

Dear Editor and Reviewers,

Thank you for your careful evaluation of our manuscript and for the constructive and insightful comments. We greatly appreciate the time and effort devoted to reviewing our work. Your suggestions have helped us improve the scientific clarity, methodological explanation, interpretation of the results, and overall presentation of the manuscript.

We have carefully considered each comment and revised the manuscript accordingly. The principal revisions include clarifying the study domain and source categories; expanding the discussion of biomass burning, residential fuel use, and persistent underground coal fires; clarifying the propagation and interpretation of observational uncertainty; adding seasonally adjusted trend analyses and quantitative measures of emission and hotspot changes; moderating claims concerning independent validation, inventory differences, and policy effects; strengthening the discussion of limitations and source attribution; and revising the figures, captions, terminology, and language throughout the manuscript.

A detailed point-by-point response is provided below, together with the relevant changes made in the revised manuscript and their corresponding line numbers. The reviewers' comments are reproduced in full, followed by our responses and the revised text.

We hope these revisions and explanations adequately address all raised concerns, and we look forward to your further evaluation of our revised manuscript.

Best Regards,

Jason Cohen

Response to Reviewer #1:

We sincerely thank the reviewer for the careful evaluation of our manuscript and for the constructive and insightful comments. We greatly appreciate the time and effort devoted to reviewing our work. The comments have helped us improve both the scientific presentation and the clarity of the manuscript. We have carefully considered all suggestions and revised the manuscript accordingly. In this response, your original comments are completely unedited and are given in **yellow highlight**, our responses are given in **blue text** and updates to the paper are given in **purple text**.

The revisions include: (1) clarifying the potential influence of biomass burning and residential solid-fuel combustion; (2) distinguishing the study domain from the facilities used in the source-category analysis; (3) clarifying the propagation of observational uncertainty into CO emission estimates and distinguishing this contribution from total emission uncertainty; (4) moderating claims regarding independent validation, EDGAR bias, policy effectiveness, and coal-fire seasonality; (5) adding a seasonally adjusted trend analysis with HAC standard errors and a seasonal Mann-Kendall robustness test; (6) quantifying the 37.9% endpoint decline and the 40.8% reduction in hotspot intensity; and (7) revising the text, captions, terminology, and statistical interpretation throughout the manuscript.

This study quantified daily gridded CO emissions in Central Asia for 2019-2024 and assessed CO trends and hotspots. This study provided actionable insights for inventory improvement and migration strategies. This topic is interesting and meaningful to readers. I only have some small suggestions.

Special comments:

1 Lines 35-37. CO is also emitted by biomass burning. What about biomass burning emissions in the research region? For example, anthropogenic biomass burning for cooking and heating from wood and yak dung in rural areas.

Response: We appreciate the reviewer's valuable suggestion. Biomass burning and residential solid-fuel combustion may contribute to regional CO concentrations and seasonal variability, particularly during the heating season. MFIEF estimates daily grid-scale CO emissions from atmospheric observations without explicitly separating individual source sectors. Contributions from these sources may therefore be included in the retrieved emissions. We have acknowledged these contributions in Section 2.1 and Section 3.1. The Discussion also clarifies that differences between MFIEF and EDGAR may reflect differences in source coverage, spatial and temporal allocation, and estimation methodology. The present analysis characterizes grid-scale emission patterns and selected source environments; it does not quantify the separate contributions of biomass burning and fossil-fuel combustion. The perturbation analysis quantifies the contribution of the assumed observational uncertainties to the uncertainty of the grid-scale emission estimates. It does not represent the total uncertainty of the inferred emissions. One strength to our approach however is a more robust uncertainty

in the emissions on the grid and how the uncertainty changes over time as well (Qin et al., 2023).

Revised manuscript, Sect. 2.1, lines 98-100: “The region also contains rural areas where residential solid fuel combustion and occasional biomass burning may contribute to seasonal CO variability.”

Revised manuscript, Sect. 3.1, lines 256-258: “Elevated CO is also found in parts of northwestern Kazakhstan, where biomass burning and residential solid fuel combustion may contribute to seasonal variations in the regional background CO, particularly during the heating season.”

Revised manuscript, Sect. 4 Discussion, lines 438–448: “These statistics characterize the numerical differences between the datasets. EDGAR estimates anthropogenic emissions using sector-specific activity data and emission factors on a monthly time scale and locates them based on a priori information, whereas MFIEF derives a grid-scale signal constrained by atmospheric observations. These differences may arise from additional or missing emission sources, inaccuracies in the spatial allocation of sources, or differences between actual emissions and bottom-up estimates. Differences in atmospheric transport, satellite sampling and retrievals, and inversion assumptions including chemistry, physics, and long-range transport may all contribute to the observed variability and uncertainty in the top-down estimate. For additional magnitude context, Wang et al. (2025) reported observation-constrained CO emission fluxes ranging approximately from 0 to $30 \mu\text{g m}^{-2} \text{s}^{-1}$, with values in relatively clean regions generally below $6 \mu\text{g m}^{-2} \text{s}^{-1}$. The regional mean fluxes obtained in this study ranged from 3.06 to $4.93 \mu\text{g m}^{-2} \text{s}^{-1}$, with peak values exceeding $10 \mu\text{g m}^{-2} \text{s}^{-1}$, and therefore are of a comparable order of magnitude and spread. This numerical comparison provides a broad reference for the inferred emission magnitudes.”

2 Figure 1. Strangely, there are no power/steel plants or coal bases in the Central Asian countries. I also suggest adding the province or country names in Figure 1; otherwise, readers may be confused by the regions mentioned in the text.

Response: We thank the reviewer for this helpful suggestion. We agree that the original Figure 1 could give the impression that the mapped industrial facilities represent the entire Central Asian study domain. The inversion covers 35° – 50° N and 70° – 100° E, whereas the source-category analysis uses facilities in Xinjiang and adjacent parts of Gansu and Qinghai. To make the category-level comparison temporally consistent, we further restricted the industrial sample to facilities that operated continuously throughout 2019–2024. After excluding intermittently operating facilities and non-coal-fired power plants, the final sample comprised 53 coal-fired power plants and 4 steel plants. We revised the Figure 1 caption and the corresponding description in Section 2.1 to state this scope explicitly and retained the geographic labels added to improve readability.

Revised manuscript, Sect. 2.1, lines 103–107: “Fig. 1 illustrates the study area together with the coal bases, steel plants, and power plants included in the source-category analysis. These facilities are located in Xinjiang and adjacent parts of Gansu and Qinghai, whereas the satellite-based inversion was conducted over the entire study

domain. For the category-level temporal analysis, only steel plants and coal-fired power plants that operated continuously throughout 2019–2024 were retained, resulting in a fixed sample of 4 steel plants and 53 coal-fired power plants.”

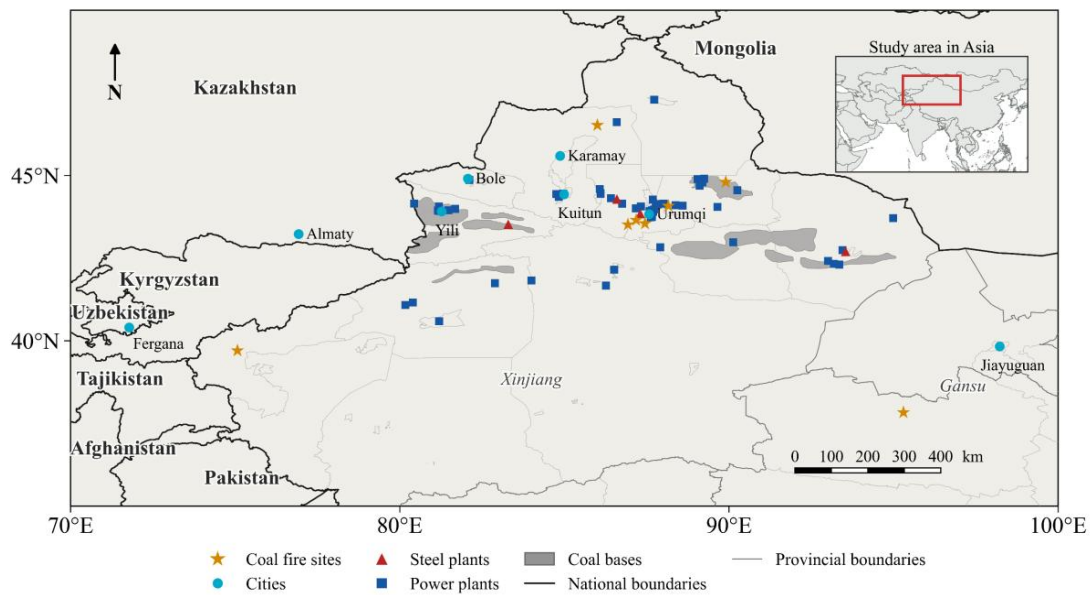


Figure 1. Study area (35°–50° N, 70°–100° E). The mapped coal bases and industrial facilities provide geographic context for the source-category analysis and are located in Xinjiang and adjacent parts of Gansu and Qinghai. For the 2019–2024 category-level statistics, only continuously operating facilities were retained, comprising 4 steel plants and 53 coal-fired power plants.

3 Line 195-200. Equation 2 was used to calculate the change in CO emissions. But uncertainty is usually not defined as the change of certain values.

Response: We appreciate the reviewer's valuable comment. We have considered that the observed CO column loadings have an amount of uncertainty and have perturbed these observations and then run these through MFIEF in an attempt to quantify the impact that this uncertainty has on the emissions range. This perturbation approach provides a quantitative estimate of how the assumed observational uncertainties propagate into the inverted emissions. We agree that it does not represent the total uncertainty of the CO emission estimates, because uncertainties in meteorology, transport, chemistry, temporal sampling, and other model assumptions are not fully included. We have therefore revised the manuscript to describe the calculated range specifically as the observational contribution to emission uncertainty and as a component of the total top-down emission uncertainty. For more details on the method as applied to NO_x see Lu et al., (2025) and as applied to CH₄ see Zheng et al., (2026) . Revised manuscript, Sect. 2.4, lines 213-214: “Accordingly, this study propagates the assumed observational uncertainties through MFIEF and quantifies the corresponding range of variation in the inverted emissions.”

Revised manuscript, Sect. 2.4, lines 222–229: “The calculated changes define the emission uncertainty range associated with propagation of the perturbed observational inputs. Specifically, the relative emission response was calculated by Eq. (2). ... The

distribution of these responses quantifies the observational contribution to emission uncertainty, and forms a minimum of the total top-down emissions uncertainty. Grid cells exhibiting responses larger than the adopted threshold are classified as exceeding this uncertainty range and are excluded from the final emission product.”

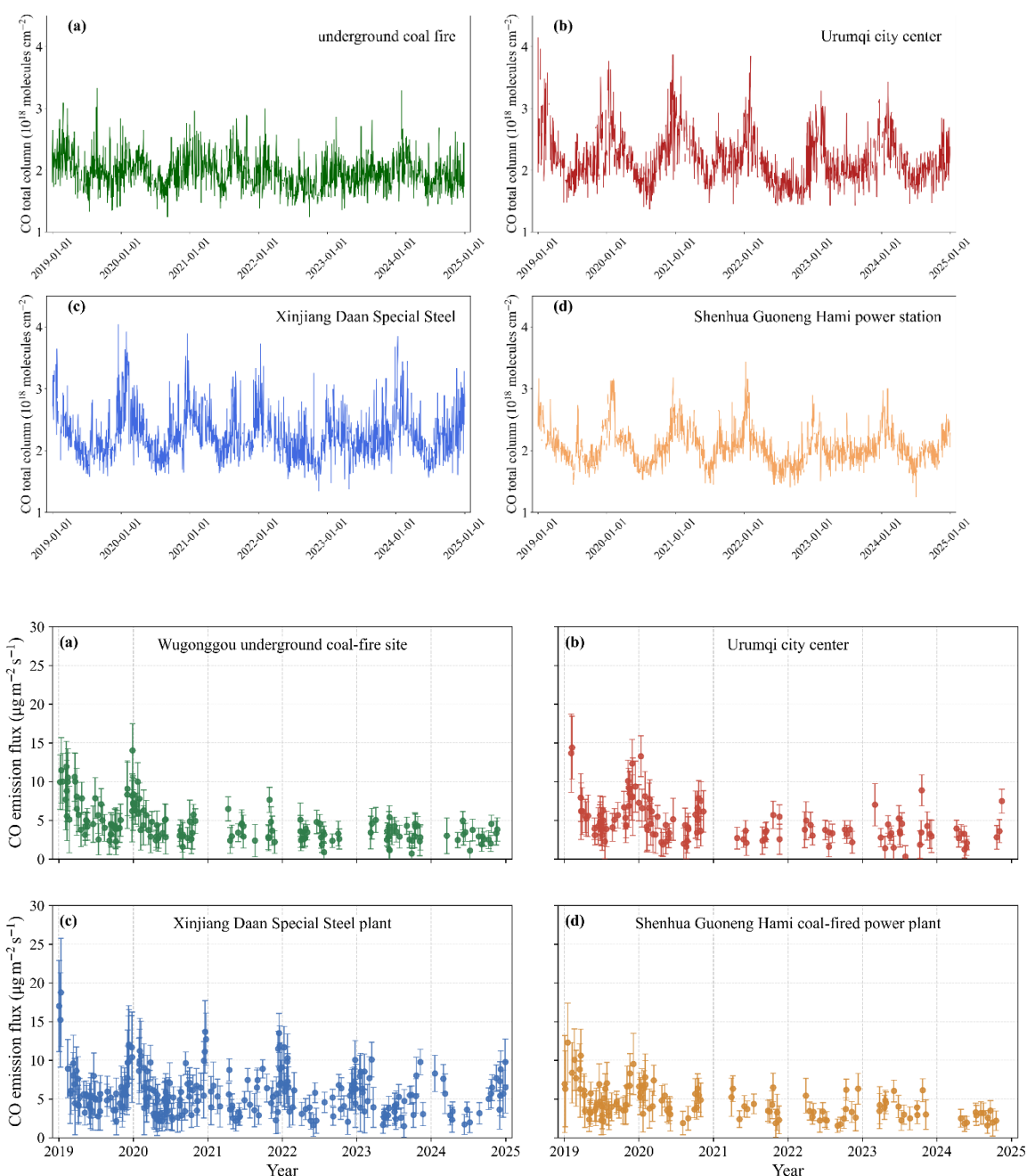
4 Figure 4A. Why coal fire also show a clear seasonal variation in CO emissions?

Response: We thank the reviewer for this comment. Figure 4a shows satellite-observed CO total columns, whereas Figure 8a shows inverted CO emission fluxes. Seasonal variation in the column does not necessarily imply the same seasonal variation in local coal-fire emissions, because the column also reflects atmospheric transport, chemical loss, and contributions from surrounding sources. We have clarified this distinction in Section 3.1. The Wugonggou CO column exhibits weaker seasonality than the other selected locations, with a monthly amplitude of 0.32×10^{18} molecules cm^{-2} compared with $0.71\text{--}0.77 \times 10^{18}$ molecules cm^{-2} at the other three locations. After re-examining the emission-flux series, we found that the strongest high-emission episodes occurred near the beginning of the study period and subsequently weakened. Wintertime sampling was sparse in some years, so the seasonal differences are interpreted as descriptive temporal patterns rather than independent trends. Wugonggou is an uncontrolled underground coal-fire source rather than an engineered industrial stack, and its decline is not used as evidence of industrial emission-control policy effectiveness. Some published studies suggest that colder conditions and oxygen availability can affect underground combustion (Song et al., 2020; Zhu and Wen, 2023), but the present satellite record cannot establish that mechanism independently.

Revised manuscript, Sect. 3.1, lines 276–279: “The Wugonggou grid exhibited weaker seasonality, with a monthly amplitude of 0.32×10^{18} molecules cm^{-2} , compared with $0.71\text{--}0.77 \times 10^{18}$ molecules cm^{-2} at the other three locations. This weaker variation is consistent with the relatively persistent nature of underground coal combustion and the weaker influence of seasonally varying urban and industrial activities in the surrounding area.”

Revised manuscript, Sect. 3.2, lines 367–375: “At Wugonggou, for example, the maximum daily emission flux of $14.0 \mu\text{g m}^{-2} \text{s}^{-1}$ occurred in December 2019, while the maximum observed values in subsequent years declined to 9.98, 7.65, 5.09, 5.42, and $4.97 \mu\text{g m}^{-2} \text{s}^{-1}$ from 2020 to 2024, respectively. Its annual mean similarly decreased from $6.38 \mu\text{g m}^{-2} \text{s}^{-1}$ in 2019 to $3.03 \mu\text{g m}^{-2} \text{s}^{-1}$ in 2024.”

Revised manuscript, Sect. 3.2, lines 376–378: “In particular, Wugonggou is an uncontrolled underground coal-fire source rather than an engineered industrial stack, and its temporal evolution should not be interpreted in terms of the same emission-control mechanisms as the urban and industrial grid cells.”



5 Figure 8. It is easy to imagine the CO emission unit in coal fire and urban centers as they are area sources. But for power plants, which are point sources, how do we calculate the emission? Or using the pixel data for the power plant location?

Response: We thank the reviewer for this important clarification. The MFIEF framework produces gridded CO emission fluxes expressed in $\mu\text{g m}^{-2} \text{s}^{-1}$ averaged over the entire TROPOMI pixel rather than facility-level point-source emission rates.

For the power plant analysis, we extracted the emission value from the grid cell containing the geographic location of each power plant. Therefore, the values shown in Figure 8 represent the atmospheric CO emission enhancement within the corresponding grid cell associated with the power plant location, rather than direct measurements of the stack emission rate. Because the spatial resolution of the TROPOMI-based

inversion is coarser than an individual facility, the extracted value represents the net emission flux from all sources within the grid cell containing the facility. It should therefore be interpreted as a grid-scale estimate associated with the facility location, rather than as a direct estimate of the stack emission rate.

Because the TROPOMI-based inversion has a spatial resolution coarser than individual facilities, the derived emissions represent the combined influence of the power plant and nearby sources within the grid cell, as well as rapid mixing and adjustments occurring within the grid scale. We have clarified this point in the revised manuscript.

Revised manuscript, Sect. 3.2, lines 362–363: “Because the inversion is performed at a spatial resolution of $0.05^\circ \times 0.05^\circ$, these estimates represent the net emissions from all sources within the corresponding grid cell rather than emissions from an individual facility.”

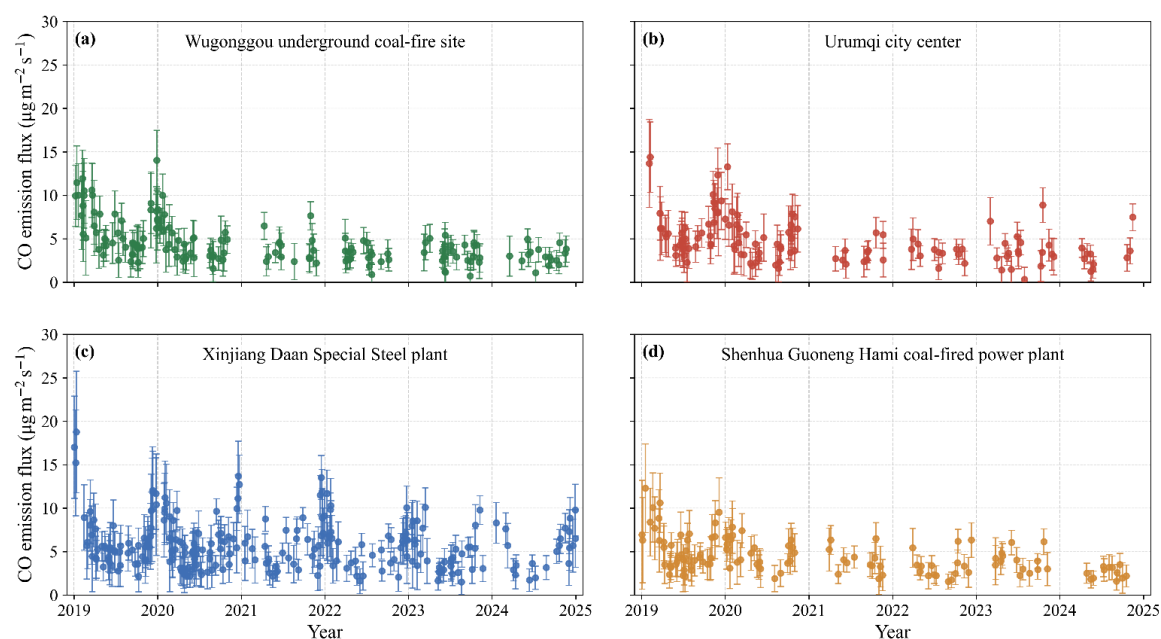


Figure 8. Temporal variations in MFIEF-derived CO emission fluxes at four representative locations from January 2019 to December 2024: (a) the Wugonggou underground coal-fire site, (b) Urumqi city center, (c) the Xinjiang Daan Special Steel plant, and (d) the Shenhua Guoneng Hami coal-fired power plant. Markers represent the mean emission fluxes, and error bars indicate ± 1 standard deviation. The values were extracted from the 0.05° grid cell containing each location.”

6 Table 1. Many cities, coal-fired sites, steel plants, and power plants are shown in Figure 1. So the data summarized in Table 1 are calculated from 1 typical site or averages of all the same types of locations?

Response: We thank the reviewer for this clarification. Table 1 reports category-level distributions rather than values from a single typical site, whereas Figure 8 presents representative individual locations. We agree that the previous description did not state the industrial-facility screening criterion clearly.

For temporal comparability, we re-screened the industrial facilities and retained only coal-fired power plants and steel plants that operated continuously throughout 2019 –

2024. Facilities that were not in operation during the complete six-year period, together with non-coal-fired power plants, were excluded. This screening yielded 53 coal-fired power plants and 4 steel plants. Table 1 and the associated text were recalculated using this fixed sample.

Using the screened sample, mean grid-scale CO emission fluxes decreased from 4.94 to 3.02 $\mu\text{g m}^{-2} \text{s}^{-1}$ at coal-fired power-plant locations and from 6.22 to 4.08 $\mu\text{g m}^{-2} \text{s}^{-1}$ at steel-plant locations between 2019 and 2024. The corresponding endpoint reductions were 38.9% and 34.4%, respectively. These values and the complete category statistics are reported in Section 3.2 and Table 1.

Revised manuscript, Sect. 3.2, lines 385–391: “Table 1 includes industrial facilities that operated continuously throughout 2019–2024. The screened sample contains 53 coal-fired power plants and 4 steel plants; the city-center and underground coal-fire categories each contain 8 locations. Based on grid-scale estimates from this fixed sample, mean emissions decreased from 6.47 to 3.74 $\mu\text{g m}^{-2} \text{s}^{-1}$ at city-center locations (42.2%), from 7.09 to 4.18 $\mu\text{g m}^{-2} \text{s}^{-1}$ at underground coal fire locations (41.0%), from 6.22 to 4.08 $\mu\text{g m}^{-2} \text{s}^{-1}$ at steel-plant locations (34.4%), and from 4.94 to 3.02 $\mu\text{g m}^{-2} \text{s}^{-1}$ at coal-fired power-plant locations (38.9%) between 2019 and 2024. The 90th-percentile values also decreased, including from 11.00 to 6.58 $\mu\text{g m}^{-2} \text{s}^{-1}$ at steel-plant locations and from 8.20 to 4.69 $\mu\text{g m}^{-2} \text{s}^{-1}$ at coal-fired power-plant locations.”

Revised manuscript, Sect. 3.2, lines 392–393: “Table 1. Annual statistics of grid-scale CO emission fluxes for selected source categories from 2019 to 2024. Steel-plant and power-plant samples include only facilities operating continuously across all six years (unit: $\mu\text{g m}^{-2} \text{s}^{-1}$).”

7 Typo errors: please check the text again during revision. Some are listed below.

a) Line 82: In addition to Kazakhstan, Kyrgyzstan, and Uzbekistan, some parts of other countries like Tajikistan, Afghanistan, and Pakistan are also shown in Figure 1. b) two “and then resampled” in line 100.

c) Line 147. (Li et al., 2025) to Li et al., (2025)

d) Use Figure or Fig. in the manuscript, not all of them.

Response: We sincerely thank the reviewer for carefully identifying these typographical and formatting issues. All of the specific errors pointed out by the reviewer have been corrected in the revised manuscript, including the study area description, duplicated wording, citation formatting, and figure citation style. We revised figure citations to use “Fig.” in the main text and “Figure” in figure captions, and conducted an additional proofreading pass to correct the identified grammatical, terminological, citation-formatting, and cross-reference issues.

Revised manuscript, Sect. 2.1, lines 95–96: “The study area spans from 35° N to 50° N latitude and 70° E to 100° E longitude, including Xinjiang and parts of Kazakhstan, Kyrgyzstan, Uzbekistan, Tajikistan, Afghanistan, Pakistan, and adjacent regions.”

Revised manuscript, Sect. 2.2, lines 119–121: “In this work, the Level 3 OFFL data product from Google Earth Engine (GEE) is filtered to ensure the quality assurance flag is above 0.5, and resampled using the Weighted Polygonal Remote Sensing Toolkit HARP (<http://stcorp.github.io/harp/doc/>).”

Revised manuscript, Sect. 2.3, lines 173–175: “This mass conservation equation for CO provides a framework for understanding and quantifying the changes in CO columns in the atmosphere, considering the various processes that influence its distribution, following Li et al. (2025).”

References

- Lu, L., Cohen, J.B., Qin, K., Li, X., He, Q., 2025. Identifying missing sources and reducing NO_x emissions uncertainty over China using daily satellite data and a mass-conserving method. *Atmos. Chem. Phys.* 25, 2291–2309. <https://doi.org/10.5194/acp-25-2291-2025>
- Qin, K., Lu, L., Liu, J., He, Q., Shi, J., Deng, W., Wang, S., Cohen, J.B., 2023. Model-free daily inversion of NO_x emissions using TROPOMI (MCMFE-NO_x) and its uncertainty: Declining regulated emissions and growth of new sources. *Remote Sens. Environ.* 295, 113720. <https://doi.org/10.1016/j.rse.2023.113720>
- Song, Z., Huang, X., Kuenzer, C., Zhu, H., Jiang, J., Pan, X., Zhong, X., 2020. Chimney effect induced by smoldering fire in a U-shaped porous channel: A governing mechanism of the persistent underground coal fires. *Process Saf. Environ. Prot.* 136, 136–147. <https://doi.org/10.1016/j.psep.2020.01.029>
- Zheng, B., Cohen, J.B., Lu, L., Hu, W., Tiwari, P., Lolli, S., Garzelli, A., Su, H., Qin, K., 2026. How can we trust TROPOMI based methane emissions estimation: calculating emissions over unidentified source regions. *Atmos. Chem. Phys.* 26, 1931–1946. <https://doi.org/10.5194/acp-26-1931-2026>
- Zhu, X., Wen, H., 2023. Numerical simulation study on the influence of air leakage on oxygen concentration in goafs of fully mechanized caving mining with shallow buried and large mining height. *Front. Earth Sci.* 11. <https://doi.org/10.3389/feart.2023.1138925>

Response to Reviewer #2:

We sincerely thank the reviewer for the careful evaluation of our manuscript and for the constructive and insightful comments. We greatly appreciate the time and effort devoted to reviewing our work. The comments have helped us improve both the scientific presentation and the clarity of the manuscript. We have carefully considered all suggestions and revised the manuscript accordingly. In this response, your original comments are completely unedited and are given in **yellow highlight**, our responses are given in **blue text** and updates to the paper are given in **purple text**.

Following these comments, we have revised the manuscript. The major changes include: (1) clarifying that the comparison with EDGAR and coal-consumption data provides contextual and physical consistency checks rather than independent validation of the MFIEF inversion; (2) expanding the discussion of the limitations of the inversion in the absence of independent ground-based CO emission measurements; (3) quantitatively characterizing the interannual decline and weakening of emission hotspots; (4) clarifying the nature of the Wugonggou underground coal fire and separating its temporal behavior from interpretations associated with anthropogenic emission-control policies; (5) clarifying the definition and purpose of the valid-observation metric in Fig. 7; (6) highlighting the relationship between power-plant coal consumption and the derived grid-scale CO emission-to-coal-consumption ratio; and (7) comprehensively revising the figures, captions, labels, terminology, and language throughout the manuscript.

Detailed responses to each comment are provided below.

Summary of “Quantifying and Attributing CO Emissions over Central Asia using TROPOMI and Explicit Observational Uncertainty” by Feng et al.

In this paper, Feng et al. apply a top-down inversion method to derive CO emissions estimates over Central Asia using TROPOMI CO observations. They incorporate a novel uncertainty analysis into their approach, which excludes emission estimates from grid cells that have high sensitivity to observational perturbations within the uncertainty range of TROPOMI. They analyse the timeseries of their top-down CO emissions in several site categories, showing considerable decreases in emissions between 2019 and later years. They then infer the emission factors of coal powerplants within the study area. Finally, they compare to bottom-up database EDGAR, which shows considerable under/over estimations depending on the location investigated.

This study builds upon previously published work, applying the inversion framework to an understudied region whilst expanding it to include a new uncertainty analysis. It advances both the methodological approach of the inversion scheme used, and the scientific knowledge of CO emissions within the study region. Emission trends within the timespan of the study are well analysed. Spatial correlations to industrial and natural sources are investigated, from which inventory and policy recommendations are justified. The figures generally support the scientific arguments presented, though the presentation of the data could be improved in most cases. Interesting and scientifically relevant conclusions are found regarding the decline in CO emissions within the region over the study period, the effect of observational uncertainty on the inversion results, and correlations between coal power station size and emission factors.

The comparison between the EDGAR anthropogenic inventory and inversion results substantially limits the conclusions that can be reached from this work because, as the authors note, “discrepancy arises because the bottom-up EDGAR inventory, by design, excludes nonindustrial sources”. Furthermore, this study would be greatly improved with a validation of the inversion scheme’s performance using independent data. In the absence of this, an expanded account of the limitations and known errors of the inversion would be sufficient, but again this limits the conclusions that can be drawn from the study. Comparison to coal consumption data shows the inversion results are physically plausible, but demonstrates nothing further than this. Overall, justifications of confirming “the consistency between satellite-inferred emissions and bottom-up activity estimates” are not valid when the inversion scheme accuracy has not been rigorously tested, the bottom-up inventories include only anthropogenic emissions, and the coal power stations have poorly-constrained emission factors.

Conclusions on the “effectiveness of MFIEF” inversion scheme should not be claimed to be shown in this study without a comparison to independent data. The decrease in annual CO emissions found (~38%) should have a significance attached to it, and efforts should be made to quantify statements such as “weakening of emission hotspots” within the abstract and conclusion.

A clearer message should be presented on the specific inventory and policy recommendations the reader should draw from the report, particularly in the conclusion section. Paragraph two of the conclusion summarises the findings well, but makes no effort to suggest where future efforts to reduce air pollution should look based on these findings. For instance, the correlation between power plant size and emission factor found from Table 2, or missing emissions from the EDGAR database could be mentioned explicitly as points of further investigation. Figure presentation is a recurring issue, with poor labelling, captions and resolution for most figures, though the data presented is mostly of a good scientific standard that is explained well in the text. There are also a considerable number of minor errors that hinder the clarity and reproducibility.

Overall, my recommendation is that this study is a good fit to the journal ACP and should be published, subject to the corrections outlined in this document.

Response: We sincerely thank the reviewer for the careful and constructive evaluation of our manuscript. We greatly appreciate the positive assessment of the scientific relevance of this work and the detailed suggestions regarding the interpretation of the inversion results, comparison with EDGAR, presentation of the figures, and clarity of the manuscript. We have carefully considered all comments and substantially revised the manuscript accordingly.

About “The comparison between the EDGAR anthropogenic inventory and inversion results” We have revised the corresponding discussion to clarify the methodological differences between the two datasets and expanded the interpretation of their discrepancies to include uncertainties associated with both bottom-up inventories and the inversion framework. Statements implying that the EDGAR comparison validates the inversion have been removed. Differences can be computed, but these differences are not validation. The statement that the coal-consumption comparison “confirms the consistency between satellite-inferred emissions and bottom-up activity estimates”

has been revised, and the statement that this comparison “supports the reliability of the inversion methodology” has been removed.

Finally, we have substantially expanded the limitations section to discuss the absence of independent CO emission measurements, propagation of satellite retrieval uncertainty through the inversion, spatial aggregation at the 0.05° grid scale, limitations in source attribution, and uncertainties associated with meteorological fields and atmospheric processes. Correspondingly, claims regarding the performance of MFIEF have been moderated throughout the manuscript. The revised manuscript therefore focuses on the capability of MFIEF to characterize regional spatial patterns, temporal variability, and differences among source environments. We have also added a numerical comparison with the CO emission-flux ranges reported by Wang et al. (2025). The regional mean fluxes obtained here, 3.06–4.93 $\mu\text{g m}^{-2} \text{ s}^{-1}$, are of a comparable order of magnitude to their reported values. This comparison provides a broad reference for the emission magnitudes rather than an independent validation of the inversion.

We have expanded the discussion of possible causes of the observed differences. These include differences in source coverage, activity data and emission factors, spatial and temporal allocation, satellite temporal sampling, retrieval uncertainty, inversion assumptions, and contributions from diffuse or non-industrial sources. We retained the underground coal-fire comparison in the current Fig. 11a (formerly Fig. 10a), but its purpose has been revised. It is now presented as a source-coverage example.

Revised manuscript, Abstract, lines 19–25: “The spatial distribution of satellite-derived grid-scale CO emission estimates reveals differences in urban emissions relative to bottom-up inventories, indicates a moderate positive association between coal consumption and inferred CO emissions at selected industrial locations, and highlights emissions associated with persistent underground coal fires, which are not explicitly represented as a separate source category in conventional anthropogenic emission inventories. This study demonstrates the potential of MFIEF as an observation-driven approach for characterizing regional CO emission patterns and temporal variability in data-sparse regions and provides complementary information for emission-inventory development and source characterization when production and location data are not available or are out of date.”

Revised manuscript, Sect. 5, lines 498–499: “The comparisons with EDGAR and reported coal consumption provide contextual assessments of spatial correspondence and physical plausibility.”

Revised manuscript, Sect. 3.2, lines 397–399: “A moderate positive association was observed between coal consumption and the inferred CO emissions ($r = 0.59$, $p = 0.058$, $N = 11$; Fig. 9a), indicating that plants with higher coal consumption generally tended to be located in grid cells with higher inferred emissions.”

Revised manuscript, Sect. 3.2, lines 362–363: “Because the inversion is performed at a spatial resolution of $0.05^\circ \times 0.05^\circ$, these estimates represent the net emissions from all sources within the corresponding grid cell rather than emissions from an individual facility.”

Revised manuscript, Sect. 4, lines 438–440: “EDGAR estimates anthropogenic emissions using sector-specific activity data and emission factors on a monthly time scale and locates them based on a priori information, whereas MFIEF derives a grid-scale signal constrained by atmospheric observations.”

Revised manuscript, Sect. 4, lines 458–459: “This difference is expected because persistent

underground coal fires are not explicitly represented as an independent source category in EDGAR and may therefore be absent or only indirectly represented within existing anthropogenic sectors.”

Revised manuscript, Sect. 5, lines 496–506: “Several limitations should be considered when interpreting the MFIEF-derived CO emissions. First, independent ground-based or facility-level CO emission measurements with spatial and temporal coverage comparable to the satellite-derived estimates are unavailable across the study region. The comparisons with EDGAR and reported coal consumption provide contextual assessments of spatial correspondence and physical plausibility. Second, the derived emissions remain sensitive to uncertainties in the satellite CO and HCHO retrievals, meteorological fields, temporal sampling, and the assumptions used to separate transport, chemical production, and surface emissions. In particular, vertical transport can alter the relationship between surface CO concentrations and satellite-observed column loadings (Wang et al., 2026). Third, the grid-scale emission flux estimated for an industrial facility includes the combined influence of sources within the corresponding grid cell. The present results are most appropriate for characterizing broad spatial patterns and relative temporal variability at the grid scale. Future integration of ground-based observations, facility-level measurements, aircraft or UAV observations, and multi-species satellite measurements would provide additional constraints on emission magnitude and source attribution.”

About “Conclusions on the “effectiveness of MFIEF” inversion scheme” We have removed statements claiming that this study demonstrates the effectiveness of MFIEF and now describe its potential for characterizing broad spatial patterns and temporal variability in data-sparse regions. Regional mean CO emissions decreased from $4.93 \mu\text{g m}^{-2} \text{s}^{-1}$ in 2019 to $3.06 \mu\text{g m}^{-2} \text{s}^{-1}$ in 2024, corresponding to a 37.9% endpoint reduction. Using 72 monthly observations, we fitted a regression with calendar-month fixed effects and heteroskedasticity- and autocorrelation-consistent standard errors. Because this analysis uses a distinct statistical model and robustness test, we retained it as a dedicated Methods subsection (Section 2.5) and expanded the subsection to state the sample size, distinguish the endpoint comparison from the fitted trend, and specify the reporting of 95% confidence intervals. Finally, we quantified the previously qualitative statement regarding hotspot weakening. Using the fixed threshold adopted in the study, 11 hotspot regions were identified. Their mean emission intensity decreased from $9.51 \mu\text{g m}^{-2} \text{s}^{-1}$ in 2019 to $5.63 \mu\text{g m}^{-2} \text{s}^{-1}$ in 2024, representing a 40.8% reduction. Because the current analysis quantifies hotspot intensity rather than a statistically tested change in hotspot area, references to hotspot contraction have been replaced by the more precise expression reduced hotspot intensity. These revisions have been made consistently in the Abstract, Methods, Results, and Conclusions.

Revised manuscript, Sect. 3.2, lines 321–336: “Regional mean CO emissions decreased from $4.93 \mu\text{g m}^{-2} \text{s}^{-1}$ in 2019 to $3.06 \mu\text{g m}^{-2} \text{s}^{-1}$ in 2024, corresponding to a 37.9% endpoint reduction. The seasonally adjusted trend was $-0.33 \mu\text{g m}^{-2} \text{s}^{-1}$ (95% confidence interval: -0.52 to $-0.15 \mu\text{g m}^{-2} \text{s}^{-1} \text{yr}^{-1}$; HAC-adjusted $p < 0.001$). A seasonal Mann–Kendall test independently supported the decreasing trend: $-0.20 \mu\text{g m}^{-2} \text{s}^{-1} \text{yr}^{-1}$ ($p < 0.001$). This overall trend toward lower emissions is further reflected in the evolution of emission hotspots, which are defined as regions where the emission flux exceeded twice the study-area mean ($6.72 \mu\text{g m}^{-2} \text{s}^{-1}$). A total of 11 such hotspots were identified, primarily situated between 39.83° – 46.58° N and 81.08° – 98.23° E. Their average

intensity decreased markedly from $9.51 \mu\text{g m}^{-2} \text{s}^{-1}$ in 2019 to $5.63 \mu\text{g m}^{-2} \text{s}^{-1}$ in 2024, representing a total reduction of 40.8% or an average annual decrease of $0.68 \mu\text{g m}^{-2} \text{s}^{-1} \text{yr}^{-1}$. The PDFs of the emissions (Fig. 6b) provide further evidence of this systematic reduction, showing a distinct leftward shift toward lower values. The distribution peak decreased from $3.34 \mu\text{g m}^{-2} \text{s}^{-1}$ in 2019 to $2.85 \mu\text{g m}^{-2} \text{s}^{-1}$ in 2024. Concurrently, the PDFs of the standard deviations (Fig. 6c) indicate a significant weakening of spatial heterogeneity, with the peak value dropping from $1.66 \mu\text{g m}^{-2} \text{s}^{-1}$ to below $1.30 \mu\text{g m}^{-2} \text{s}^{-1}$. In summary, the results depicted in Fig. 6 clearly highlight a marked downward trend in CO emissions between 2019 and 2024, characterized by reduced hotspot intensity, a systematic shift of the emission distribution to lower values, and reduced spatial variability. These observed changes may reflect the combined effects of emission-control policies, industrial restructuring, energy use, meteorology, sampling, and non-industrial sources; their individual contributions cannot be separated here. Together, these results indicate a statistically significant regional decline, reduced hotspot intensity, and weaker spatial variability over the study period.”

About “the specific inventory and policy recommendations the reader should draw from the report”

We have revised the Conclusions to state the practical implications more cautiously. The revised text emphasizes that the retained observation-driven estimates provide complementary constraints on emerging, spatially variable, and temporally dynamic sources that may be inadequately represented in conventional inventories. It also identifies the need for independent ground-based, facility-level, aircraft or UAV, and multi-species satellite observations to provide stronger constraints on emission magnitude and source attribution. We do not interpret the present analysis as quantitatively separating the effects of emission-control policies, industrial activity, energy use, meteorology, sampling, and non-industrial sources.

Revised manuscript, Sect. 5, lines 490–506: “These findings demonstrate the value of observation-driven inversion for regional emission characterization. The uncertainty analysis quantifies the observational contribution to emission uncertainty and identifies grid-cell estimates that do not meet the adopted stability criteria. The retained estimates provide complementary constraints on regional emission patterns and source environments, particularly for characterizing emerging, spatially variable, and temporally dynamic emission sources that may be inadequately represented in conventional inventories. Future applications to other satellite-retrieved species will require species-specific evaluation of retrieval sensitivity, chemistry, transport, and independent validation data. ... Future integration of ground-based observations, facility-level measurements, aircraft or UAV observations, and multi-species satellite measurements would provide additional constraints on emission magnitude and source attribution.”

High-level comments/corrections

The text includes a high volume of typos and poor phrasing, and will need to be reviewed again.

Response: Thank you for this important comment. We conducted an additional line-by-line proofreading of the manuscript and corrected the specific grammatical, terminological, punctuation, unit-formatting, reference-formatting, and figure/table cross-reference issues identified below. We also revised several sentences for clarity and consistency.

The captions provided are generally insufficient. The authors should consider these more carefully and perhaps refer to guides on good practices for caption writing. The resolution of all figures is poor and includes areas of very blurry text. The resolution must be improved, and where possible, vector files should be generated. Axis labels and titles are often inadequate and sometimes missing altogether. Rainbow colourbars are used throughout, which gives a false impression of the data.

Response: Thank you for this helpful comment. We revised the figures and captions to improve resolution and readability, added missing axis labels and units, enlarged legends and annotations, replaced rainbow colour scales with sequential or diverging colour scales where appropriate, and expanded the captions to define the variables, time periods, spatial sampling, and screening thresholds. The figure-specific revisions are described individually below.

Cities (e.g. Urumqi) referred to in the text are never labelled on a map, and it is often not clear they are cities without contextual hints from the surrounding text.

Response: Major cities referred to in the text, including Urumqi and other principal urban locations used in the analysis, have now been labelled in Fig. 1.

The revised Fig. 1 and its caption are provided in the manuscript at lines 109–112.

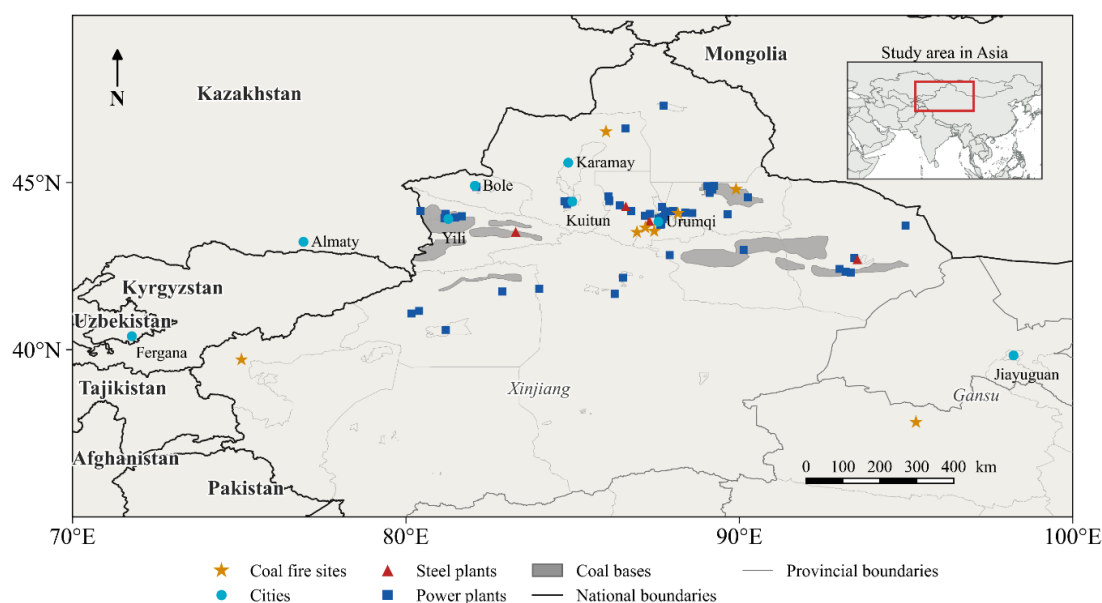


Figure 1. Study area (35°–50° N, 70°–100° E). The mapped coal bases and industrial facilities provide geographic context for the source-category analysis and are located in Xinjiang and adjacent parts of Gansu and Qinghai. For the 2019–2024 category-level statistics, only continuously operating facilities were retained, comprising 4 steel plants and 53 coal-fired power plants.

Major comments/corrections

Further discussion on the limitations of the direct comparison to EDGAR data are needed, with exploration of possible reasons that may account for differences observed beyond bottom-up inventory inaccuracies. Figure 9b shows a difference plot in emissions despite EDGAR explicitly not including non-anthropogenic emissions, making this direct comparison surprising at best and

misleading at worst. The comparison in Figure 10a between EDGAR emissions and derived inversion fluxes seems misplaced, unless the authors are arguing for the (natural?) underground coal fires to be included in anthropogenic emission inventories in future.

Thank you for your valuable suggestion. We agree that the original presentation could be interpreted as treating MFIEF and EDGAR as directly comparable estimates of the same quantity. This was not our intention, and we have revised both the text and figure interpretation accordingly.

The purpose of the comparison at the underground coal-fire locations is now explicitly framed as an illustration of differences in source coverage rather than evidence that EDGAR “underestimates” coal-fire emissions. EDGAR is an anthropogenic inventory and does not explicitly represent persistent underground coal-fire emissions as a separate source category, whereas the observation-driven inversion responds to the total atmospheric CO enhancement within a grid cell.

We have therefore revised Section 4 to emphasize that differences between MFIEF and EDGAR may result from multiple factors, including differences in source coverage, activity data and emission factors, spatial allocation, temporal representation, satellite retrieval uncertainty, and inversion uncertainty. We no longer interpret the MFIEF–EDGAR difference as evidence that one dataset is necessarily more accurate than the other. Figure 11a (formerly Fig. 10a) is retained as a source-coverage example, but its caption and discussion have been revised to make clear that it is not a validation of EDGAR or MFIEF.

These qualifications are now stated in Sect. 4, lines 438–448 and 458–468, and in the Fig. 11 caption, lines 470–475.

Natural coal fires are mentioned several times in the paper, but it is often unclear if these are natural or human-controlled/influenced events. An explanation of these events, including the context they occur within the study region, is missing from the introduction, and should be highlighted more clearly throughout the text. Does the coal fire site represent a source independent of human regulation controls in Figures 4a and 8a? If so, specify this in the text and comment on it. If not, why not? They show considerable decline, so if they are natural this weakens the proposed link between the observed declines at anthropogenic facilities and new emissions regulations.

Response: We thank the reviewer for identifying this ambiguity. We agree that describing the underground coal fires simply as “natural” was too restrictive. Underground coal fires in Xinjiang may begin through spontaneous oxidation of exposed coal and may also be initiated or intensified when mining exposes coal seams and fractures. The satellite observations used here cannot distinguish these causal mechanisms. We have therefore revised the manuscript to consistently describe these sources as “persistent underground coal fires” rather than assigning them exclusively to a natural origin. We have also added a brief explanation in the Introduction describing their occurrence and potential natural and anthropogenic influences.

Regarding the Wugonggou coal-fire site shown in Figures 4a and 8a, we interpret its temporal variation separately from that of regulated industrial stack sources. The manuscript now describes Wugonggou as an uncontrolled underground coal-fire source rather than an engineered industrial stack, and its temporal evolution is not interpreted using the same emission-control mechanisms as the urban and industrial grid cells. The observed regional changes may reflect multiple influences, including industrial activity, residential heating, meteorology, satellite sampling, and underground coal-fire activity.

Revised manuscript, Introduction, lines 69-70: “Persistent underground coal fires in Xinjiang may begin through spontaneous oxidation of exposed coal and may also be initiated or intensified when mining exposes coal seams and fractures.”

Revised manuscript, Introduction, lines 70-72: “Because these processes are partially related to human activity, we use the neutral term ‘persistent underground coal fires’ and interpret them separately from regulated industrial sources.”

Revised manuscript, Sect. 3.2, lines 376-378: “In particular, Wugonggou is an uncontrolled underground coal-fire source rather than an engineered industrial stack, and its temporal evolution should not be interpreted in terms of the same emission-control mechanisms as the urban and industrial grid cells.”

“number of valid observational days” presumably means average number of valid observational days per gridcell within one month. This needs to be explicitly mentioned, otherwise there is no information about how one defines a valid day over the entire domain. It isn’t clear why average emissions and average number of valid days are presented together in Figure 7-is there a connection between the two? There doesn’t appear to be, but “The joint evolution of these two metrics highlights both the long-term decline in regional CO emissions and the persistence of seasonal emission peaks” implies that there is some link that includes the valid day data that I have missed.

Response: We thank the reviewer for pointing out this ambiguity. The reviewer’s interpretation is correct: the metric represents the mean number of days per month for which a valid emission estimate is available in each grid cell, averaged across the study domain. The valid-day series was included as a sampling diagnostic rather than as a physical driver of CO emissions. Its purpose is to show whether the temporal variation in the regional mean emissions could be associated with changes in satellite sampling and the availability of valid inversion results.

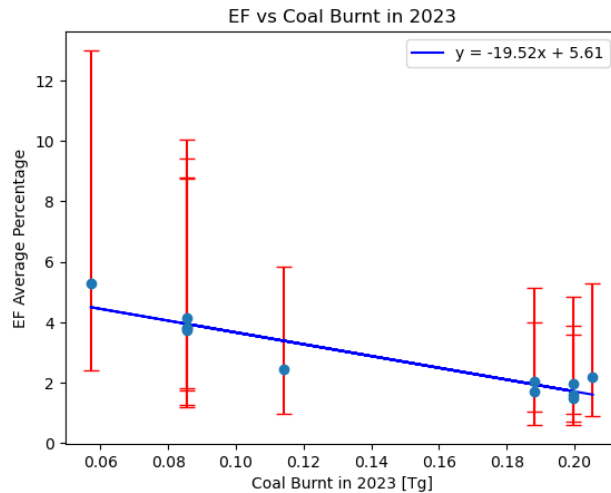
Revised manuscript, Sect. 3.2, lines 342–343: “Fig. 7 presents the monthly regional mean CO emissions and the mean number of valid emission-estimate days per grid cell from January 2019 to December 2024.”

Revised manuscript, Sect. 3.2, lines 343–345: “A valid day is defined as a day for which a grid cell retains an emission estimate after the quality and perturbation-response screening described above; the plotted valid-day metric is the spatial mean of this monthly count across the study domain.”

Revised manuscript, Sect. 3.2, lines 351–352: “The valid-day series is presented only as a sampling diagnostic and is not interpreted as a physical driver of emission variability.”

Revised manuscript, Fig. 7 caption, lines 354–358: “Figure 7. Monthly mean CO emission flux and temporal sampling coverage over the study region from January 2019 to December 2024. (a) Spatially averaged monthly CO emission flux derived from grid cells satisfying the observational-quality criteria. (b) Mean number of valid observation days per grid cell per month, averaged over all grid cells in the study region, including grid cells with no valid observations in a given month. The valid-day series is shown as a diagnostic of temporal sampling coverage rather than as an explanatory driver of emission variability.”

The correlation between power plant size and emission factor presented in Table 2 is an interesting result which isn’t highlighted. The table structure buries this because it is ordered by latitude rather than Tons of CO (2023). This could be more clearly visualised in a graph (see mock graph I quickly generated from the data in Table 2 below).



Response: Thank you for this suggestion. We replaced the former Table 2 with a two-panel Figure 9. After re-examining the operating status of the 12 originally listed plants, we excluded one plant whose operation in 2023 could not be confirmed. For the remaining 11 plants, coal consumption showed a moderate positive association with grid-scale CO emissions ($r = 0.59$, $p = 0.058$; Fig. 9a), although this association did not reach the conventional significance level of 0.05. The grid-scale CO emission-to-coal-consumption ratio ranged from approximately 14 to 53 g CO kg⁻¹ coal and showed a strong inverse association with coal consumption ($r = -0.93$, $p < 0.001$; Fig. 9b). We interpret this comparative ratio cautiously because coal consumption appears in the denominator and the grid-cell estimates may include contributions from nearby sources. The revised Figure 9 is reproduced below.

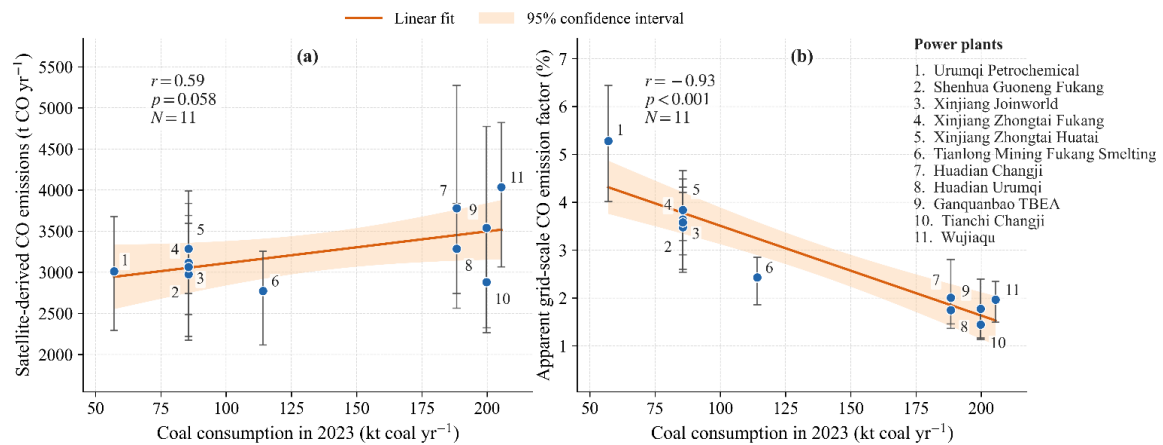


Figure 9. Relationship between satellite-derived CO emissions and coal consumption at selected coal-fired power plants. (a) Comparison between annual coal consumption and MFIEF-derived CO emission for individual power-plant grid cells. (b) Relationship between coal consumption and the grid-scale CO emission-to-coal-consumption ratio. Circles and vertical error bars represent the means and 25th–75th percentile ranges, respectively.

Revised manuscript, Sect. 3.2, lines 397–399: “A moderate positive association was observed between coal consumption and the inferred CO emissions ($r = 0.59$, $p = 0.058$, $N = 11$; Fig. 9a), indicating that plants with higher coal consumption generally tended to be located in

grid cells with higher inferred emissions.”

Revised manuscript, Sect. 3.2, lines 402–409: “The grid-scale CO emission-to-coal-consumption ratio, calculated as the satellite-derived CO emission from the $0.05^{\circ} \times 0.05^{\circ}$ grid cell containing each power plant divided by the reported coal consumption of that plant, ranged from approximately 14 to 53 g CO kg⁻¹ coal. This ratio is used only as a comparative metric rather than a facility-specific emission factor, because the satellite-derived emissions represent the combined emissions from all sources within the corresponding grid cell rather than actual stack emissions. The ratio showed a strong inverse relationship with coal consumption ($r = -0.93$, $p < 0.001$, $N = 11$; Fig. 9b). However, because coal consumption is included in the denominator, the relationship is partly affected by mathematical coupling.”

Figures

Figure 1 requires an inset to place the region geographically in the world. Major cities referred to in the text should be named on the map. Borders should be clarified as international or internal, with internal border linewidths thickened if they aren't removed entirely, and countries could be named to make the figure more interpretable. It is unclear if the power and steel plants labelled are only for those within China/Xinjiang or if none exist outside this region. Legend text should be larger and less blurry.

Response: We thank the reviewer for the detailed comments on the figures. We have comprehensively revised all figures and captions. Specifically, we increased figure and label resolution, generated vector-format figures where possible, added missing axis labels and units, enlarged legends and annotations, replaced rainbow color scales with perceptually more appropriate sequential or diverging color scales, and expanded the captions so that each figure can be interpreted more independently from the main text. Geographic labels have also been added where appropriate.

We have revised Fig. 1. Country and city labels, the geographic scope, and the facility-selection scope are clarified in the figure and caption. The mapped industrial facilities are those used for the source-category analysis, not a complete inventory for all of Central Asia. See Sect. 2.1 and the Figure 1 caption. The revised Figure 1 is reproduced below.

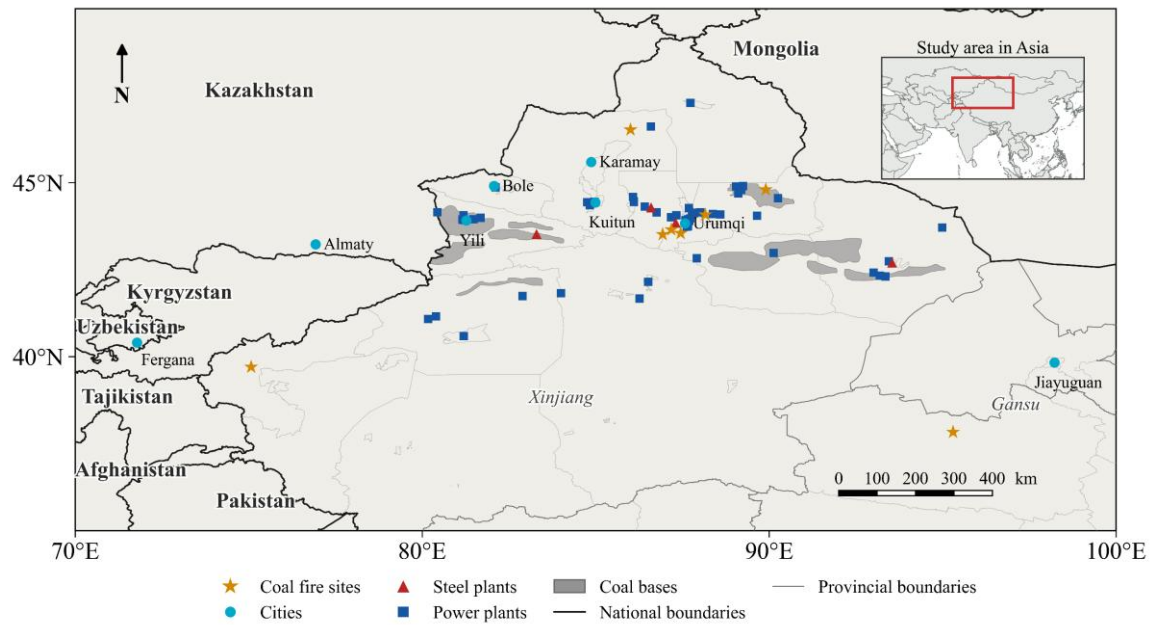


Figure 1. Study area (35°–50° N, 70°–100° E). The mapped coal bases and industrial facilities provide geographic context for the source-category analysis and are located in Xinjiang and adjacent parts of Gansu and Qinghai. For the 2019–2024 category-level statistics, only continuously operating facilities were retained, comprising 4 steel plants and 53 coal-fired power plants.

Figure 2 should be restructured to show the flow of the methodology from top to bottom. The box labelled “the same method” should be removed and a parallel line of processes should be added to show the same steps. It isn’t clear if prior CO emissions and wind data also provide input to the random perturbation path (as I would guess), or if this is independent of these inputs (as the current figure suggests).

Response: Figure 2 presents the workflow and the original-observation and perturbed-observation paths, together with the common inversion inputs. The revised Figure 2 is reproduced below.

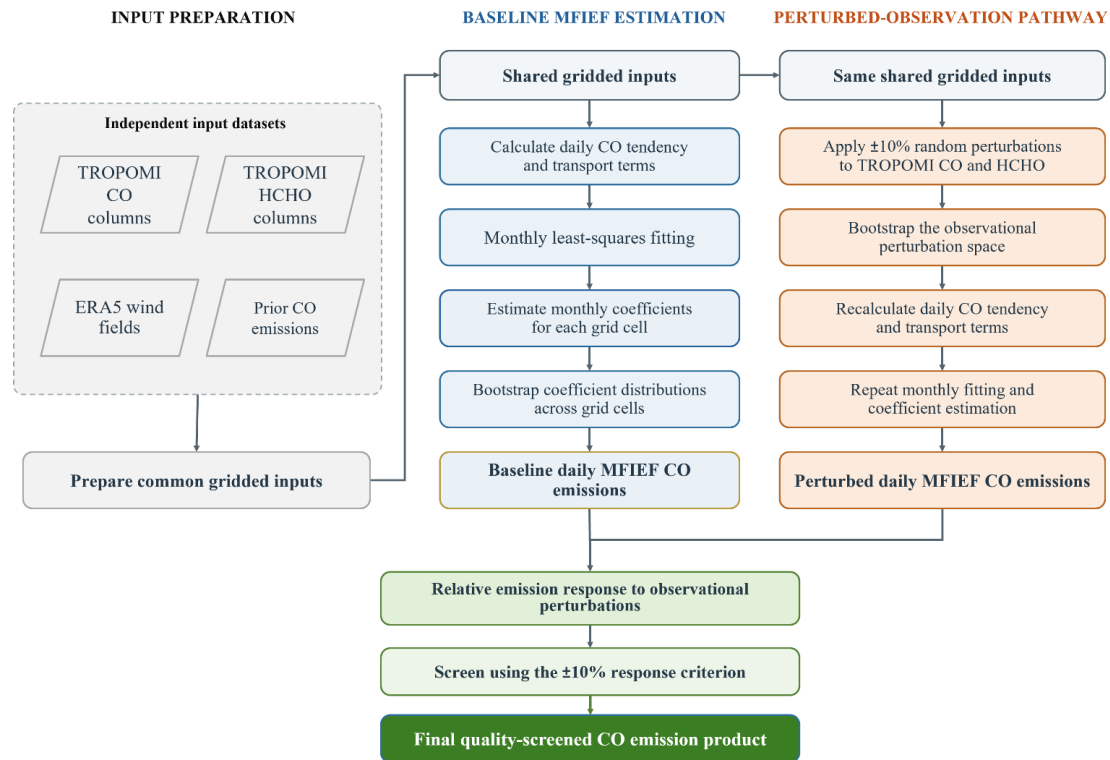


Figure 2. Workflow of MFIEF-based CO emission estimation and uncertainty-propagation analysis based on $\pm 10\%$ perturbations of TROPOMI CO and HCHO observations.

Figure 3a uses a rainbow colourbar, which gives a false visual impression of the CO concentrations and should be changed. The colourbar label should have a title as well as units. The 19-24 text should be modified to something clearer or removed. Figure 3b requires axis labels. The legend and labels that are currently present should be larger. The figure should be created as a vector to improve its resolution. The captions for Figure 3 are both insufficient.

Response: The color scale, colorbar title and units, year label, axis labels, legend size, and caption have been revised. The revised Figure 3 is reproduced below.

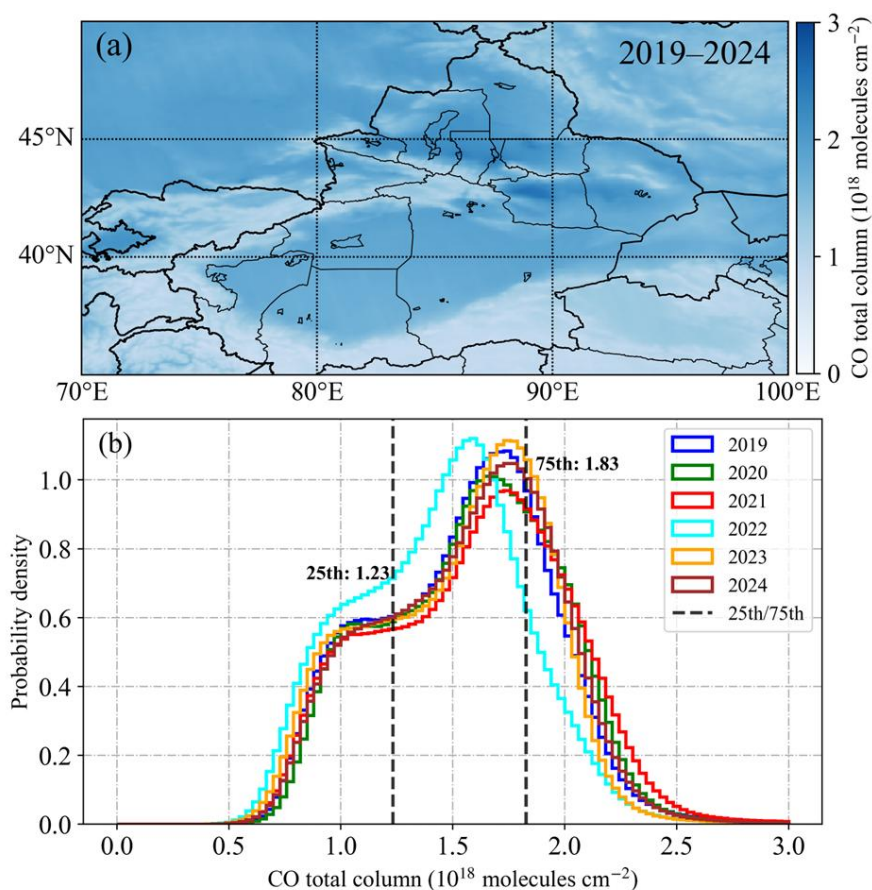


Figure 3. Spatial and statistical distributions of TROPOMI CO total columns over the study area during 2019–2024. (a) Mean spatial distribution over the six-year period. (b) Annual probability density distributions of CO total-column values over the study area.

Figure 4 requires axis labels. The date tick marks could be simplified to one per year at the start of the year. It is not clear over what area the TROPOMI CO total column data was averaged to generate the timeseries (average of nearest gridcell, closest point of each overpass, etc.). Legends should be improved, it is not clear Urumqi is a city or that “coal fire” represents (natural?) underground coal fires, or that “power plant” refers to a coal power plant. It isn’t clear if steel and coal plants represent a mean of all plants in the region, or if not which one they refer to. Resolution needs to be improved substantially, and the figure should be generated as a vector.

Response: Axis labels and units have been added and the sampling is clarified as the $0.05^\circ \times 0.05^\circ$ grid cell containing each selected coordinate. The labels identify the selected underground coal-fire site, urban center, steel plant, and coal-fired power plant. See Sect. 3.1 and the Figure 4 caption. The revised Figure 4 is reproduced below.

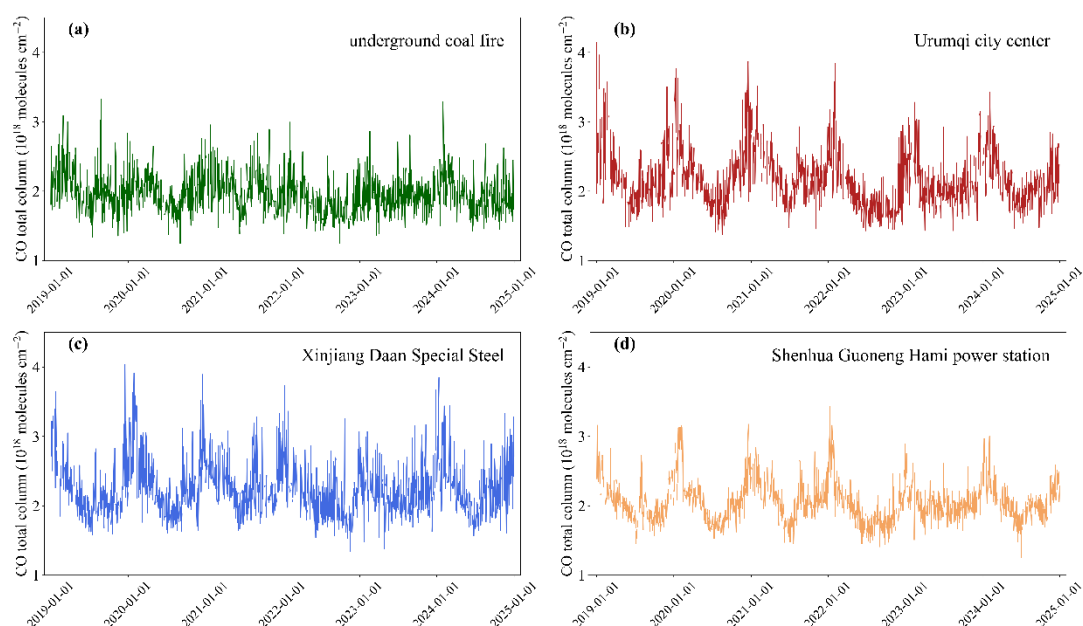


Figure 4. Time series of TROPOMI CO total columns in the $0.05^{\circ} \times 0.05^{\circ}$ grid cells containing (a) the Wugonggou underground coal-fire site, (b) the center of Urumqi, (c) Xinjiang Daan Special Steel, and (d) Shenhua Guoneng Hama Power Station during 2019–2024.

Figure 5a should change colourbar and change label to AAOD443/AAOD880 Ratio. Remove (a) in bottom left corner of figure. Figure 5b is very hard to read, I can't easily interpret the meaning or point of this figure. It requires axis labels with units. The legend should make clear that the $\pm 10\%$ and $\pm 20\%$ lines refer to uncertainty thresholds of TROPOMI data. Resolution needs to be improved. Figure 5c should change the colourbar, which should be different to figure 5a to show that it is different data plotted. Colourbar label should be added. Captions should be improved.

Response: The Figure 5 panels, labels, color scales, uncertainty-threshold annotation, and caption have been revised to distinguish the AAOD ratio, perturbation-response analysis, and retained-grid results. The revised Figure 5 is reproduced below.

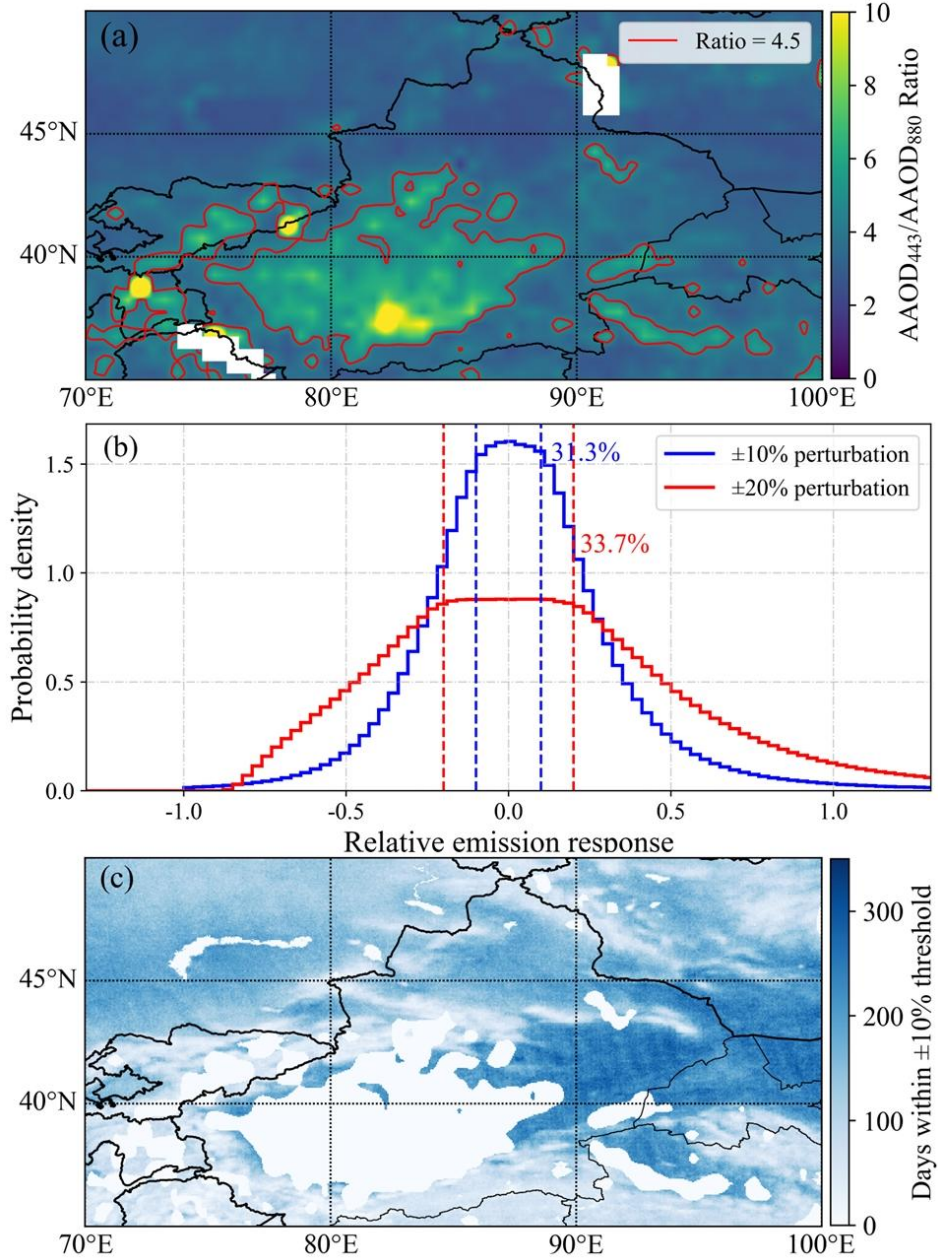


Figure 5: (a) Spatial distribution of the mean AAOD₄₄₃/AAOD₈₈₀ ratio during 2019–2024. Grid cells with ratios greater than 4.5 were identified as dust-dominated areas and excluded from the subsequent CO emission inversion. (b) Probability density distributions of the relative emission changes resulting from ±10% and ±20% perturbations in the TROPOMI CO and HCHO total-column inputs. The blue and red dashed lines indicate corresponding uncertainty bounds of ±0.1 and ±0.2, respectively. The annotated percentages represent the proportions of estimates remaining within the corresponding uncertainty bounds. (c) number of valid retrievals after ±10% observational-uncertainty filtering.

All panels in Figure 6 should have axis labels added. Rainbow colourbars should be changed. “emission” and “std” titles should be improved.

Response: Axis labels and units have been added, the color scales and panel titles have been revised, and the caption defines the annual screening criterion and variables. The revised Figure 6 is reproduced below.

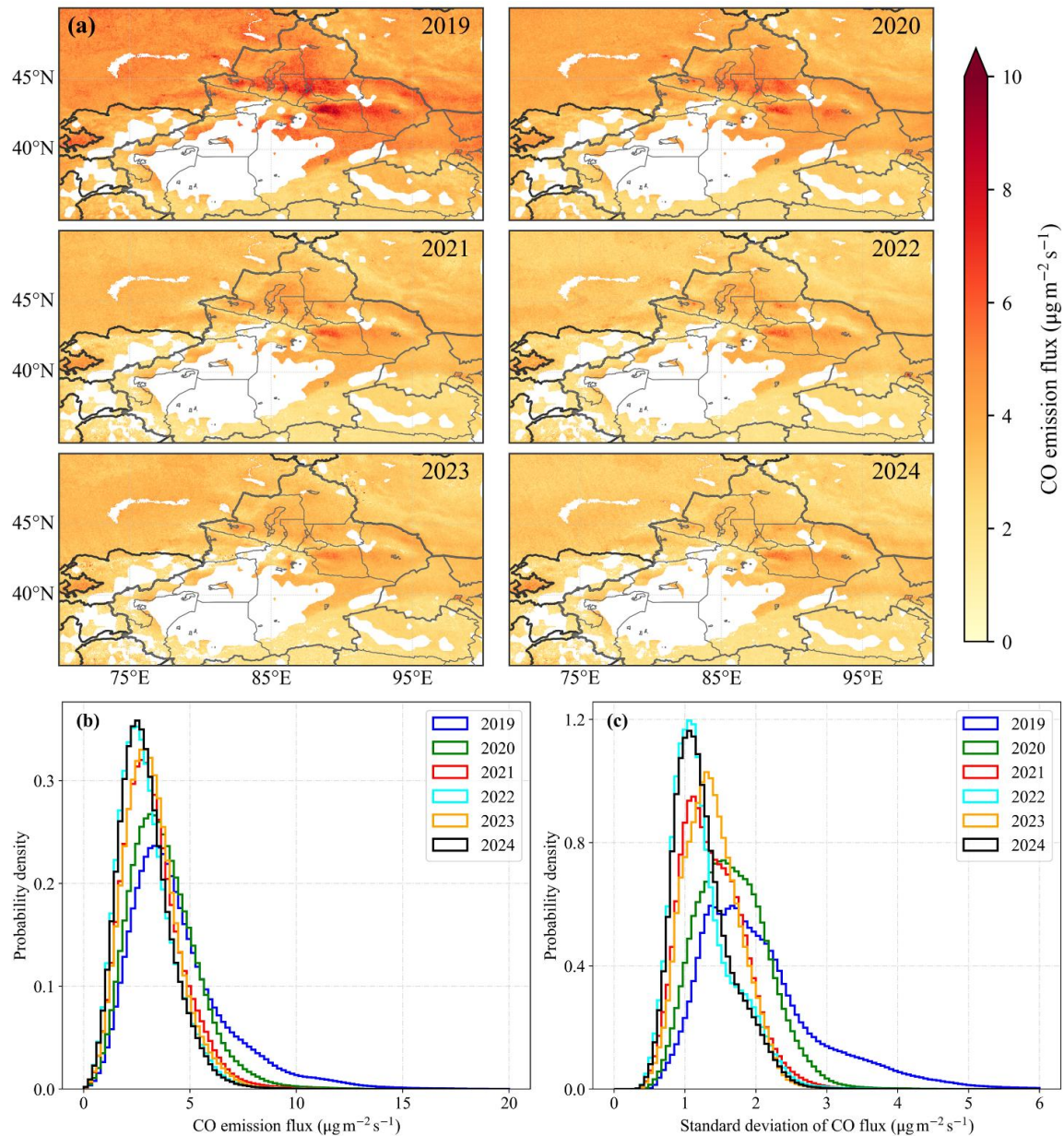


Figure 6. Annual characteristics of MFIEF-derived CO emissions from 2019 to 2024. (a) Annual mean CO emission flux; grid cells with fewer than 10 valid emission-estimate days in a year are shown as missing values. (b) Probability density distributions of annual mean CO emission fluxes. (c) Probability density distributions of the interannual standard deviation of CO emission fluxes. Emission-flux units are $\mu\text{g m}^{-2} \text{s}^{-1}$.

Figure 7 left axis should have units added. “per grid” presumably means “per gridcell”.

Response: Correct. The left-axis units have been added and the valid-day metric is defined as the domain average of the number of days per month with a valid grid-cell emission estimate. It is shown only as a sampling diagnostic. The revised Figure 7 is reproduced below.

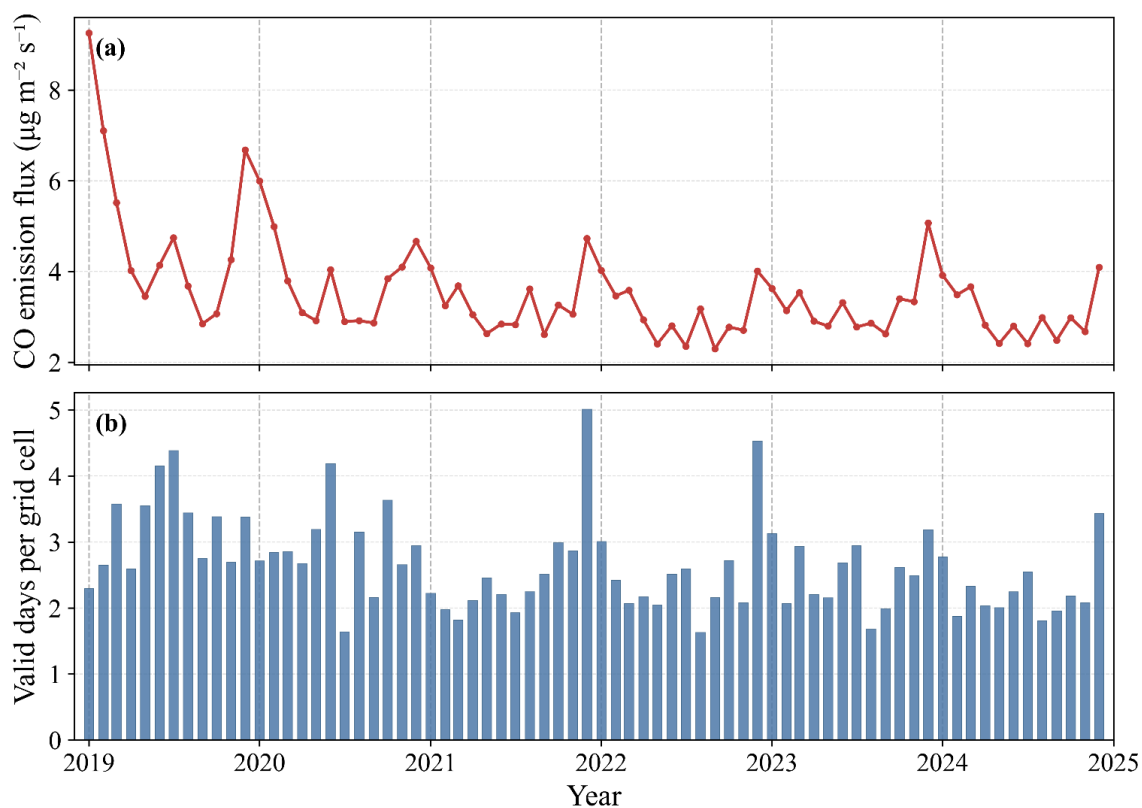


Figure 7. Monthly mean CO emission flux and temporal sampling coverage over the study region from January 2019 to December 2024. (a) Spatially averaged monthly CO emission flux derived from grid cells satisfying the observational-quality criteria. (b) Mean number of valid observation days per grid cell within each month. The valid-day series is shown as a diagnostic of temporal sampling coverage rather than as an explanatory driver of emission variability.

Figure 8 legends need to be improved, see Figure 4 comments. Axis labels are needed.

Response: Axis labels and units have been added, and the panels identify the four selected locations explicitly. The text now uses “locations” and “grid-scale emission flux” rather than implying direct point-source stack measurements. The revised Figure 8 is reproduced below.

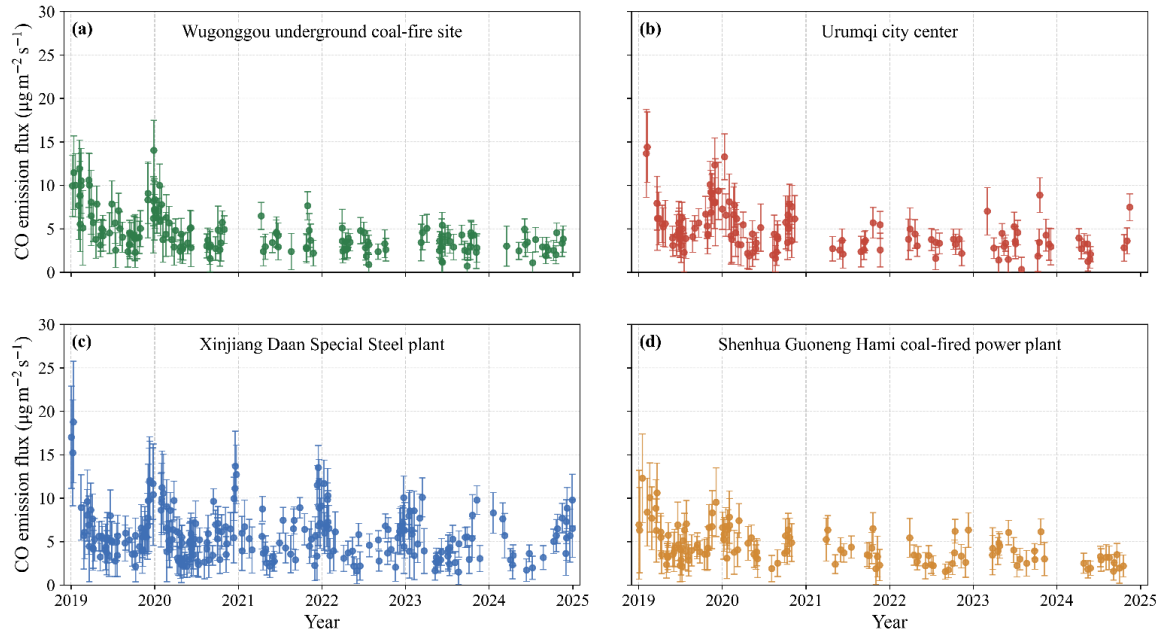


Figure 8. Temporal variations in MFIEF-derived CO emission fluxes at four representative locations from January 2019 to December 2024: (a) the Wugonggou underground coal-fire site, (b) Urumqi city center, (c) the Xinjiang Daan Special Steel plant, and (d) the Shenhua Guoneng Hami coal-fired power plant. Markers represent the mean emission fluxes, and error bars indicate ± 1 standard deviation. The values were extracted from the 0.05o grid cell containing each location.

Figure 9 has the same problems as Figure 6.

Response: Because the former Table 2 is now presented as Figure 9, the former Figure 9 has been renumbered Figure 10. Axis and colorbar labels and the diverging difference scale have been revised, and the caption defines the variables, period, and sign of the MFIEF–EDGAR difference. The revised Figure 10 is reproduced below.

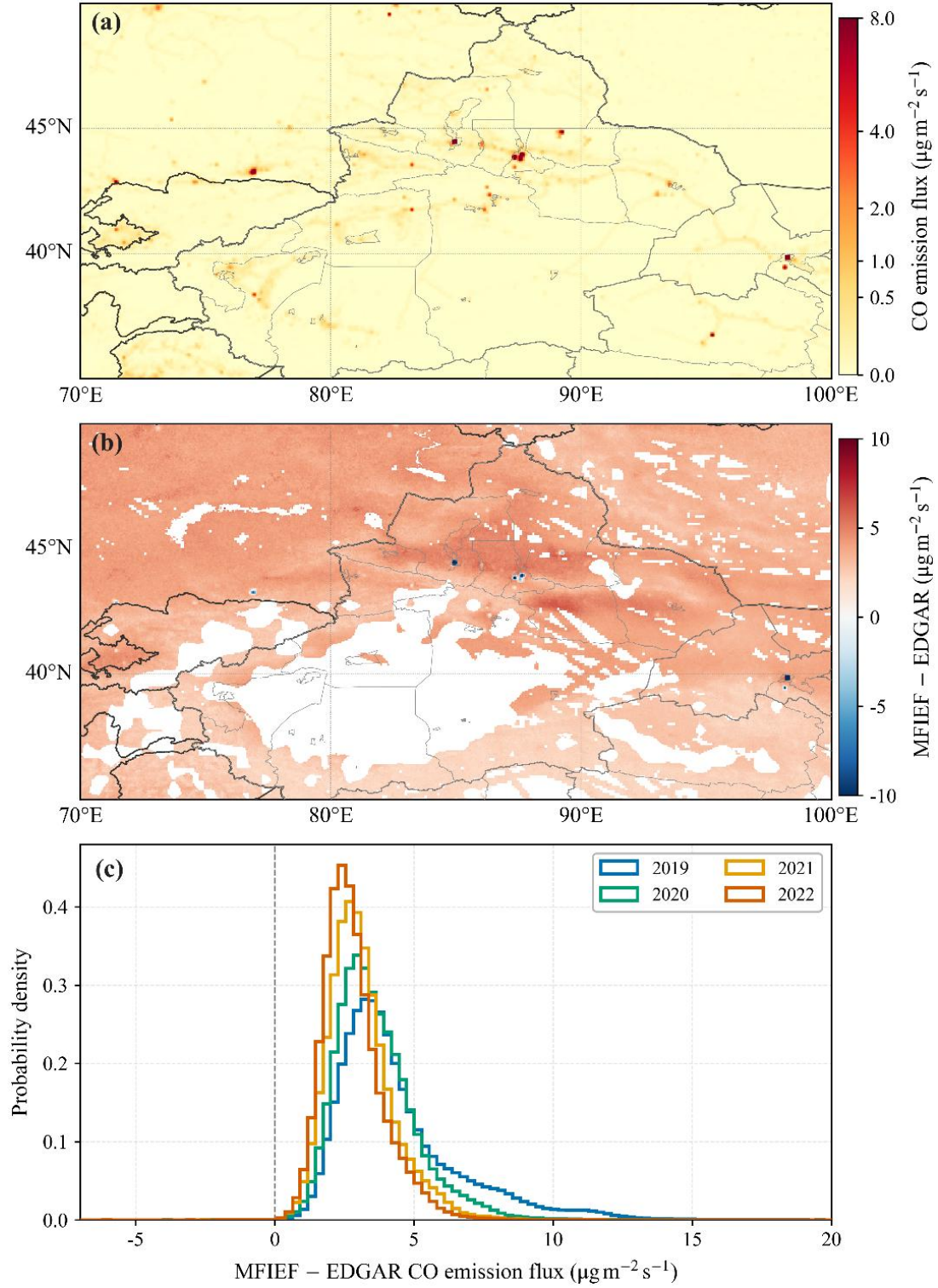


Figure 10. Comparison of CO emission estimates from MFIEF and EDGAR during their overlapping period from 2019 to 2022. (a) Mean EDGAR anthropogenic CO emission flux. (b) Mean grid-cell difference calculated as MFIEF minus EDGAR. Positive values indicate larger MFIEF-derived fluxes, whereas negative values indicate larger EDGAR emissions. (c) Probability density distributions of the monthly grid-cell differences. The differences reflect the distinct source coverage, spatial allocation, temporal representation, and estimation methods of MFIEF and EDGAR.

Figure 10 needs axis labels. 10a should have y axis limits standardised across the 4 panels for ease of comparison. The drastically different y axis scale in the bottom left panel of figure 10b should be mentioned in the caption.

Response: Because the new coal-consumption analysis is Figure 9, the former Figure 10 has been renumbered Figure 11. Axis labels and units have been added; the coal-fire panels use a common y-axis scale, and the distinct Jiayuguan scale is identified in the text and caption. The revised Figure 11 is reproduced below.

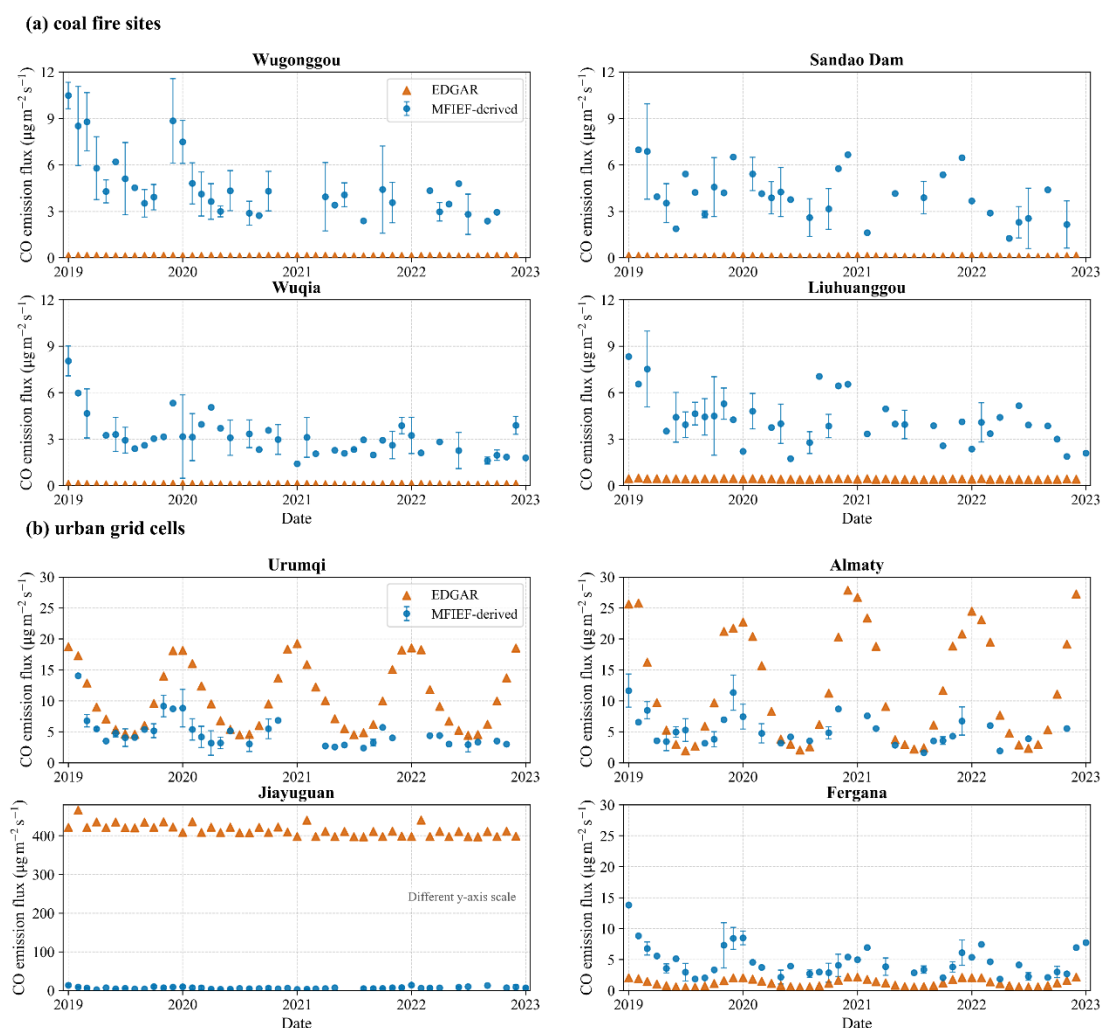


Figure 11. Comparison of MFIEF-derived grid-scale CO emission fluxes and EDGAR anthropogenic CO emissions at selected locations. (a) Persistent underground coal-fire sites, which are not explicitly represented as a separate source category in EDGAR. (b) Selected urban centers. MFIEF-derived emissions are presented as monthly means, with vertical bars indicating the within-month variability (standard deviation) of daily estimates, whereas EDGAR values represent monthly inventory estimates. The comparison illustrates differences in source coverage and temporal representation between the two datasets.

Minor comments/corrections

Line	Type	Original	Correction	Comment	Response
12	Addition	marked interannual decline	—	Across the study period	Revised in the Abstract, line 13. “Regional mean CO emissions decreased by 37.9% between 2019 and 2024, with a significant seasonally adjusted trend of $-0.33 \mu\text{g m}^{-2} \text{s}^{-1} \text{yr}^{-1}$ ”, and quantified in Sect. 3.2, lines 321–336.
22	Addition	significant environmental problems	—	Expand on this	Revised manuscript, Introduction, lines 28–31: “Air pollution from combustion-related emissions can adversely affect human health, atmospheric chemistry, and ecosystem stability, whereas greenhouse gas accumulation contributes to long-term climate change and associated environmental risks.”
25	Modify	... major sink of the hydroxyl radical (OH) affecting the concentrations of...	... major sink of the hydroxyl radical (OH), which in turn affects the concentrations of...	—	Revised at lines 35–36.
29	Restructure	Once emitted into the atmosphere, CO undergoes chemistry in-situ including oxidation into	—	CO can't be produced if it's already been emitted, needs rephrasing.	The atmospheric-chemistry description was rewritten at lines 37–42.

		carbon dioxide (CO ₂) and production from the oxidation of volatile organic compounds			
31	Modify	In unpolluted atmosphere...	Under unpolluted atmospheric conditions...	—	Revised at lines 41–42.
32	Remove	...in winter...	—	—	The sentence was rewritten and qualified at lines 43–47.
33	Modify	...higher values during times controlled mostly by...	...elevated concentrations...	—	Revised at lines 43–46.
37	Modify	...due to when and where...	...when...	—	Revised at lines 44–46.
41	Add	China's MEIC system	—	Define acronym	MEIC is defined at lines 50–52.
44	Add	WMO GAW	—	Define acronym	WMO GAW is defined at lines 56–57.
48	Add	(Jiang et al., 2015; Kopacz et al., 2010)	—	Add more, recent citations	Added Tang et al. (2022) as an additional recent reference for CTM inversions. Revised manuscript, Introduction, lines 60–62.
49	Remove	integrate atmospheric transport and chemistry to infer emissions from concentration fields	—	duplicated	The duplicated wording was removed at lines 60–63.
60	Modify	...the emissions of CO is both...	...the emission of CO is...	—	The sentence was rewritten in the Introduction.

61	Modify	Though	thought	—	The typographical error was removed when the sentence was rewritten.
65	Modify	As related...	Those related...	—	The paragraph was restructured at lines 73–82.
65	Restructure	non-power generating industries, as well as for industry and power production	—	So... all forms of industry then?	The paragraph was restructured at lines 73–82.
84	Restructure	extensive coal mining as well as many industries which use the coal and are powered by the coal.	—	Repeat coal too much	The study-area and source-category description was revised at lines 95–107.
86	Remove	Also	—	—	The redundant transition was removed.
94	Add	TROPOMI	TROPOspheric Monitoring Instrument (TROPOMI)	—	TROPOMI is defined at lines 114–115.
95	Modify	Column loading	Total column	—	The TROPOMI total-column product and daily processing are clarified in Sect. 2.2, lines 114–121.
95	Modify	Retrieved daily	retrieved	—	The TROPOMI total-column product and daily processing are clarified in Sect. 2.2, lines 114–121.
96	Restructure	...2.3 μm spectral range of the shortwave infrared (SWIR)...	—	This is the absorption feature of CO (which is in the SWIR), not the spectral range	The instrument description was rewritten in Sect. 2.2, lines 114–119.

97	Remove	CH ₄ , and aerosols	—	Not relevant to this study	The irrelevant species reference was removed.
100	Remove	And then resampled	—	repeated	The duplicated resampling phrase was removed at lines 119–121.
101	Restructure	The data were then resampled with a resolution of 0.05°×0.05°.	—	Is the resampling repeated? Confusingly worded here.	The resampling procedure is clarified at lines 119–121.
103	Modify	HCHO is one of VOCs...	—	—	The HCHO description was revised in Sect. 2.2.
105	Restructure	lifetime of HCHO columns is far shorter than those of CO columns,	—	Lifetime is for chemicals, not concentrations	The HCHO lifetime statement was revised in Sect. 2.2.
108	Modify	its uncertainty is larger than for CO	It is less sensitive to HCHO than CO	—	The uncertainty statement was revised in Sect. 2.2.
109	Modify	In this work, concurrent data of	In this work, concurrent total column observations	—	The repeated phrase was removed.
110	Modify	GEE is utilized, resampled by HARP	GEE is utilized and resampled by HARP	—	The GEE wording was revised at lines 119–121.
112	Restructure	wind:	—	I'm not a fan of these headers to the paragraphs. Wind should at least be capitalised.	The meteorological-variable description was reformatted and clarified in Sect. 2.2.
115	Modify	The 05:00 and 06:00 UTC times correspond to the observation periods of the	—	Presumably this only applies over the region of interest in this study?	The temporal sampling of the meteorological fields was clarified in Sect. 2.2.

		TROPOMI satellite			
119	Modify	with a resolution of $0.1^{\circ} \times 0.1^{\circ}$, with subsequent resampling at $0.05^{\circ} \times 0.05^{\circ}$.	with a resolution of $0.1^{\circ} \times 0.1^{\circ}$ subsequently resampled to $0.05^{\circ} \times 0.05^{\circ}$.	—	The sentence was revised and stated consistently as $0.05^{\circ} \times 0.05^{\circ}$.
123	Add	nine distinct angles	Nine distinct viewing angles	—	The sentence was revised to specify the nine variables
126	Restructure	The L3 AAOD data product used in this paper has a resolution of 0.25° and is resampled to 0.05° to distinguish dust-affected regions.	—	How will this resampling help to distinguish dust-affected regions? No more information is added by resampling like this.	The resampling was performed only to match the spatial resolution of the emission estimates and other datasets, rather than to provide additional spatial information. The AAOD screening and its interpretation are clarified at lines 285–295.
130	Modify	Equation (2)	—	Do you mean equation 1? If not, why reference eq 2 first?	Equation (2) is now defined as the observational contribution to emission uncertainty at lines 213–229.
132	Modify	approximation of	approximate	—	The term was revised to “approximate” in the Methods.
132	Restructure	weighting terms of the underlying partial differential equations that govern the transformation between CO and its driving factors,	—	Unclear what this means. Partial differential equations aren't mentioned before, and aren't explicitly represented in eq 1. By weighting	The weighting terms and their roles were clarified in the equation description.

				terms do you mean the alpha terms?	
142	Modify	these terms	Emission, production, and transport terms	—	Emission, transport, and chemical terms were explicitly defined in Sect. 2.3.
142	Modify	Eq. (1)	Equation (1)	—	Eq. (1)” was changed to “Equation (1)” where required by the sentence structure
143	Modify	after converted the unit into	in units of	—	The units and conversion were clarified in Sect. 2.3.
149	Restructure	The terms of this equation are initialized using this equation,	—	Repetition of this equation, reference the specific equation (eq 1 I think)	The repetitive sentence was removed/restructured.
151	Modify	grid-by-grid	—	Unclear what this means. Gridcell-by-gridcell?	“grid-by-grid” was replaced with “grid-cell” terminology throughout the manuscript.
151	Restructure	The fitting is based on multiple least squares regression grid-by-grid on a month-by-month basis, reflecting the differences in albedo, elevation, land-surface use, emissions type, and differences in confounding variables on each pixel specifically	—	Too long sentence	The sentence was divided and rewritten in Sect. 2.4, lines 213–229.

		grouping all daily data from TROPOMI and the ECMWF reanalysis where they all occur.			
154	Add	weighting terms	weighting terms α_i	—	The weighting terms were clarified in Sect. 2.3.
156	Restructure	To ensure physically realistic and mathematically consistent values for the coefficients α_2 , α_3 , and α_4 , physical cutoffs are applied. Any value that is not positive ($\alpha > 0$) is set to NaN.	To ensure physically realistic and mathematically consistent values for the coefficients α_2 , α_3 , and α_4 , any value that is not positive ($\alpha > 0$) is set to NaN.	If using more cutoffs than this, list them.	The threshold-selection explanation was revised; we do not claim an unperformed multi-threshold sensitivity analysis.
159	Modify	The resulting equation	Equation (1)	—	The equation reference was corrected.
159	Modify	Grids	gridcells	Across the paper too	“Grids” was replaced by “grid cells” or “grid-cell” as grammatically appropriate.
213	Modify	biomass incomplete combustion	—	Specify wildfires, agricultural, both?	Biomass-burning sources are specified at lines 46–47, 99–100, and 256–258.
214	Add	The statistical distribution of the CO total column	—	Averaged over what time period?	The spatial and temporal basis of the statistical distribution is clarified in Sect. 3.1, lines 251–263.
224	Add	The CO total column on a daily basis is	—	Over what spatial location?	The spatial domain and averaging are clarified at lines 251–

		provided in Fig. 4 for the following sites...		Closest TROPOMI observation? Averaged over gridbox?	263.
224	Add	Wugonggou Coal Fire in Fukang	—	Brief background on what this is. Natural? Long-term? Remote? Underground?	Background and interpretation were added at lines 69–72, 271–279, and 376–378.
225	Add	representative power plant, a representative steel plant	—	How are they determined to be representative? Which ones were chosen?	The four selected grid cells and their coordinates are specified at lines 271–273.
225	Add	center of Urumqi	—	Presumably a city? Where is it located?	The location is specified as the center of Urumqi at lines 271–273.
248	Add	thereby minimizing potential biases in CO emission estimates associated with dust contamination	—	Provide a source and/or explain further	The unsupported placeholder citations were removed; the dust-screening interpretation is now stated at lines 285–295.
260	Add	135 days	135 valid days	—	The valid-day metric is defined at lines 342–345 and in the Fig. 7 caption, lines 354–358.
271	Remove	Been	—	—	The grammatical error was removed when the sentence was rewritten.
272	Remove	with a reduction of approximately 38%	a reduction of 38%	—	The reduction is quantified at lines 321–336.
273	Modify	which are defined as	which we define as regions where	—	The hotspot threshold is explicitly defined

		regions where the concentration is exceeded twice the study area's mean	the concentration exceeded twice the study area's mean		at lines 324–326.
276	Modify	average annual decrease of 0.68 $\mu\text{g m}^{-2} \text{s}^{-1}$	average annual decrease of 0.68 $\mu\text{g m}^{-2} \text{s}^{-1}$ per year	—	The annual and endpoint changes are quantified at lines 326–328.
348	Add	Overall, these emission factors align well with typical ranges for CO from coal combustion reported in established bottom-up inventories such as EDGAR	—	What are these emission factors from EDGAR and MEIC? Include numbers and errors.	The unsupported statement was removed and replaced by the qualified summary at lines 332–336.
354	Modify	A comparison between this emission and EDGAR	—	What emission?	The text now specifies grid-scale CO emission flux at lines 360–365.
366	Restructure	In the figure, the blue areas of the figure are mainly distributed.	—	rephrase	The sentence was rephrased at lines 367–375.
375	Add	These statistics suggest that our inversion generally yields higher CO emissions than EDGAR	—	Why are emissions higher, why so distributed?	The possible causes and the limits on attribution are stated at lines 373–378.
395	Modify	we applied the MFIEF	—	Define MFIEF or at least mention it is an inversion algorithm	MFIEF is defined at lines 83–86 and written in full at the start of the Conclusions, lines 477–478.
399	Remove	Both	—	—	The redundant word

					was removed when the conclusion was rewritten.
401	Modify	all of which leads to	Both of which contribute to	—	The sentence was rewritten in the Conclusions.
402	Modify	The results reveal	We find	—	The sentence was rewritten in the Conclusions.
403	Modify	that it is never the less	That CO is nevertheless	—	The unsupported absolute claim was removed; the revised statement appears at lines 490–495.
404	Add	with stronger emissions during winter and early spring	—	Add some statistics	The supporting statistics are reported at lines 321–336 and 395–409.
408	Modify	urban agglomerations	Urban centers	—	The terminology was revised to “urban centers” or “city-center grid cells,” as appropriate.
421	Modify	different directions	sources	—	The source terminology was revised in the Discussion and Conclusions.
421	Modify	on-the-ground	Ground-based	—	“On-the-ground” was replaced with “ground-based” at lines 496–506.
424	Remove	the issue of	—	—	The redundant phrase was removed.