Supporting Information for

Current potential of CH_4 emission estimates using TROPOMI in the Middle East

Mengyao Liu[1]*, Ronald van der A[1], Michiel van Weele[1], Lotte Bryan[1,2], Henk Eskes[1], Pepijn Veefkind[1, 2], Yongxue Liu[3], Xiaojuan Lin[1,4], Jos de Laat[1], Jieying Ding[1]

[1] KNMI, Royal Netherlands Meteorological Institute, De Bilt, The Netherlands

[2] Delft University of Technology, Delft, The Netherlands

[3] School of Geographic and Oceanographic Sciences, Nanjing University, Nanjing, China

[4] Department of Earth System Science, Ministry of Education Key Laboratory for Earth System Modeling, Tsinghua University, Beijing, China

* Correspondence to: Mengyao Liu (mengyao.liu@knmi.nl)

This file includes 7 pages containing

1. Part A: Build daily AOD to select reliable TROPOMI XCH_4 observations
2. Part B: Procedure to build the non-divergent wind field
3. Part C: Uncertainty estimation and sources with high confidence
4. Figures S1 and S2
5. Table S1

Part A. Build daily AOD to select reliable TROPOMI XCH_4 observations

The way to find AOD for each pixel is consisted of two steps. (1) Check if any MODIS observations of AOD at 550nm are available within a 0.2° radial area of the pixel center. Use the median value of these MODIS AODs to stand for the AOD of this pixel if the presumption is true. (2) If there are no available MODIS observations, use the AOD value from global hourly Atmospheric Composition Reanalysis 4 (EAC4) that is constrained by MODIS observation instead. The method to build the constrained daily AOD from EAC4 is as follow:

First, we build monthly mean of MODIS AOD on a grid of 0.75°, which is the same resolution as EAC4 dataset. For the grid cell without any observations in a month, we apply a spatial interpolation staring from surrounding ± 2 grid cells until finding the available MODIS observations. In the meanwhile, the monthly mean of EAC4 AOD is generated by sampling grid cells at the same overpass time as MODIS pixels and preforming the same spatial interpolation process. Thus the consistence between monthly MODIS AOD and CAMS AOD is ensured. Then the daily AOD of a TROPOMI pixel that has no MODIS observation is estimated by:

$$AOD_T = \frac{AOD^m_{MODIS}}{AOD^m_{ECA}} \cdot AOD^d_{ECA} \qquad \text{(S1)}$$

AOD_T stands for the estimated daily AOD for a TROPOMI pixel missing a MODIS observation. AOD^m_{MODIS} and AOD^m_{ECA} are monthly means of MODIS and ECA AOD after the same spatial interpolation aforementioned, respectively. AOD^d_{ECA} is the daily AOD from ECA. The calculation of Eq. (S1) is on a grid of 0.75°.

Part B. Procedure to build the non-divergent wind field

The Flux-Divergence method operates on the assumption that any divergence in the flux of a target gas results from the presence of its sources and sinks. However, the flux (F) is determined by the product of the gas (ΔXCH_4^{PBL} in this study, use V to represent the vertical column density in Eq. (1)) and the wind velocity (w).

By applying the chain rule, we obtain:

$$\nabla \cdot F = \nabla \cdot (V\vec{w}) = (\nabla V) \cdot \vec{w} + V(\nabla \cdot \vec{w}) \quad \text{(S2)}$$

A segment of the flux divergence is attributed to variations in the wind field, unrelated to the emission or deposition. At a global level, wind divergence arises from the movement between low-pressure and high-pressure zones. Additionally, larger-scale factors like mountain ranges, coastlines, and other topographic features can contribute to divergence. On a more localized level, interpolating the ECMWF wind field grid onto a regular grid may also introduce some divergence in the wind field. In order to find a flux-field for which $\nabla \cdot \vec{w}$ is minimized, and thus $\nabla(\nabla \cdot \vec{w})$ approaches zero, we followed the method proposed by Sims (2018):

Assume $\vec{w}$ contains divergence. Iterate:

1. Compute the divergence of $\overrightarrow{w_k}$: $\nabla \cdot \overrightarrow{w_k}$

2. Compute the gradient of the divergence: $\nabla(\nabla \cdot \overrightarrow{w_k})$.

3. Update vector field: $\overrightarrow{w_{k+1}} = \overrightarrow{w_k} + m\nabla(\nabla \cdot \overrightarrow{w_k})$.

m is used to scale the update. Stopping Criterion: $|\overrightarrow{w_{k+1}} - \overrightarrow{w_k}| < 10^{-5}$ or $k > 10000$.

The principle of this algorithm is based on the observation that an area has negative divergence when more flux enters a domain than leaves it. Conversely, a domain has positive divergence when it has a net outflow of particles. Therefore, divergence can be diminished by adding flow away from the area with a negative divergence, and towards areas with a positive divergence. Repeatedly adding the gradient of the divergence to a vector field achieves this, as the gradient is an arrow pointing from low-divergence areas to high-divergence areas.

This algorithm is similar to the Newton-Rhapson technique that iteratively approximates the root of a function by continuously adding the derivative of a function to an initial guess for the root. As the gradient of the divergence becomes smaller when a local minimum is reached, the update to the vector field also becomes smaller.

A side note on the implementation of the algorithm is that the units of the original vector field (F) and the vector field we use as update ($\nabla(\nabla \cdot F)$) are different. In the specific implementation of wind vector field ($\vec{w}$), the units of the original vector field are m/s, whereas the gradient of the divergence has units 1/(m·s). To compensate for this, it was chosen to multiply the gradient of the divergence by the dimensions of the grid cells Δx and Δy. Thus, the scaling factor then becomes $m \cdot \Delta x \cdot \Delta y$.

It is also important to note that the algorithm is applied to the wind field after interpolation, as interpolation also introduces small-scale divergence. In each iteration, the absolute value of the divergence is summed over the research domain. When the decline of this metric is smaller than 10^{-5}, the algorithm is terminated. The sum over the divergence lies in the order of magnitude of 10^{-1}. A limit of 10000 iterations is also implemented, in case the metric does not converge.

Part C. Uncertainty estimation and sources with high confidence

To mitigate the influence of singular values from linear regressions or time series and estimate uncertainty for each source, we use a Monte Carlo experiment to each possible source that pass the spatial filter and have been corrected by spatial correction in each year. The procedure is as follows:

(1) We randomly choose 80% of the sampling days from a time series in a year as a subset to this grid cell. We derive a new emission, E_i, and count the ratio, R_i, of the number of days that have larger normalized D_d^S to the total days of the subset.

(2) Repeat step (1) 30 times for a time series that has more than 20 sampling days while 10 times for the one that have fewer days.

(3) Take one-standard deviation of $\{R_i\}$ as an uncertainty of a grid cell and the median R of $\{R_i\}$ to be the criteria to select sources with high confidence.

If R for a grid cell is greater than 0.5, this source is regarded as a source with high confidence. We also investigate the influence of the number of the iteration, and find the result typically become stable from 20 times, and 30 times can ensure the result robust as well as the efficiency of the calculation.

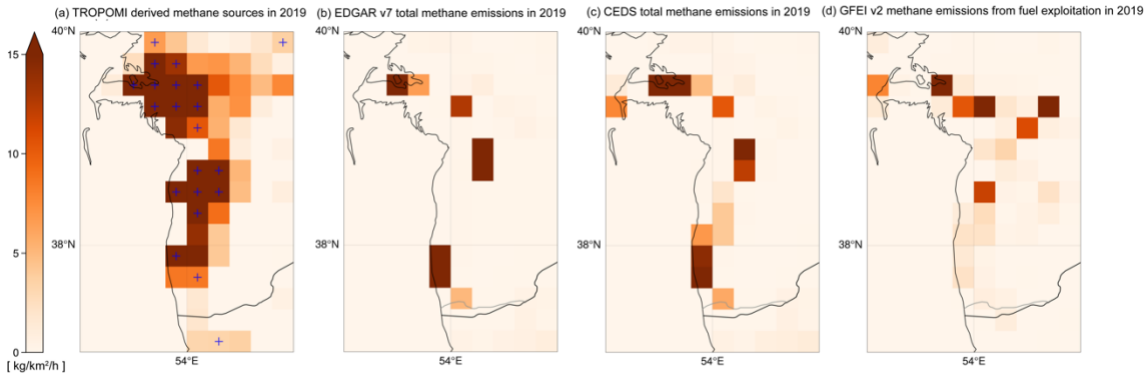

Figure S1. (a) The spatial distributions of annul methane emissions derived from TROPOMI XCH_4 over west of Turkmenistan in 2019. The sources passing the temporal filter are marked with blue "+". (b) EDGAR v7.0 averaged annual methane total emission in 2019. (c) CEDS v_2021_04_21 averaged annual total methane emissions in 2019. (d) GEFI v2 averaged annual methane emissions from the fuel exploitation in 2019.

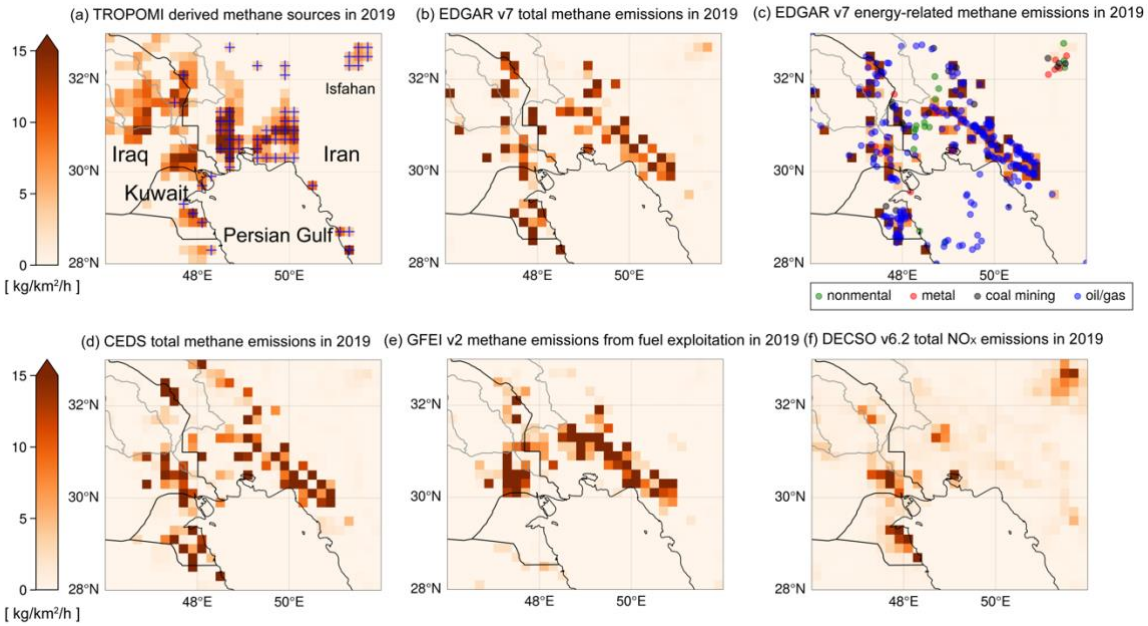

Figure S2. (a) The spatial distributions of annul methane emissions derived from TROPOMI XCH_4 over the coastal area of Iran, Iraq and Kuwait in 2019. The sources passing the temporal filter are marked with blue "+". (b) EDGAR v7.0 averaged annual methane total emission in 2019. (c)Energy-related methane emissions from EDGAR v7.0 overlapped with the industrial heat sources identified by VIIRS instrument. (d) CEDS v_2021_04_21 averaged annual total methane emissions in 2019. (e) GEFI v2 averaged annual methane emissions from the fuel exploitation in 2019. (f) Averaged annual DECSO v6.2 NO_X total emission in 2019.

Table S1. The annual emissions of EDGAR and TROPOMI

West of Turkmenistan [37.0°N, 53.0°E, 40°N, 55°E]				
kt/yr	2018	2019	2020	2021
EDGAR total	591.82	600.96	745.54	524.94
EDGAR energy	590.52	599.64	744.22	523.61
TROPOMI all sources	2825.86	2429.45	2426.45	2691.91
TROPOMI constant sources	1632.78	1544.48	1484.28	1412.61
Tehran [35.2°N, 50.6°E, 36°N, 52°E]				
kt/yr	2018	2019	2020	2021
EDGAR total	132.18	130.75	132.69	140.32
EDGAR energy	24.55	20.80	20.50	25.30
TROPOMI all sources	219.46	272.99	201.60	187.08
TROPOMI constant sources	147.20	207.39	177.82	153.39
Isfahan [32.4°N, 51.2°E, 32.8°N, 52.0°E]				
kt/yr	2018	2019	2020	2021
EDGAR total	32.00	31.56	31.81	34.77
EDGAR energy	10.35	9.46	9.24	11.63
TROPOMI all sources	128.51	113.35	111.52	126.30
TROPOMI constant sources	113.30	102.50	107.85	114.59
Iraq & Iran coastal area [29.6°N, 47.0°E, 32.6°N, 51°E]				
kt/yr	2018	2019	2020	2021
EDGAR total	4795.74	4327.36	4168.02	5101.83
EDGAR energy	4724.91	4254.67	4093.91	5026.12
TROPOMI all sources	3570.15	3483.61	3220.13	3780.61
TROPOMI constant sources	1036.85	1156.05	926.73	1280.26
Riyadh [24.4°N, 46.4°E, 25°N, 47°E]				
kt/yr	2018	2019	2020	2021
EDGAR total	263.71	267.11	275.59	279.53
EDGAR energy	65.65	64.33	68.60	67.22
TROPOMI all sources	171.25	184.10	176.59	187.86
TROPOMI constant sources	99.77	87.98	85.37	87.72