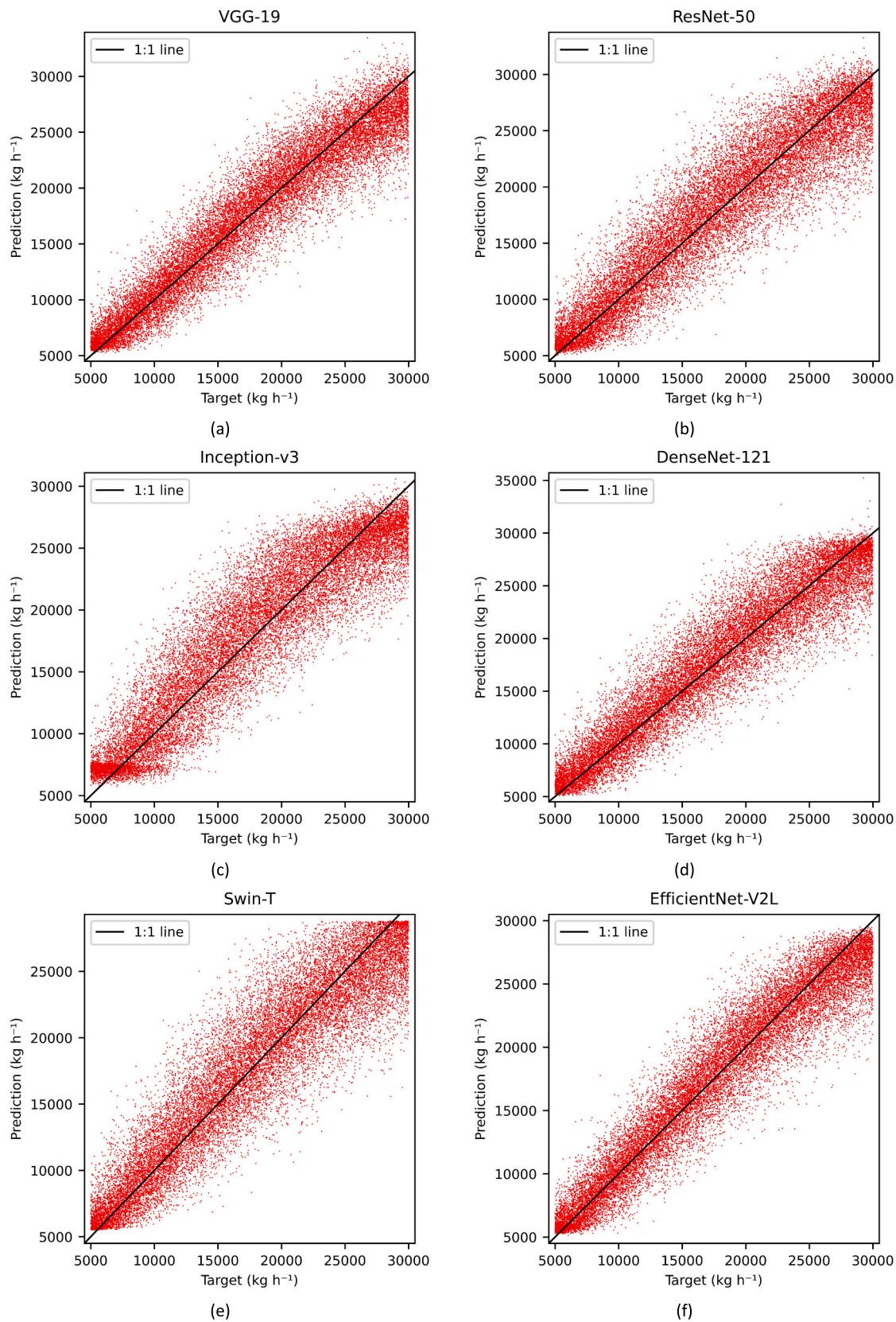


Fig. 6. Scatter plots presenting relationship of predicted and real values of methane source rate obtained from (a) VGG-19, (b) ResNet-50, (c) Inception-v3, (d) DenseNet-121, (e) Swin-T, and (f) EfficientNet-V2L.

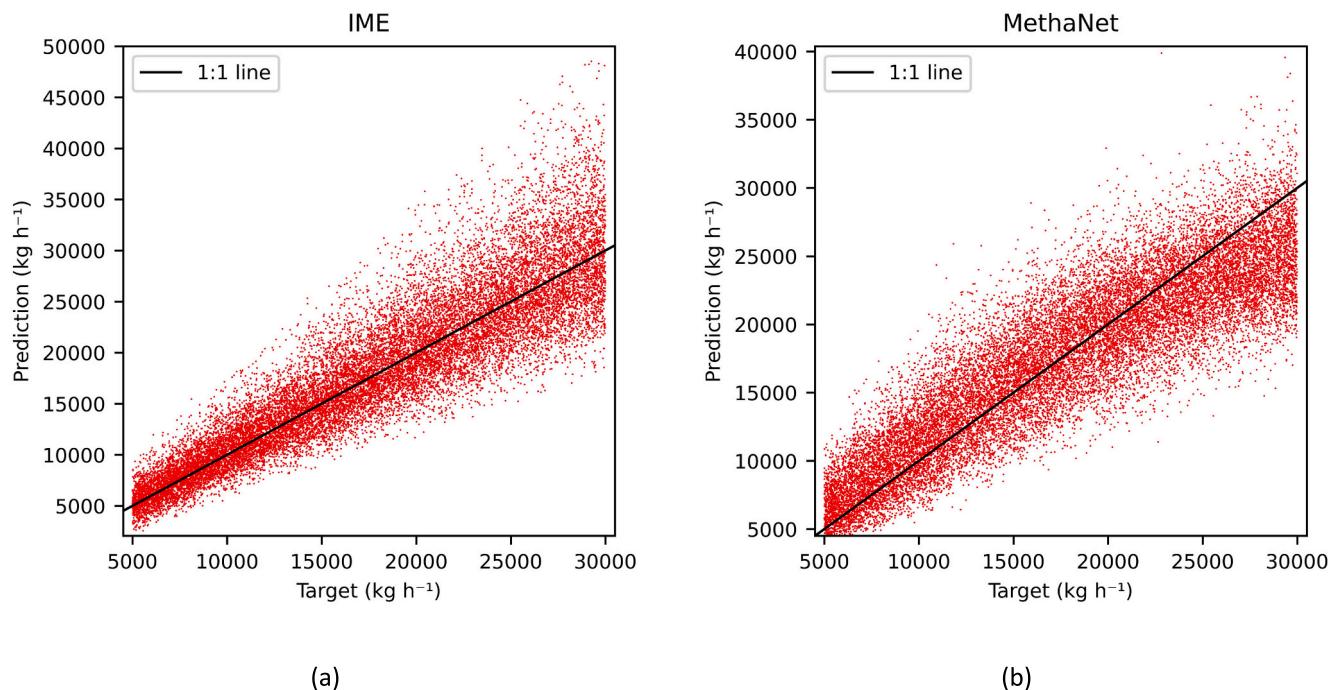


Fig. 7. Relationship of predicted and real values of methane source rate obtained from (a) IME and (b) MethaNet.

Table 5
Performance of state-of-the-art methane quantification approaches compared to fine-tuned EfficientNet-V2L.

Method	RMSE (kg h ⁻¹)	MAPE (%)	R (%)
IME	3388	13.97	90.69
MethaNet	3341	16.03	88.73
EfficientNet-V2L (fine-tuned)	2101	10.05	95.70

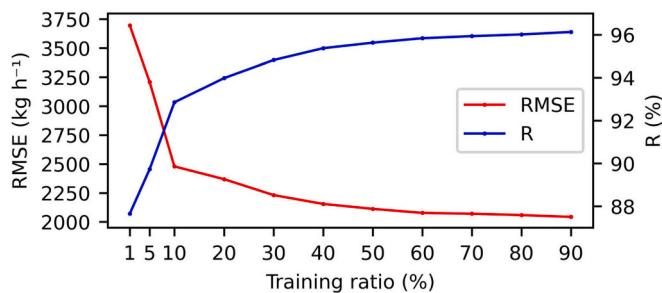


Fig. 8. Performance of deep model with different sampling ratios during the training stage.

4.3. Evaluation using real cases

We estimated the source rates of 27 real plumes in Sentinel-2 scenes for four locations over the Permian Basin, Hassi Messaoud, and Körpeze (two locations), to assess the practical value of the proposed method. Methane column enhancement images over these 27 samples are presented in Fig. 9.

The proposed method results are validated by comparing them to IME quantifications, using the approach of Varon et al. (2021). The plume area is manually adjusted by removing background noise to obtain more precise validation results using IME. Meanwhile, the plume scenes remain intact to be fed to the EfficientNet-V2L model for a more realistic assessment. The plumes and their corresponding estimated source rates with the IME and proposed approach (with fine-tuned

EfficientNet-V2L model) for one sample at each location are depicted in Fig. 10. The estimated source rates using the IME method and EfficientNet-V2L (fine-tuned) are noticeably close. This slight difference between EfficientNet-V2L (fine-tuned) and IME indicates capability of the proposed method to quantify methane source rates for real Sentinel-2 cases.

To further evaluate performance of the proposed model, it is cross validated with the IME results. The predicted source rates of the proposed model are compared to the IME results in scatter plot (Fig. 11). The estimated source rates using the IME method and EfficientNet-V2L (fine-tuned) are in agreement and indicate root mean square deviation (RMSD), mean absolute percentage deviation (MAPD), and Pearson R of 515.86 kg h⁻¹, 6.0%, and 96.75%, respectively.

5. Discussion

The results demonstrate the efficacy of deep learning architectures for methane source rate quantification. This discussion section provides an analysis of the potential of the employed deep learning models, a comparison of the best performing deep learning approach to existing methods, a discussion of the limitations and restrictions of the proposed method, and an exploration of the applicability of the proposed method for other satellites.

5.1. Proposed method performance analysis

Quantitative comparison of learning strategies indicates a superior performance of fine-tuning approaches compared to the other approaches examined in this study. However, the learning from-scratch strategy also results in acceptable estimation (above 90% R) when compared to the fine-tuning strategy. The transfer learning strategy, in which all weights of the base model are frozen, demonstrates the poorest performance among all. This transfer learning could be more beneficial when insufficient training data is available. The benchmark plume dataset, produced in this study, however, offers an enriched dataset, which is of great benefit during the training stage and produces accurate results even when the models are trained from-scratch. Accordingly, compared to the transfer approaches with frozen base models, from-

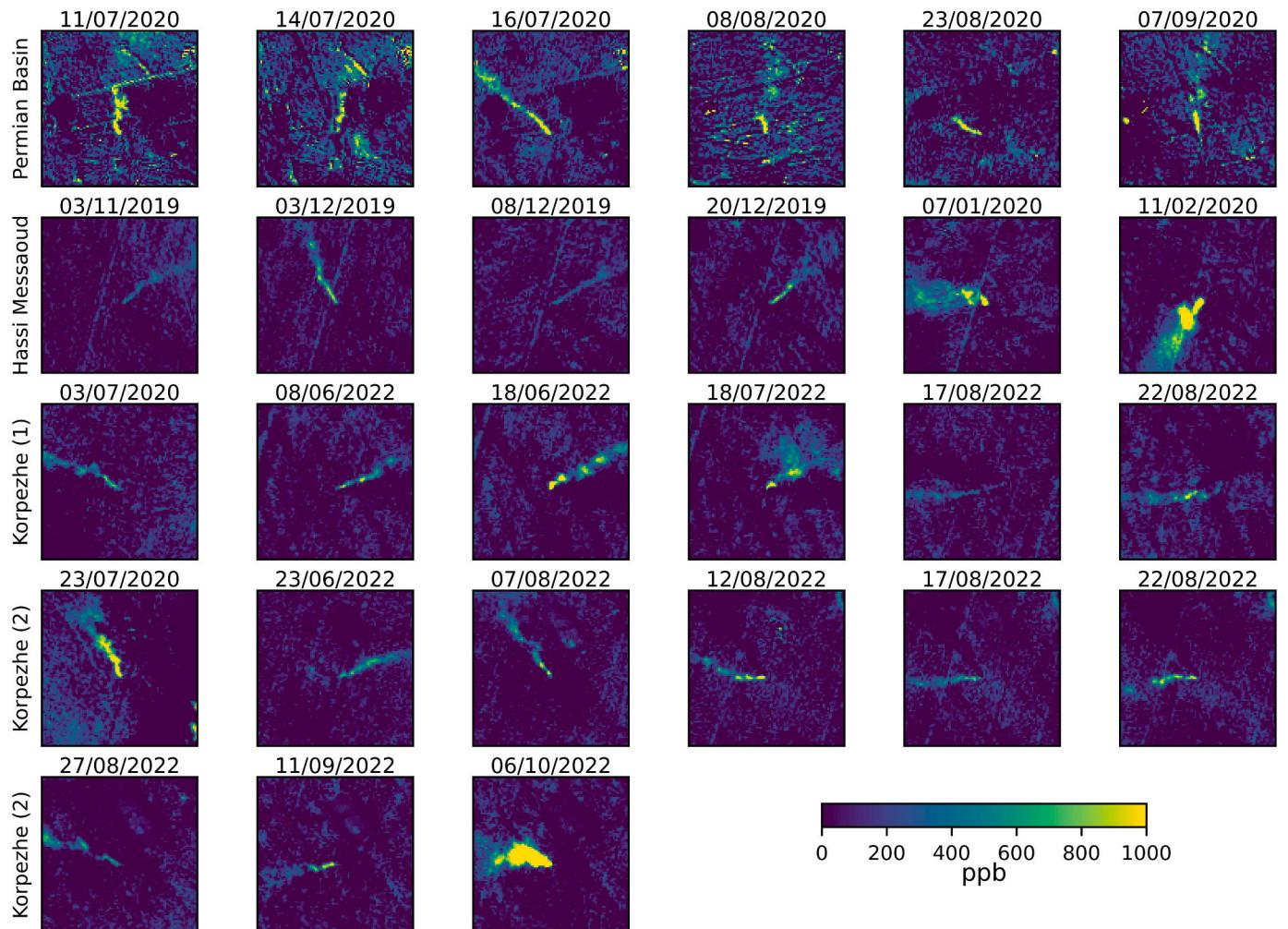


Fig. 9. Methane column concentration images obtained from Sentinel-2 data over four different platforms at the Permian Basin (first row), Hassi Messaoud (second row), Korpezhe (1) (third row), and Korpezhe (2) (fourth and fifth rows).

scratch methods reach noticeably improved performance. Comparing deep learning models, ResNet-50 shows better performance with the transfer learning strategy. Whereas VGG-19 obtains more precise methane source rates for from-scratch strategy. Among all obtained deep networks, the fine-tuned EfficientNet-V2L has the highest accuracy with 2101 kg h^{-1} RMSE, 10.05% MAPE, and 95.70% correlation (R).

5.2. Comparison with other methods

In comparison to existing approaches, the proposed deep learning method outperformed traditional techniques (IME and MethaNet) for methane source rate quantification. The results of the study suggest that the proposed deep learning method can provide more accurate and precise methane source rate estimates. For lower source rates, the IME estimations are close to the real values; however, as the flux rate increases, they become sparse and the inaccuracy rises (Fig. 7 (a)). The reason for the increased inaccuracy in IME method with higher flux rates is due to the quantification error, which is often expressed as a specific ratio or percentage of the source rate. This error is a product of uncertainties in methane retrieval, wind speed, and the IME model (Varon et al., 2019). Consequently, the quantification error rises with the source rate. In contrast, deep learning-based approaches produce a consistent amount of error across different source rates, as opposed to a proportional variation which is commonly observed in traditional methods (IME and CSF).

Accordingly, it is concluded that the proposed deep learning approaches provide higher accuracy compared to the widely accepted IME approach for methane quantification, particularly at high source rates, and it also eliminates the necessity for wind speed data. MethaNet predictions are generally sparse over all flux rates (Fig. 7 (b)) in comparison with the proposed approach (Fig. 6 (f)).

5.3. Limitations

While the proposed approach has shown promising results in estimating the methane plume source rate from Sentinel-2 satellite data, there are limitations to be considered. In particular, the proposed method is limited to the specified range of source rates (5000 kg h^{-1} to $30,000 \text{ kg h}^{-1}$) and may not work effectively for anomalous point sources outside of this range.

Although one-time training of the deep learning models is adequate to estimate methane plume source rate of unseen and new point sources, the proposed approach is based on deep learning models and requires sufficient time for training the models at first. Moreover, producing benchmark data for the supervised learning method can also be time-consuming. Although existing models, such as IME and CSF, are faster than the proposed approach, they still require the methane plume simulation step to determine effective wind speed parameters. Thus, even existing methods require a simulation step.

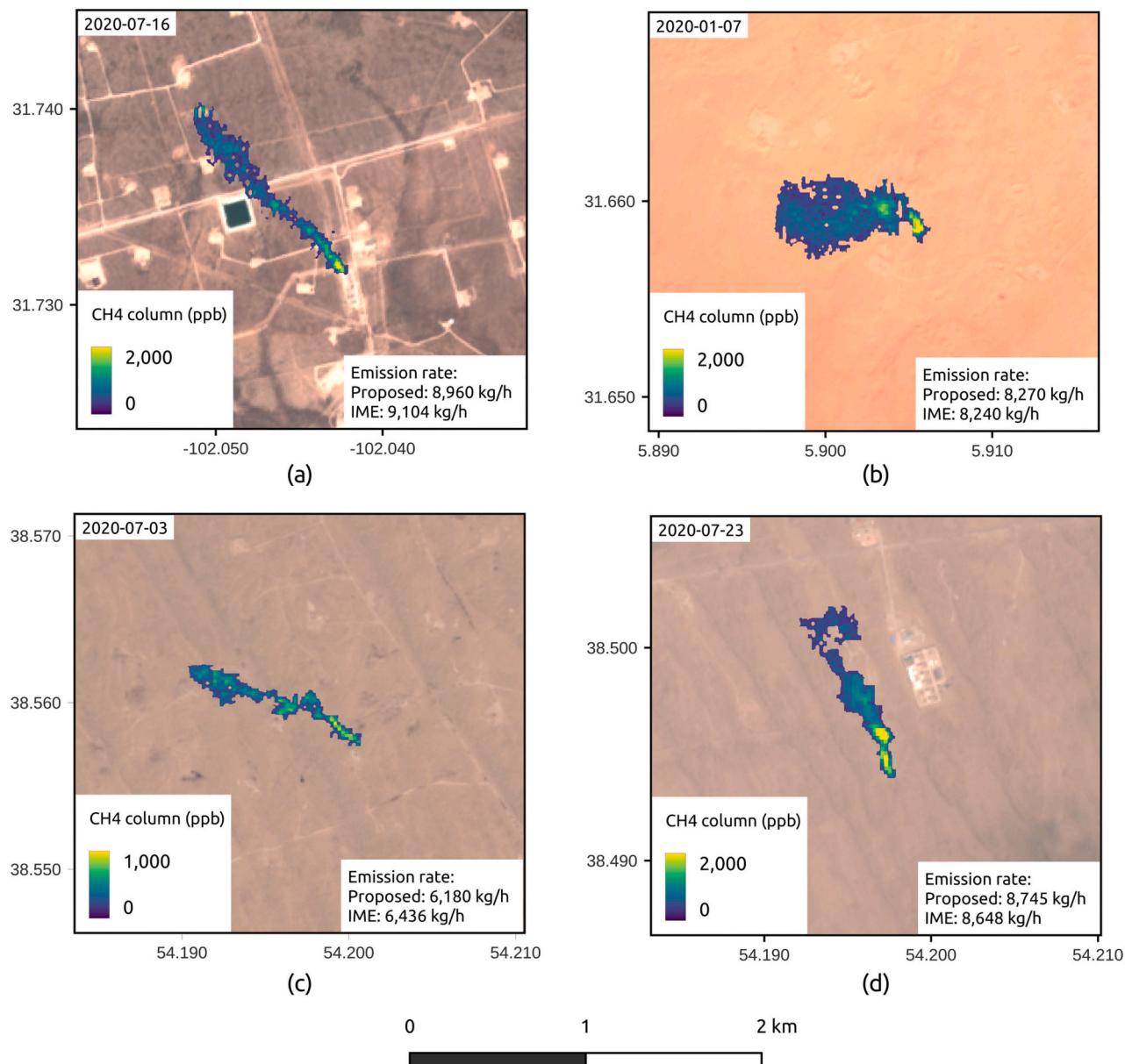


Fig. 10. Methane plumes and source rates quantified using IME and fine-tuned EfficientNet-V2L for real Sentinel-2 samples of (a) Permian Basin (b) Hassi Messaoud, (c) Korpezhe (1) and (d) Korpezhe (2).

5.4. Applicability for other satellites

The proposed method has shown promising results in estimating methane source rates from Sentinel-2 data using deep learning techniques. Previous studies have utilized machine and deep learning models to estimate methane emissions from different satellite sensors such as AVIRIS (Jongaramrungruang et al., 2022), PRISMA (Joyce et al., 2022), and TROPOMI (Schuit et al., 2023). These studies have shown great potential for machine and deep learning models in association with different satellite sensors to monitor methane point sources. This suggests that the proposed method could also be applicable for determining methane source rates using other satellite datasets. However, applying the proposed method to other satellite sensors would require modifications in all steps, including plume simulation, real background noise production, benchmark dataset generation, and training deep learning models. Consequently, while the proposed method has demonstrated excellent results in estimating methane source rates from Sentinel-2 data, its applicability to other satellite sensors remains to be

explored. Future studies could investigate the performance of the proposed method on other satellite sensors capable of detecting methane plumes by modifying all the necessary steps.

6. Conclusion

In this study, we produced an enriched benchmark dataset of Sentinel-2 methane plumes that can be utilized in future experiments aiming to monitor large point sources of methane captured by Sentinel-2, given its limitations in detecting diffuse or area emitters. The capability of several well-known deep learning models for automatic estimation of methane flux rate was investigated leveraging this dataset. The benchmark dataset was generated by aggregating simulated methane plumes and real background noise to imitate the real condition of Sentinel-2 plume images.

The potential of state-of-the-art deep learning architectures to estimate methane flux rate from Sentinel-2 images was evaluated. We used VGG-19, ResNet-50, Inception-v3, DenseNet-121, Swin-T, and

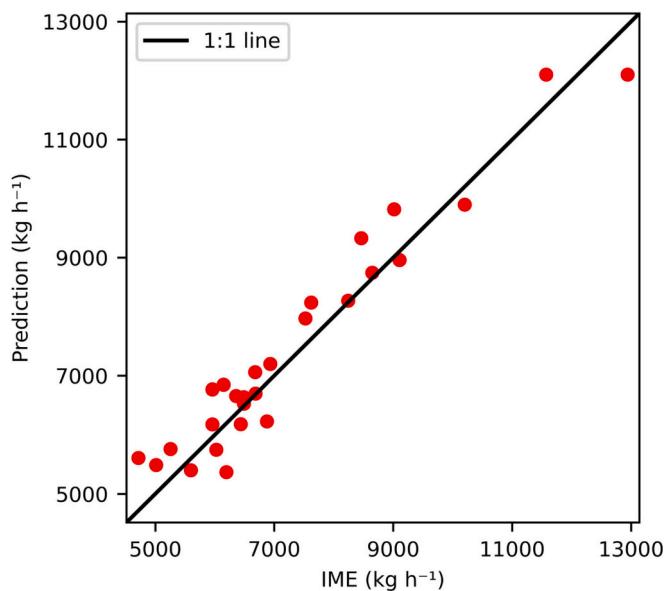


Fig. 11. Scatter plot of the relationship between predictions of the proposed method and IME using Sentinel-2 real images.

EfficientNet-V2L in several experiments and modified them for the regression tasks. Three different learning strategies, namely learning from scratch, transfer learning with fine-tuning, and transfer learning without fine-tuning, were examined for each network. The results indicated poor performance of the transfer learning with frozen base layers (without fine-tuning) compared to the other two approaches. Among all examined architectures with different strategies, the fine-tuned EfficientNet-V2L was superior and resulted in 2101 kg h^{-1} RMSE, 10.05% MAPE, and 95.70% Pearson R value.

The investigated deep learning approaches were further compared to conventional methods for methane flux rate quantification, including IME and a deep convolutional network called MethaNet. The best deep model in the current study (fine-tuned EfficientNet-V2L) surpasses both of these well-known quantification approaches by more than 1200 kg h^{-1} in RMSE, 3% in MAPE, and 5% R. In addition, the proposed algorithm, unlike IME, eliminates the necessity for auxiliary wind speed data, resulting in an automatic approach with a higher precision. More importantly, this study establishes a valuable benchmark dataset for Sentinel-2 methane plumes, facilitating future research on monitoring large point sources of methane using this satellite imagery. Moving forward, considering the limitations of Sentinel-2 in detecting diffuse emitters, future work could focus on developing hybrid models that combine Sentinel-2 data with other satellite sensors or ground-based measurements to enable comprehensive monitoring of methane emissions across diverse source types. Additionally, investigations into the transferability and generalization of the trained models to different geographic regions or varying environmental conditions would be valuable for expanding the applicability and practicality of this approach.

Funding

This study was funded by the Natural Sciences and Engineering Research Council (NSERC) Discovery under Grant to M. Mahdianpari (Grant No. RGPIN-2022-04766).

CRediT authorship contribution statement

Ali Radman: Conceptualization, Methodology, Investigation, Visualization, Software, Writing – review & editing. **Masoud Mahdianpari:** Conceptualization, Methodology, Visualization, Investigation,

Supervision, Funding acquisition, Writing – review & editing. **Daniel J. Varon:** Conceptualization, Methodology, Investigation, Writing – review & editing. **Fariba Mohammadimanesh:** Conceptualization, Methodology, Investigation, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

- Ansari, Mohsen, Homayouni, Saeid, Safari, Abdolreza, Niazmardi, Saeid, 2021. A new convolutional kernel classifier for hyperspectral image classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 14, 11240–11256. <https://doi.org/10.1109/JSTARS2021.3123087>.
- Blaylock, Brian K., Horel, John D., Crosman, Erik T., 2017. Impact of Lake breezes on summer ozone concentrations in the salt Lake Valley. *J. Appl. Meteorol. Climatol.* 56 (2), 353–370. <https://doi.org/10.1175/JAMC-D-16-0216.1>.
- Bouilala, Wadii, Ghadheri, Hamza, Khan, Mehshan Ahmed, Ahmed, Fawad, Ahmad, Jawad, 2021. A novel CNN-LSTM-based approach to predict urban expansion. *Ecol. Inform.* 64 (September), 101325 <https://doi.org/10.1016/j.ecoinf.2021.101325>.
- Buchwitz, Michael, Schneising, Oliver, Reuter, Maximilian, Heymann, Jens, Krautwurst, Sven, Bovensmann, Heinrich, Burrows, John P., et al., 2017. Satellite-derived methane hotspot emission estimates using a fast data-driven method. *Atmos. Chem. Phys.* 17 (9), 5751–5774. <https://doi.org/10.5194/acp-17-5751-2017>.
- Conley, S., Franco, G., Faloona, I., Blake, D.R., Peischl, J., Ryerson, T.B., 2016. Methane emissions from the 2015 aliso canyon blowout in Los Angeles, CA. *Science* 351 (6279), 1317–1320. <https://doi.org/10.1126/science.aaf2348>.
- Cusworth, Daniel H., Thorpe, Andrew K., Ayasse, Alana K., Stepp, David, Heckler, Joseph, Asner, Gregory P., Miller, Charles E., 2022. Strong methane point sources contribute a disproportionate fraction of total emissions across multiple basins in the United States. *Proc. Natl. Acad. Sci.* 119 (38), e2202338119 <https://doi.org/10.1073/pnas.2202338119>.
- Duren, Riley M., Thorpe, Andrew K., Foster, Kelsey T., Rafiq, Talha, Hopkins, Francesca M., Yadav, Vineet, Bue, Brian D., et al., 2019. California's methane super-emitters. *Nature* 575 (7781), 180–184. <https://doi.org/10.1038/s41586-019-1720-3>.
- Ehret, Thibaud, Truchis, Aurélien De, Mazzolini, Matthieu, Morel, Jean-Michel, d'Aspremont, Alexandre, Lauvaux, Thomas, Facciolo, Gabriele, 2021. Global Tracking and Quantification of Oil and Gas Methane Leaks from Recurrent Sentinel-2 Imagery. *ArXiv Preprint ArXiv:2110.11832*.
- Etminan, M., Myhre, G., Highwood, E.J., Shine, K.P., 2016. Radiative forcing of carbon dioxide, methane, and nitrous oxide: a significant revision of the methane radiative forcing. *Geophys. Res. Lett.* 43 (24) <https://doi.org/10.1002/2016GL071930>.
- Fioletov, V.E., McLinden, C.A., Krotkov, N., Li, C., 2015. Lifetimes and emissions of SO 2 from point sources Estimated from OMI. *Geophys. Res. Lett.* 42 (6), 1969–1976. <https://doi.org/10.1002/2015GL063148>.
- Frankenberg, Christian, Thorpe, Andrew K., Thompson, David R., Hulley, Glynn, Kort, Eric Adam, Vance, Nick, Borchardt, Jakob, 2016. Airborne methane remote measurements reveal heavy-tail flux distribution in four corners region. *Proc. Natl. Acad. Sci.* 113 (35), 9734–9739. <https://doi.org/10.1073/pnas.1605617113>.
- Gorrono, Javier, Varon, Daniel J., Irakulis-Loitxate, Itziar, Guanter, Luis, 2023. Understanding the potential of Sentinel-2 for monitoring methane point emissions. *Atmos. Meas. Tech.* 16 (1), 89–107. <https://doi.org/10.5194/amt-16-89-2023>.
- He, Kaiming, Zhang, Xiangyu, Ren, Shaoqing, Sun, Jian, 2016. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778.
- Hosseini, Benyamin, Mahdianpari, Masoud, Brisco, Brian, Mohammadimanesh, Fariba, Salehi, Bahram, 2022. WetNet: a spatial-temporal ensemble deep learning model for wetland classification using Sentinel-1 and Sentinel-2. *IEEE Trans. Geosci. Remote Sens.* 60, 1–14. <https://doi.org/10.1109/TGRS.2021.3113856>.
- Huang, Bo, Zhao, Bei, Song, Yimeng, 2018. Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery. *Remote Sens. Environ.* 214 (September), 73–86. <https://doi.org/10.1016/j.rse.2018.04.050>.
- Huang, Gao, Liu, Zhuang, van der Maaten, Laurens, Weinberger, Kilian Q., 2018. Densely Connected Convolutional Networks arXiv. <http://arxiv.org/abs/1608.06993>.
- Intergovernmental Panel on Climate Change, 2014. Climate change 2013 – the physical science basis: Working group I contribution to the fifth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9781107415324>.
- Irakulis-Loitxate, Itziar, Guanter, Luis, Maasakers, Joannes D., Zavalera-Araiza, Daniel, Aben, Ilse, 2021. Satellites Unveil Easily-Fixable Super-Emissions in One of the World's Largest Methane Hotspot Regions. <https://doi.org/10.31223/X56G7R>. Preprint Submitted to EarthArXiv.

- Jacob, Daniel J., Turner, Alexander J., Maasakkers, Joannes D., Sheng, Jianxiong, Sun, Kang, Liu, Xiong, Chance, Kelly, Aben, Ilse, McKeever, Jason, Frankenberg, Christian, 2016. Satellite observations of Atmospheric methane and their value for quantifying methane emissions. *Atmos. Chem. Phys.* 16 (22), 14371–14396. <https://doi.org/10.5194/acp-16-14371-2016>.
- Jacob, Daniel J., Varon, Daniel J., Cusworth, Daniel H., Dennison, Philip E., Frankenberg, Christian, Gautam, Ritesh, Guanter, Luis, et al., 2022. Quantifying methane emissions from the global scale down to point sources using satellite observations of Atmospheric methane. *Atmos. Chem. Phys.* 22 (14), 9617–9646. <https://doi.org/10.5194/acp-22-9617-2022>.
- Jongaramrunguang, Siraput, Frankenberg, Christian, Matheou, Georgios, Thorpe, Andrew K., Thompson, David R., Kuai, Le, Duren, Riley M., 2019. Towards accurate methane point-source quantification from high-resolution 2-D plume imagery. *Atmos. Meas. Tech.* 12 (12), 6667–6681. <https://doi.org/10.5194/amt-12-6667-2019>.
- Jongaramrunguang, Siraput, Thorpe, Andrew K., Matheou, Georgios, Frankenberg, Christian, 2022. MethaNet – an AI-driven approach to quantifying methane point-source emission from high-resolution 2-D plume imagery. *Remote Sens. Environ.* 269 (February), 112809 <https://doi.org/10.1016/j.rse.2021.112809>.
- Joyce, Peter, Villena, Cristina Ruiz, Huang, Yahui, Webb, Alex, Gloor, Manuel, Wagner, Fabien Hubert, Chipperfield, Martyn P., Guilló, Rocío Barrio, Wilson, Chris, Boesch, Hartmut, 2022. Using a deep neural network to detect methane point sources and quantify emissions from PRISMA hyperspectral satellite images. *Gases/Remote Sensing/Data Processing and Information Retrieval*, Preprint. <https://doi.org/10.5194/egusphere-2022-924>.
- Karim, Zainolabadien, van Zyl, Terence L., 2021. Deep/Transfer learning with feature space ensemble networks (FeatSpaceEnsNets) and average ensemble networks (AvgEnsNets) for change detection using DInSAR Sentinel-1 and optical Sentinel-2 satellite data fusion. *Remote Sens.* 13 (21), 4394. <https://doi.org/10.3390/rs13214394>.
- Keshk, Hatem Magdy, Yin, Xu-Cheng, 2020. Change detection in SAR images based on deep learning. *Int. J. Aeronaut. Space Sci.* 21 (2), 549–559. <https://doi.org/10.1007/s42405-019-00222-0>.
- Krings, T., Gerilowski, K., Buchwitz, M., Hartmann, J., Sachs, T., Erzinger, J., Burrows, J. P., Bovensmann, H., 2013. Quantification of methane emission rates from coal mine ventilation shafts using airborne remote sensing data. *Atmos. Meas. Tech.* 6 (1), 151–166. <https://doi.org/10.5194/amt-6-151-2013>.
- Krings, T., Gerilowski, K., Buchwitz, M., Reuter, M., Tretner, A., Erzinger, J., Heinze, D., Pflüger, U., Burrows, J.P., Bovensmann, H., 2011. MAMAP – a new spectrometer system for column-averaged methane and carbon dioxide observations from aircraft: retrieval algorithm and first inversions for point source emission rates. *Atmos. Meas. Tech.* 4 (9), 1735–1758. <https://doi.org/10.5194/amt-4-1735-2011>.
- Kumar, Satish, Kingwill, William, Mouton, Rozanne, Adamczyk, Wojciech, Huppertz, Robert, Sherwin, Evan D., 2022. Guided transformer network for detecting methane emissions in Sentinel-2 satellite imagery. In *Climate Change AI* <https://www.climatechange.ai/papers/neurips2022/94>.
- Liu, Ze, Lin, Yutong, Cao, Yue, Hu, Han, Wei, Yixuan, Zhang, Zheng, Lin, Stephen, Guo, Baining, 2021. Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows arXiv.
- Mahdianpari, Masoud, Salehi, Bahram, Rezaee, Mohammad, Mohammadmiranesh, Fariba, Zhang, Yun, 2018. Very deep convolutional neural networks for complex land cover mapping using multispectral remote sensing imagery. *Remote Sens.* 10 (7), 1119. <https://doi.org/10.3390/rs10071119>.
- Matheou, Georgios, Bowman, Kevin W., 2016. A recycling method for the large-Eddy simulation of plumes in the Atmospheric boundary layer. *Environ. Fluid Mech.* 16 (1), 69–85. <https://doi.org/10.1007/s10652-015-9413-4>.
- Matheou, Georgios, Chung, Daniel, 2014. Large-Eddy simulation of stratified turbulence. Part II: application of the stretched-vortex model to the Atmospheric boundary layer. *J. Atmos. Sci.* 71 (12), 4439–4460. <https://doi.org/10.1175/JAS-D-13-0306.1>.
- Memon, Nimrabanu, Parikh, Hemani, Patel, Samir B., Patel, Dhruvesh, Patel, Vibha D., 2021. Automatic land cover classification of multi-resolution dualpol data using convolutional neural network (CNN). *Remote Sens. Appl. Soc. Environ.* 22 (April), 100491 <https://doi.org/10.1016/j.rsease.2021.100491>.
- Montzka, S.A., Dlugokencky, E.J., Butler, J.H., 2011. Non-CO₂ Greenhouse gases and climate change. *Nature* 476 (7358), 43–50. <https://doi.org/10.1038/nature10322>.
- Nottrott, Anders, Kleissl, Jan, Keeling, Ralph, 2014. Modeling passive scalar dispersion in the Atmospheric boundary layer with WRF large-Eddy simulation. *Atmos. Environ.* 82 (January), 172–182. <https://doi.org/10.1016/j.atmosev.2013.10.026>.
- Nunalee, Christopher G., Kosović, Branko, Bieringer, Paul E., 2014. Eulerian dispersion modeling with WRF-LES of plume impingement in neutrally and stably stratified turbulent boundary layers. *Atmos. Environ.* 99 (December), 571–581. <https://doi.org/10.1016/j.atmosev.2014.09.070>.
- Omara, Mark, Zavala-Araiza, Daniel, Lyon, David R., Hmiel, Benjamin, Roberts, Katherine A., Hamburg, Steven P., 2022. Methane Emissions from US low production oil and Natural gas well sites. *Nat. Commun.* 13 (1), 2085. <https://doi.org/10.1038/s41467-022-29709-3>.
- Pandey, Sudhanshu, Gautam, Ritesh, Houweling, Sander, van der Gon, Hugo Denier, Sadavarte, Pankaj, Borsdorff, Tobias, Hasekamp, Otto, 2019. Satellite observations reveal extreme methane leakage from a natural gas well Blowout. *Proc. Natl. Acad. Sci.* 116 (52), 26376–26381. <https://doi.org/10.1073/pnas.1908712116>.
- Prather, Michael J., Holmes, Christopher D., Hsu, Juno, 2012. Reactive greenhouse gas scenarios: systematic exploration of uncertainties and the role of Atmospheric Chemistry: ATMOSPHERIC CHEMISTRY AND GREENHOUSE GASES. *Geophys. Res. Lett.* 39 (9), n/a-n/a. <https://doi.org/10.1029/2012GL051440>.
- Radman, Ali, Akhoondzadeh, Mehdi, Hosseiny, Benyamin, 2022. “Monitoring and predicting temporal changes of urmia Lake and its basin using satellite multi-sensor data and deep-learning Algorithms”. *PFG – journal of photogrammetry, remote sensing and geo-information. Science* 90 (3), 319–335. <https://doi.org/10.1007/s41064-022-00203-1>.
- Rayner, P.J., Utetme, S.R., Crowell, S., 2014. Constraining regional greenhouse gas emissions using geostationary concentration measurements: a theoretical study. *Atmos. Meas. Tech.* 7 (10), 3285–3293. <https://doi.org/10.5194/amt-7-3285-2014>.
- Russakovskiy, Olga, Deng, Jia, Hao, Su, Krause, Jonathan, Satheesh, Sanjeev, Ma, Sean, Huang, Zhiheng, et al., 2015. ImageNet large scale visual recognition challenge. *Int. J. Comput. Vis.* 115 (3), 211–252. <https://doi.org/10.1007/s11263-015-0816-y>.
- Sánchez-García, Elena, Gorroño, Javier, Irakulis-Loitxate, Itziar, Varon, Daniel J., Guanter, Luis, 2022. Mapping methane plumes at very high spatial resolution with the WorldView-3 satellite. *Atmos. Meas. Tech.* 15 (6), 1657–1674. <https://doi.org/10.5194/amt-15-1657-2022>.
- Saunois, Marielle, Stavert, Ann R., Poulet, Ben, Bousquet, Philippe, Canadell, Josep G., Jackson, Robert B., Raymond, Peter A., et al., 2020. The global methane budget 2000–2017. *Earth Syst. Sci. Data* 12 (3), 1561–1623. <https://doi.org/10.5194/essd-12-1561-2020>.
- Schuit, B.J., Maasakkers, J.D., Bijl, P., Mahapatra, G., Van den Berg, A.-W., Pandey, S., Lorente, A., Borsdorff, T., Houweling, S., Varon, D.J., McKeever, J., Jervis, D., Girard, M., Irakulis-Loitxate, I., Gorroño, J., Guanter, L., Cusworth, D.H., Aben, I., 2023. Automated detection and monitoring of methane super-emitters using satellite data. *Atmos. Chem. Phys. Discuss.* [preprint]. <https://doi.org/10.5194/acp-2022-862>.
- Schwandner, Florian M., Gunson, Michael R., Miller, Charles E., Carn, Simon A., Eldering, Annmarie, Krings, Thomas, Verhulst, Kristal R., et al., 2017. Spaceborne detection of localized carbon dioxide sources. *Science* 358 (6360), eaam5782. <https://doi.org/10.1126/science.aam5782>.
- Sherwin, Evan, Rutherford, Jeffrey, Chen, Yuanlei, Aminfarid, Sam, Kort, Eric, Jackson, Robert, Brandt, Adam, 2022. Single-blind validation of space-based point-source methane emissions detection and quantification. *Atmospheric Sciences, Preprint*. <https://doi.org/10.31223/X5DH09>.
- Shindell, Drew, Kyulenstierna, Johan C.I., Vignati, Elisabetta, van Dingenen, Rita, Amann, Markus, Klimont, Zbigniew, Anenberg, Susan C., et al., 2012. Simultaneously mitigating near-term climate change and improving human health and food security. *Science* 335 (6065), 183–189. <https://doi.org/10.1126/science.1210026>.
- Simonyan, Karen, Zisserman, Andrew, 2015. Very Deep Convolutional Networks for Large-Scale Image Recognition. *ArXiv:1409.1556 [Cs]*, April. <http://arxiv.org/abs/1409.1556>.
- Skamarock, William, Klemp, Joseph, Dudhia, Jimy, Gill, David, Barker, Dale, Wang, Wei, Huang, Xiang-Yu, Duda, Michael, 2008. A description of the advanced research WRF version 3. UCAR/NCAR, Application/pdf. <https://doi.org/10.5065/D68S4MVH>.
- Szegedy, Christian, Liu, Wei, Jia, Yangqing, Sermanet, Pierre, Reed, Scott, Anguelov, Dragomir, Erhan, Dumitru, Vanhoucke, Vincent, Rabinovich, Andrew, 2015. In: *Going Deeper with Convolutions*, pp. 1–9.
- Tan, Mingxing, Le, Quoc V., 2020. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *arXiv*. <http://arxiv.org/abs/1905.11946>.
- Tan, Mingxing, Le, Quoc V., 2021. EfficientNetV2: Smaller Models and Faster Training. *arXiv*. <http://arxiv.org/abs/2104.00298>.
- Thompson, D.R., Thorpe, A.K., Frankenberg, C., Green, R.O., Duren, R., Guanter, L., Hollstein, A., Middleton, E., Ong, L., Ungar, S., 2016. Space-based remote imaging spectroscopy of the also canyon CH 4 superemitter. *Geophys. Res. Lett.* 43 (12), 6571–6578. <https://doi.org/10.1002/2016GL069079>.
- Varon, Daniel J., Jacob, Daniel J., McKeever, Jason, Jervis, Dylan, Durak, Berke O.A., Xia, Yan, Huang, Yi, 2018. Quantifying methane point sources from fine-scale satellite observations of atmospheric methane plumes. *Atmos. Meas. Tech.* 11 (10), 5673–5686. <https://doi.org/10.5194/amt-11-5673-2018>.
- Varon, Daniel J., Jervis, Dylan, McKeever, Jason, Spence, Ian, Gains, David, Jacob, Daniel J., 2021. High-frequency monitoring of anomalous methane point sources with multispectral Sentinel-2 satellite observations. *Atmos. Meas. Tech.* 14 (4), 2771–2785. <https://doi.org/10.5194/amt-14-2771-2021>.
- Varon, Daniel J., McKeever, J., Jervis, D., Maasakkers, J.D., Pandey, S., Houweling, S., Aben, I., Scarpelli, T., Jacob, D.J., 2019. Satellite discovery of anomalously large methane point sources from Oil/Gas production. *Geophys. Res. Lett.* 46 (22), 13507–13516. <https://doi.org/10.1029/2019GL083798>.
- WRF Users' Guide, 2023. Accessed August 28, 2022. https://www2.mmm.ucar.edu/wrf/users/docs/user_guide_V3/user_guide_V3.8/contents.html.
- Zhang, Yao, Hui, Jian, Qin, Qiming, Sun, Yuanheng, Zhang, Tianyuan, Sun, Hong, Li, Minzan, 2021. Transfer-learning-based approach for leaf chlorophyll content estimation of winter wheat from hyperspectral data. *Remote Sens. Environ.* 267 (December), 112724 <https://doi.org/10.1016/j.rse.2021.112724>.
- Zhang, Zhan, Sherwin, Evan D., Varon, Daniel J., Brandt, Adam R., 2022. Detecting and quantifying methane emissions from oil and gas production: algorithm development with ground-truth calibration based on Sentinel-2 satellite imagery. *Atmos. Meas. Tech.* 15 (23), 7155–7169. <https://doi.org/10.5194/amt-15-7155-2022>.