

Supporting Information: Evaluating When and Where TROPOMI XCH₄ Uncertainties Yield Consistent Flux Estimation over East, South, and Central Asia

Wei Hu^{1,2}, Jason Blake Cohen^{1,2*}, Lingxiao Lu^{1,2}, Bo Zheng^{1,2}, Pravash Tiwari^{1,2}, Simone Lolli³, Andrea Garzelli⁴, Jun Li⁵, Kai Qin^{1,2*}

¹School of Environment and Spatial Informatics, China University of Mining and Technology, Xuzhou 221116, China

²Shanxi Key Laboratory of Environmental Remote Sensing Applications, China University of Mining and Technology, Xuzhou 221116, China

³ CNR-Institute of Methodologies for Environmental Analysis (IMAA), Contrada S. Loja, Tito Scalo, 85050, Italy

⁴ Department of Information Engineering and Mathematics, University of Siena, Siena 53100, Italy

⁵ National Institute of Clean and Low Carbon Energy, Beijing 102211, China

Correspondence to: Jason B. Cohen (jasonbc@alum.mit.edu ; jasonbc@cumt.edu.cn) and Kai Qin (qinkai@cumt.edu.cn)

This file contains four parts:

Part I. Supplementary Text (Pages S2 to S3)

Part II. Supplementary Tables (Pages S4 to S5)

Part III. Supplementary Figures (Pages S6 to S9)

Part IV. Reference (Page S10)

Part I. Supplementary Text

Text S1. Methodology support

At the same time, in order to check whether the filtered data is random or selective, we examined the locations of identified strong emission sources (mining activities) in China (Maus et al., 2022, 2020; Tang and Werner, 2023; Xiao et al., 2021a, 2021b). Among 2016 distinct emission sources previously identified within China, TROPOMI available grids records values for 1987 sources, accounting for 98.6%, while Grid0.5 has values for 1871 sources, accounting for 92.8%. Coal Mines locations overlap with Grids with relative change ratio > 5 and valid grid days are shown in Figure S5. A large number of high valid days implies that a substantial amount of TROPOMI data can be used in those regions to successfully invert emissions. Looking at four of the largest such centers of high emitting coal mines, selected due to their large and relatively stable sources of methane emissions, demonstrate clearly that three of the groupings (Northwest China, Central India, and the Shanxi-Shaanxi-Inner Mongolia triangle) all exhibit a medium to large number of valid days. However, the fourth region, Southwest China and Vietnam do not exhibit zero days, but instead just a very small number of days, consistent with the fact that much of the data in this region is contaminated by clouds and/or smoke. This points out that under clear conditions, emissions can still be computed, although care is required, and that long-term averages may not be representative.

Text S2. Influence of terrain roughness

Considering that one grid contains multiple elevation values, we define the surface roughness as the standard deviation of the surface elevation within the grid cells to the $0.25^\circ \times 0.25^\circ$ grid cells of XCH₄ dataset (Vanselow et al., 2024). A regression analysis was made of the surface roughness and relative error of divergence (uniform $\pm 10\%$ perturbation), with regressions over ocean and land performed separately. In addition, considering the effects of local extreme values, grids with relative errors greater than two times their own std are not involved in the statistical analysis. For the land pixels, the R^2 is 0.03 (p-value < 0.05), which shows that there is negligible correlation between surface roughness and relative error of divergence for the land pixels, while over the ocean pixels, R^2 is 0.00 (p-value = 0.2). It can be seen that the surface roughness is not the main cause of the divergence error within the 0.25° grid.

Text S3. Different TROPOMI retrieval versions affection

In order to ensure the consistency of the data streams, only OFFL data are used in this experiment from 2019-01-01 to 2023-12-31. Considering that the time span covers a number of different retrieval versions, the data of different versions are categorized. The principle of grouping is based on improvements in the algorithm or the addition of new retrieval data. For days that are in the two group divisions, neither group contains them. The grouping and time span are shown in Table S2.

Group B updates the CH₄ cross section, the a-posteriori bias corrected CH₄ is independent of any reference data, and has new improved L1b input. Group C added the sun glint mode. Group D compared to group C, a 3rd order polynomial of surface reflectance spectral dependence was used, and has a new improved L1b input. Considering the inclusion of sun glint retrieval (Lorente et al., 2022) over ocean areas since group C, four groups ABCD compare over land and two groups CD compare over ocean. Ocean pixels were filter by “processing_quality_flags” with “sun_glint_warning”. In addition, it is worth noting that this study filtered the sun-glint grid not only for oceans, but also for inland lakes.

We counted the grids that met the conditions (i.e., relative change ratio ≤ 0.1 , obey uniform 10% perturbation), and plotted the spatial distribution of the valid grid ratio of each group. Since there are fewer valid grids, the color bar is set at 0-0.3. As can be seen from the figure, land group B has a higher valid grid ratio than other land groups, which has only 131 days, while the rest of the groups are all above 200 days. However, land group D also has a higher valid grid ratio than land group C, and its valid days are higher than group C, indicating that the range of days covered by the version is not the reason for the valid grid ratio. Overall, there are higher valid grid ratio in the northwest of China (Xinjiang, Shaanxi, etc.), the southwestern part of Mongolia, and the eastern part of India. Conversely, in the southern part of China, as well as on the sea surface, the valid grid ratio is lower. Considering that the number of days in group B is less than that of other groups, we are more conservative in concluding that the different inversion versions do not have much impact on our method.

The percentage of grids which satisfy the relative error of divergence ranges from 23.4%-25.2% over land to 20.1%-23.9% over the ocean as shown in Figure S4. The reason why the valid grid ratio of ocean is similar to that of land is because the grids selected by the sun-glnt condition include inland lakes, and the valid ratio of these grids is higher than that of the ocean, so the average value is raised. Besides, what can be seen is there is no significant distinct between different retrieval version.

Part II. Supplementary Tables

Table S1. Valid grid ratio of different random number distributions and different wind fields.

(350-975hpa)	Uniform			Normal (mean=0, std=0.036/0.0036)		Lognormal (mean=-2.5, sigma=0.5)	Beta (a=2, b=5)
Date	$\pm 10\%$	$\pm 5\%$	$\pm 1\%$	$\pm 10\%$	$\pm 1\%$	$\pm 10\%$	$\pm 10\%$
2019-01-01	23.4%	23.4%	23.4%	34.3%	34.4%	18.9%	35.1%
2020-03-01	29.8%	29.8%	29.8%	42.8%	42.8%	23.8%	43.3%
2021-05-01	30.6%	30.7%	30.6%	43.0%	43.0%	24.2%	43.6%
2022-07-01	20.8%	20.8%	20.8%	31.2%	31.2%	17.0%	32.2%
2023-09-01	23.0%	23.0%	22.9%	33.6%	33.6%	18.8%	34.6%
2023-11-01	28.2%	28.2%	28.2%	40.5%	40.5%	22.6%	41.1%
10m	Uniform			Normal (mean=0, std=0.03/0.003)		Lognormal (mean=-2.5, sigma=0.5)	Beta (a=2, b=5)
Date	$\pm 10\%$	$\pm 5\%$	$\pm 1\%$	$\pm 10\%$	$\pm 1\%$	$\pm 10\%$	$\pm 10\%$
2019-01-01	30.1%	30.1%	30.1%	43.4%	43.4%	23.9%	43.9%
2020-03-01	34.5%	34.5%	34.5%	47.4%	47.4%	28.0%	47.9%
2021-05-01	27.1%	27.1%	27.1%	39.0%	38.9%	21.4%	39.9%
2022-07-01	20.8%	20.8%	20.8%	30.7%	30.7%	17.2%	31.8%
2023-09-01	25.1%	25.1%	25.1%	36.0%	36.0%	20.6%	36.9%
2023-11-01	26.8%	26.8%	26.8%	39.2%	39.2%	21.5%	40.1%
100m	Uniform			Normal (mean=0, std=0.03/0.003)		Lognormal (mean=-2.5, sigma=0.5)	Beta (a=2, b=5)
Date	$\pm 10\%$	$\pm 5\%$	$\pm 1\%$	$\pm 10\%$	$\pm 1\%$	$\pm 10\%$	$\pm 10\%$
2019-01-01	29.2%	29.2%	29.2%	42.4%	42.4%	23.2%	43.0%
2020-03-01	32.6%	32.6%	32.6%	45.4%	45.4%	26.5%	45.9%
2021-05-01	26.0%	26.0%	26.0%	37.8%	37.8%	20.5%	38.7%
2022-07-01	20.0%	20.0%	20.0%	29.7%	29.7%	16.5%	30.9%
2023-09-01	24.1%	24.1%	24.1%	35.0%	35.0%	19.8%	35.9%
2023-11-01	26.0%	26.0%	26.0%	38.3%	38.3%	20.8%	39.2%

Table S2. Each grouping with time span and specific retrieval version.

Groups	Time range	Version
A	[2019-01-01, 2021-07-01)	01.02.02, 01.03.00, 01.03.01, 01.03.02, 01.04.00
B	(2021-07-01, 2021-11-14)	02.02.00
C	(2021-11-14, 2022-07-17)	02.03.01
D	(2022-07-17, 2023-11-26)	02.04.00, 02.05.00

95

Table S3. Statistical information of relative change ratio of each grouping (uniform $\pm 10\%$).

Group	Days	Region	Mean	Median	Q1	Q3	Max	Std
A	899	Land	23.4	23.5	20.2	26.7	38.7	5.0
B	131	Land	24.8	24.5	21.9	27.1	34.9	3.9
C	245	Land	25.2	25.5	22.1	27.9	41.6	5.2
D	424	Land	24.0	23.8	21.6	27.0	34.9	4.3
C	245	Ocean	20.1	17.3	12.8	25.5	72.5	11.9
D	424	Ocean	23.9	23.3	16.7	29.5	94.2	11.8

Part III. Supplementary Figures

100

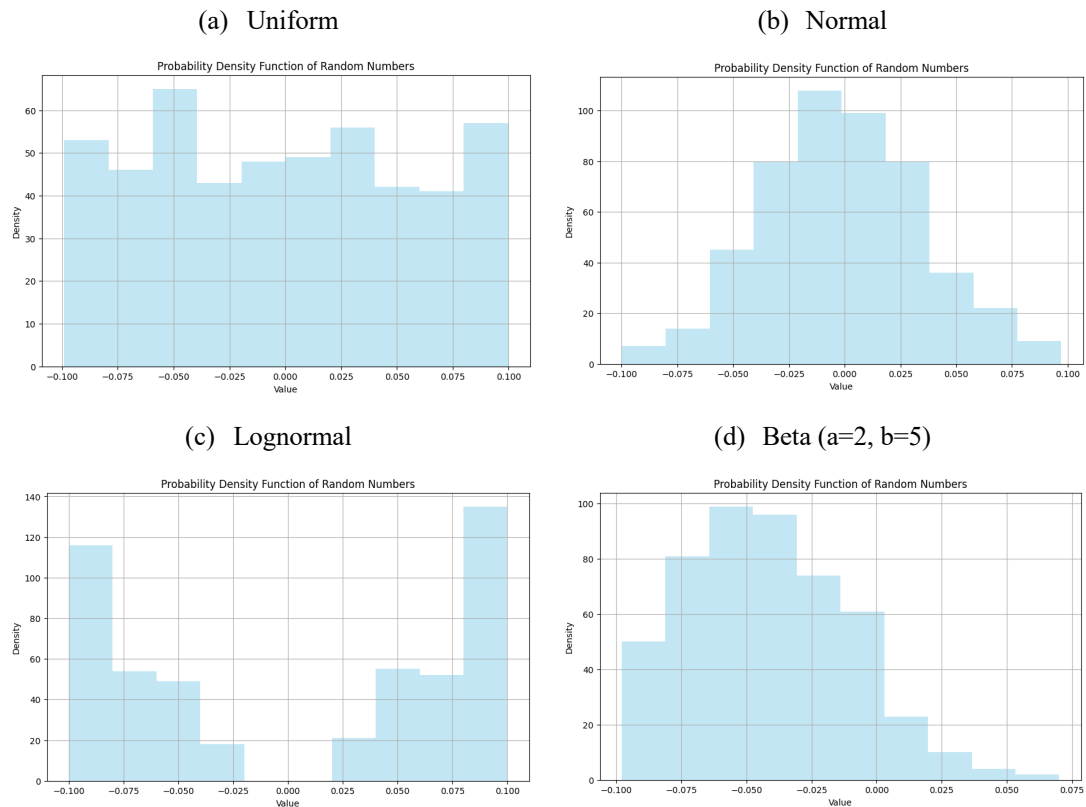


Figure S1. Random number distribution (1% perturbation case).

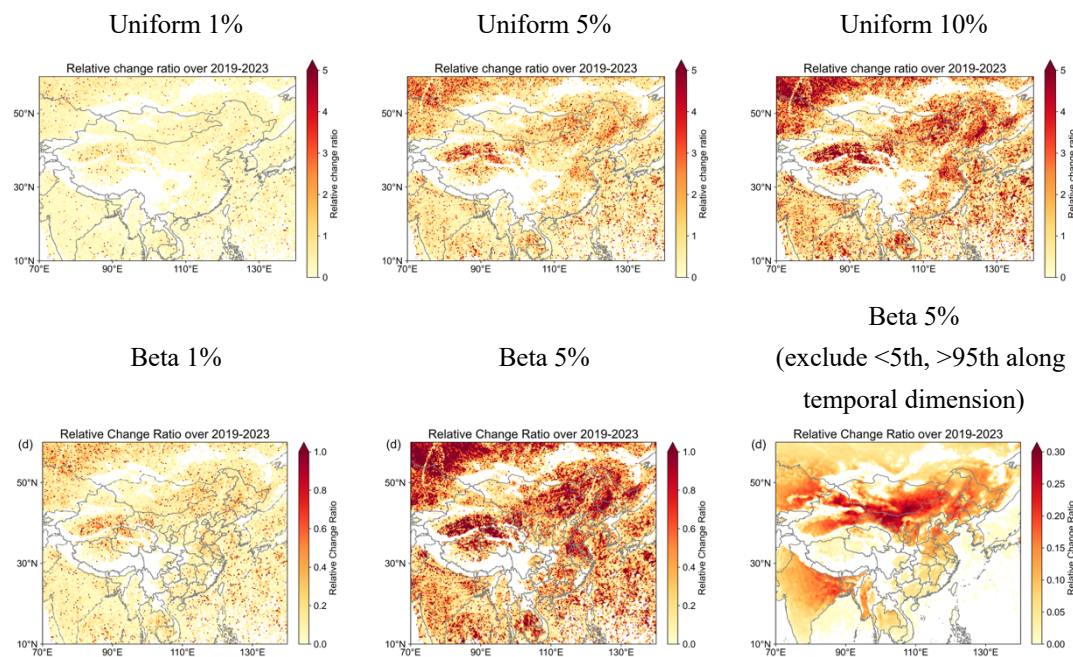
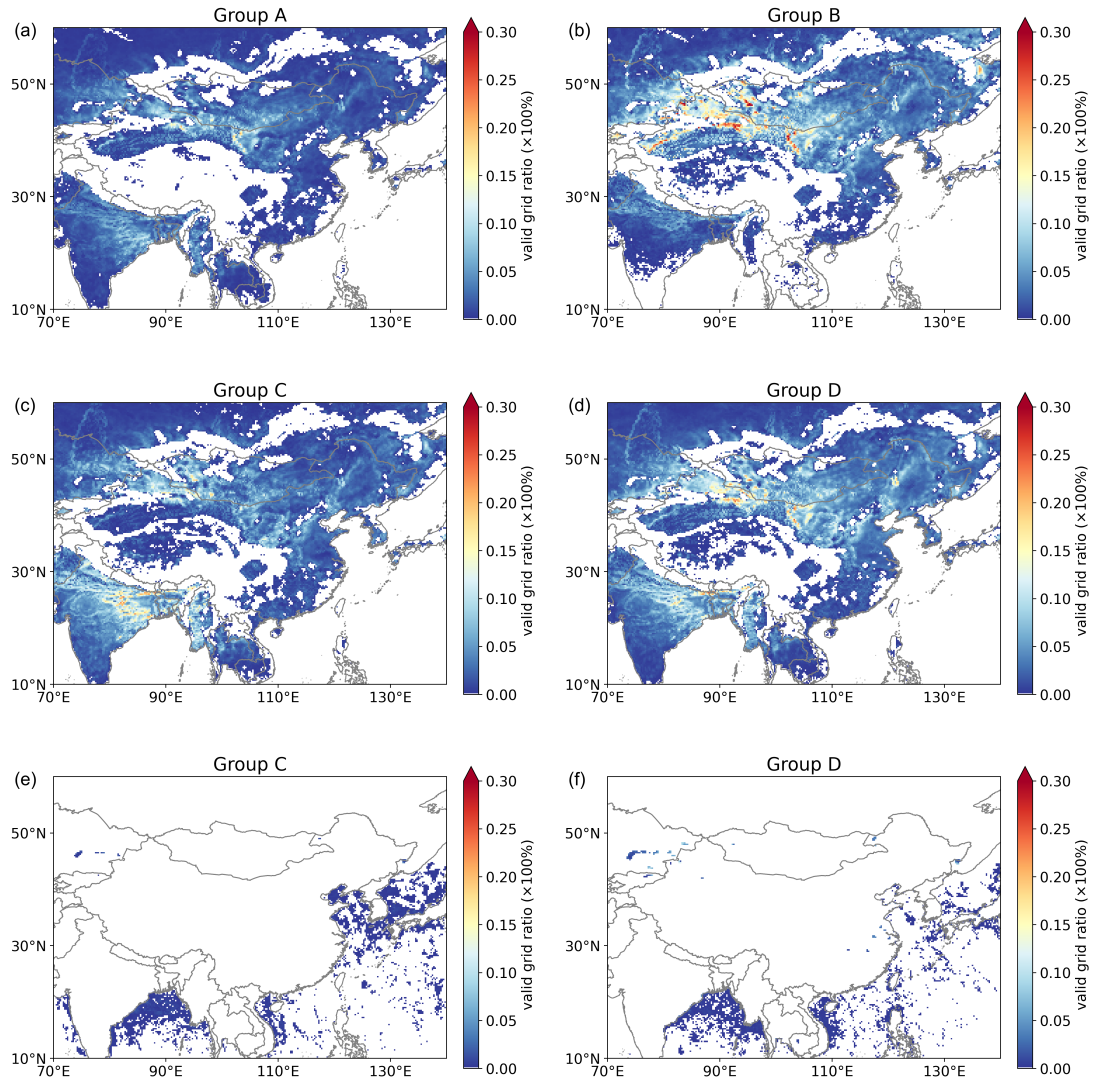
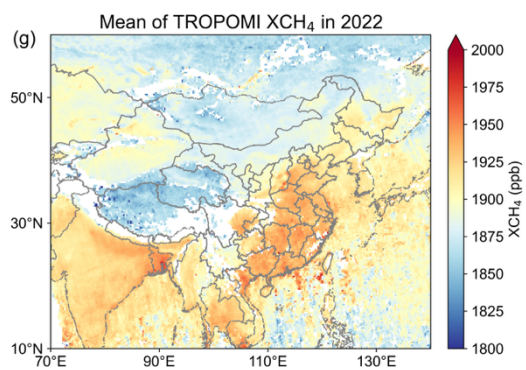
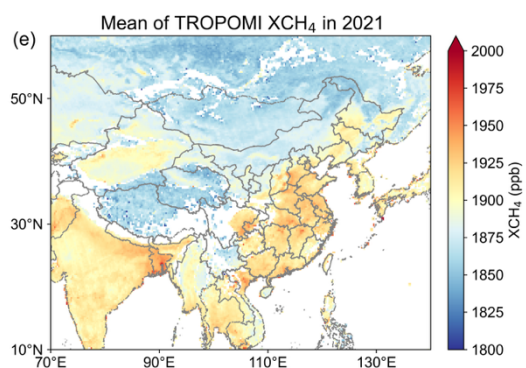
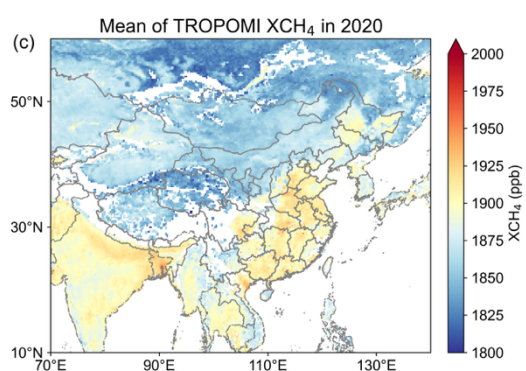
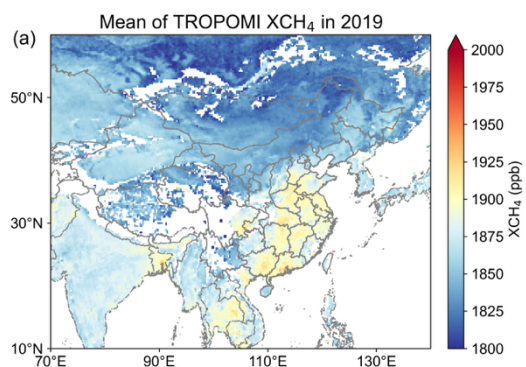


Figure S2. Relative change ratio over 2019-2023 with different x% perturbation.

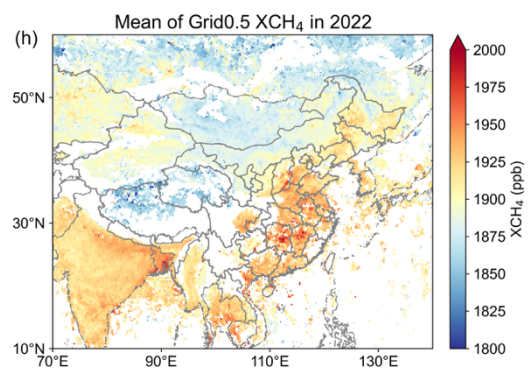
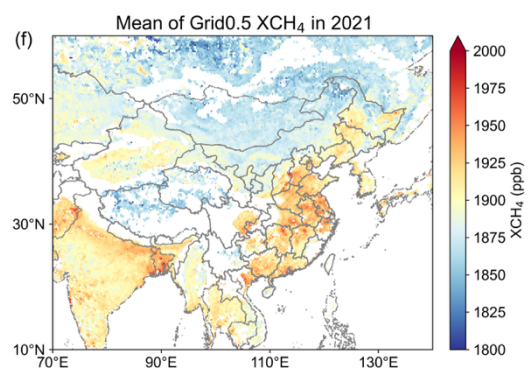
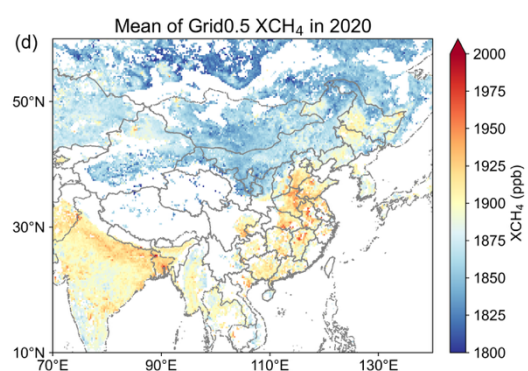
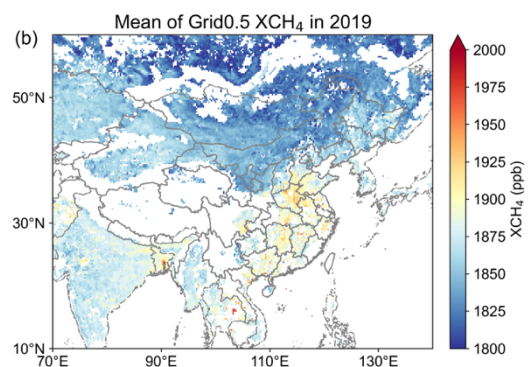


105 **Figure S3.** Spatial distribution of valid grid ratio over land and ocean in different group (uniform $\pm 10\%$ perturbation).

TROPOMI



Grid 0.5



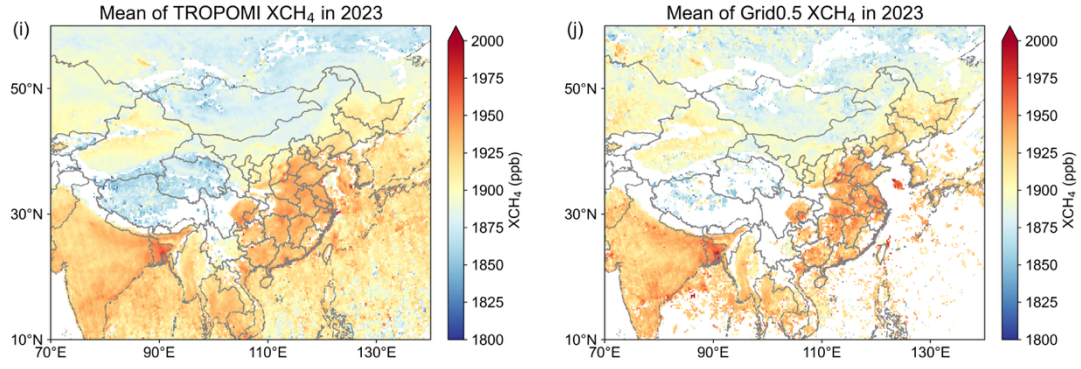


Figure S4. Mean of TROPOMI (left) and Grid0.5 (right) XCH₄ among 2019-2023 (beta perturbation).

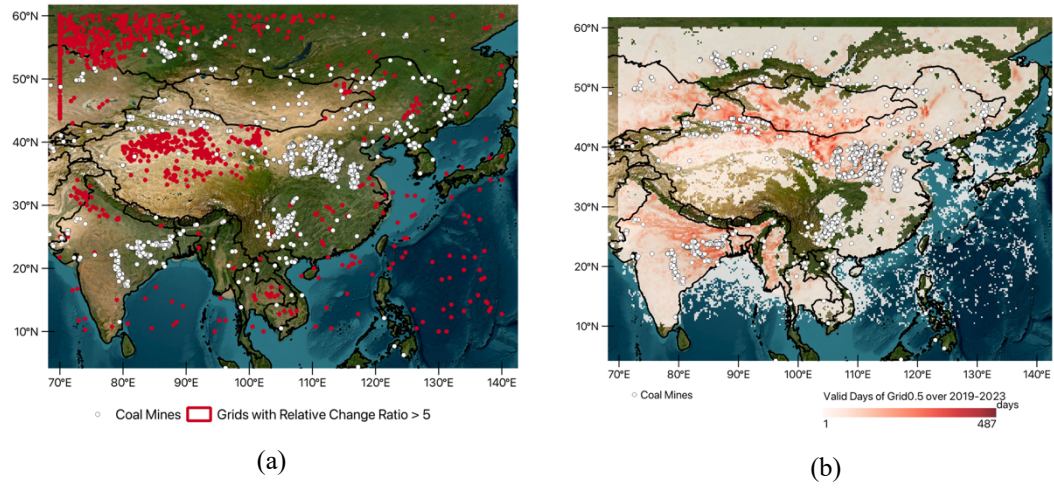


Figure S5. Coal Mines locations overlap with Grids with relative change ratio > 5 (a) and with valid days of Grid0.5 (b).

References

- Lorente, A., Borsdorff, T., Martinez-Velarte, M.C., Butz, A., Hasekamp, O.P., Wu, L., Landgraf, J., 2022. Evaluation of the methane full-physics retrieval applied to TROPOMI ocean sun glint measurements. *Atmospheric Measurement Techniques* 15, 6585–6603.
<https://doi.org/10.5194/amt-15-6585-2022>
- Maus, V., Giljum, S., Da Silva, D.M., Gutschlhofer, J., Da Rosa, R.P., Luckeneder, S., Gass, S.L.B., Lieber, M., McCallum, I., 2022. An update on global mining land use. *Sci Data* 9, 433.
<https://doi.org/10.1038/s41597-022-01547-4>
- Maus, V., Giljum, S., Gutschlhofer, J., Da Silva, D.M., Probst, M., Gass, S.L.B., Luckeneder, S., Lieber, M., McCallum, I., 2020. A global-scale data set of mining areas. *Sci Data* 7, 289.
<https://doi.org/10.1038/s41597-020-00624-w>
- Tang, L., Werner, T.T., 2023. Global mining footprint mapped from high-resolution satellite imagery. *Commun Earth Environ* 4, 134. <https://doi.org/10.1038/s43247-023-00805-6>
- Vanselow, S., Schneising, O., Buchwitz, M., Reuter, M., Bovensmann, H., Boesch, H., Burrows, J.P., 2024. Automated detection of regions with persistently enhanced methane concentrations using Sentinel-5 Precursor satellite data. *Atmos. Chem. Phys.* 24, 10441–10473.
<https://doi.org/10.5194/acp-24-10441-2024>
- Xiao, W., Chen, W., Deng, X., 2021a. Coupling and coordination of coal mining intensity and social-ecological resilience in China. *Ecological Indicators* 131, 108167.
<https://doi.org/10.1016/j.ecolind.2021.108167>
- Xiao, W., Zheng, W., Zhao, Y., Chen, J., Hu, Z., 2021b. Examining the relationship between coal mining subsidence and crop failure in plains with a high underground water table. *J Soils Sediments* 21, 2908–2921. <https://doi.org/10.1007/s11368-021-02991-2>