

Reply on RC2

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 emission factors; 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, $3.06\text{--}4.93\ \mu\text{g m}^{-2}\text{ s}^{-1}$, 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 17–19: “Reported coal-consumption data and EDGAR were used to examine the physical consistency and spatial representation of the satellite-derived estimates.”

Revised manuscript, Sect. 5, lines 500–501: “The comparisons with EDGAR and reported coal consumption provide contextual assessments of spatial correspondence and physical plausibility.”

Revised manuscript, Sect. 3.2, lines 396–397: “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 360–361: “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 437–438: “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, line 497: “Several limitations should be considered when interpreting the MFIEF-derived CO emissions.”

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 319-331: “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 $0.68 \mu\text{g m}^{-2} \text{s}^{-1} \text{yr}^{-1}$.”

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

The revised Conclusions now distinguish between observation-driven source signals and conventional anthropogenic inventory categories, identify improved representation of persistent non-industrial and diffuse sources as an inventory priority, and recommend facility-level measurements to investigate the apparent relationship between power-plant activity and grid-scale apparent emission factors.

The revised discussion now highlights three priorities: (1) improved representation of persistent non-industrial and diffuse CO sources that are not explicitly represented in conventional anthropogenic inventories; (2) improved spatial allocation and updating of rapidly changing industrial sources; and (3) further investigation of the apparent relationship between power-plant activity level and derived CO emission factors using facility-level observations.

We have also moderated the policy interpretation of the regional declining trend because the present observational analysis cannot quantitatively separate the effects of emission-control policies, changes in industrial activity, energy-use changes, meteorological variability, and changes in non-industrial sources.

Revised manuscript, Sect. 5, lines 483-489: “Together, 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 unstable grid-cell estimates. The retained estimates provide complementary constraints on regional emission patterns and source environments, with a particular added value in quantifying emissions from sources which are changing in their location, are being actively controlled, sources which are new and not yet identified, sources which may have important day-to-day variability, and how reliable current

best technologies can be for such applications. Future applications to other satellite-retrieved species will require species-specific evaluation of retrieval sensitivity, chemistry, transport, and independent validation data.”

Revised manuscript, Sect. 5, lines 495–497: “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 your valuable suggestion. We have carefully proofread the entire manuscript. In addition to the specific corrections identified by the reviewer, grammatical errors, inconsistent terminology, abbreviations, punctuation, units, references, and figure/table citations have been systematically corrected throughout the manuscript.

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 your valuable suggestion. The manuscript text and captions have been revised to specify variables, units, time periods, spatial sampling, and threshold definitions.

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.

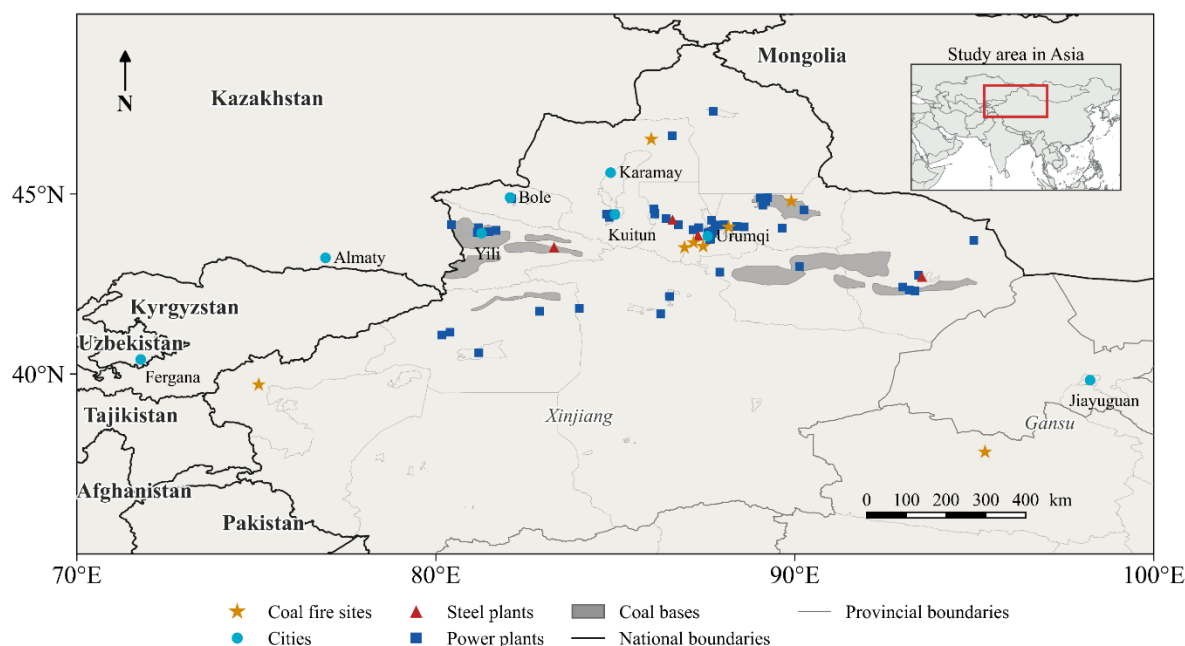


Figure 1. Study area (35°–50° N, 70°–100° E). The mapped coal bases and industrial facilities provide geographic context for the source-category analysis and are located in Xinjiang and adjacent parts of Gansu and Qinghai. For the 2019–2024 category-level statistics, only continuously operating facilities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

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 67-69: “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 69-70: “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 374-376: “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.

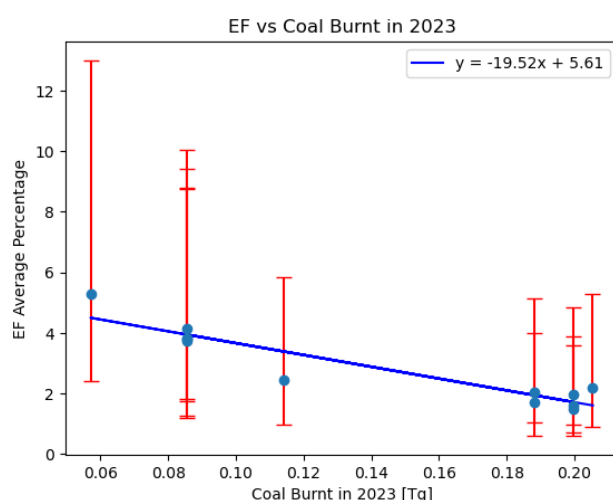
Response: We thank the reviewer for pointing out this ambiguity. The reviewer’s interpretation is correct: the metric represents the mean number of days per month for which a valid emission estimate is available in each grid cell, averaged across the study domain. The valid-day series was included as 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 340–341: “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 341–343: “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 349–350: “The valid-day series is presented only as a sampling diagnostic and is not interpreted as a physical 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.

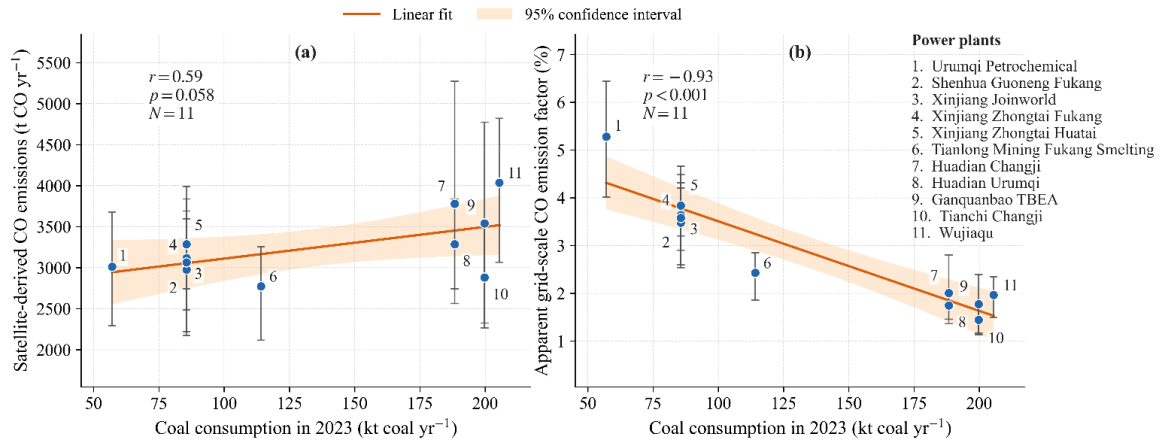


Figure 9. Relationship between satellite-derived CO emissions and coal consumption at selected coal-fired power plants. (a) Comparison between annual coal consumption and MFIEF-derived CO emission fluxes for individual power-plant grid cells. (b) Relationship between coal consumption and the grid-scale 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 396–397: “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 400–406: “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.

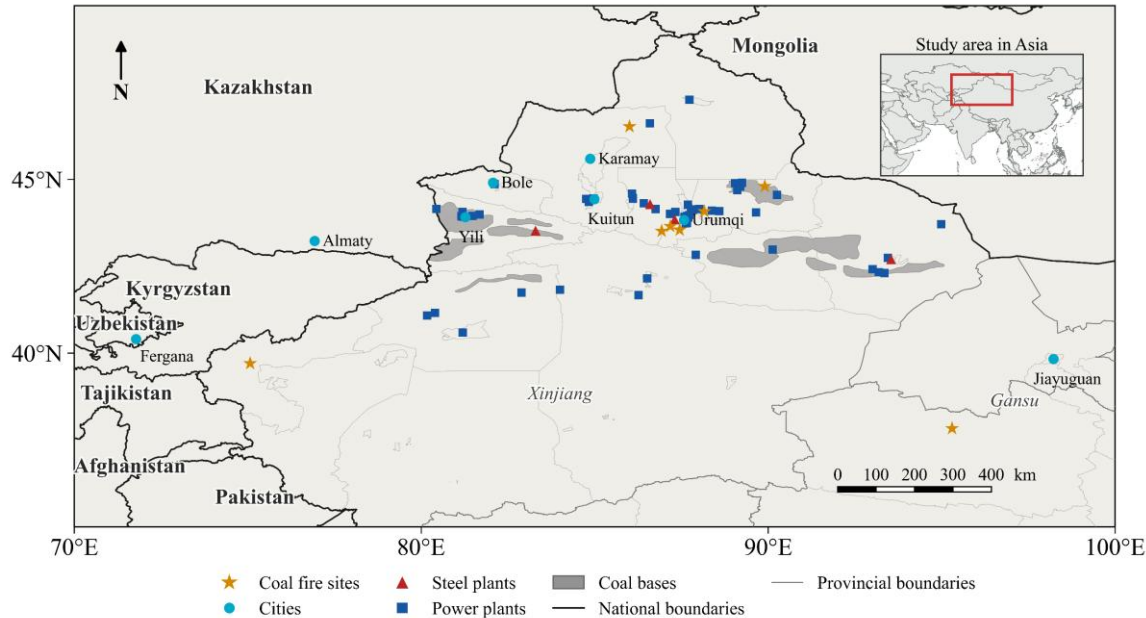


Figure 1. Study area (35°–50° N, 70°–100° E). The mapped coal bases and industrial facilities provide geographic context for the source-category analysis and are located in Xinjiang and adjacent parts of Gansu and Qinghai. For the 2019–2024 category-level statistics, only continuously operating facilities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

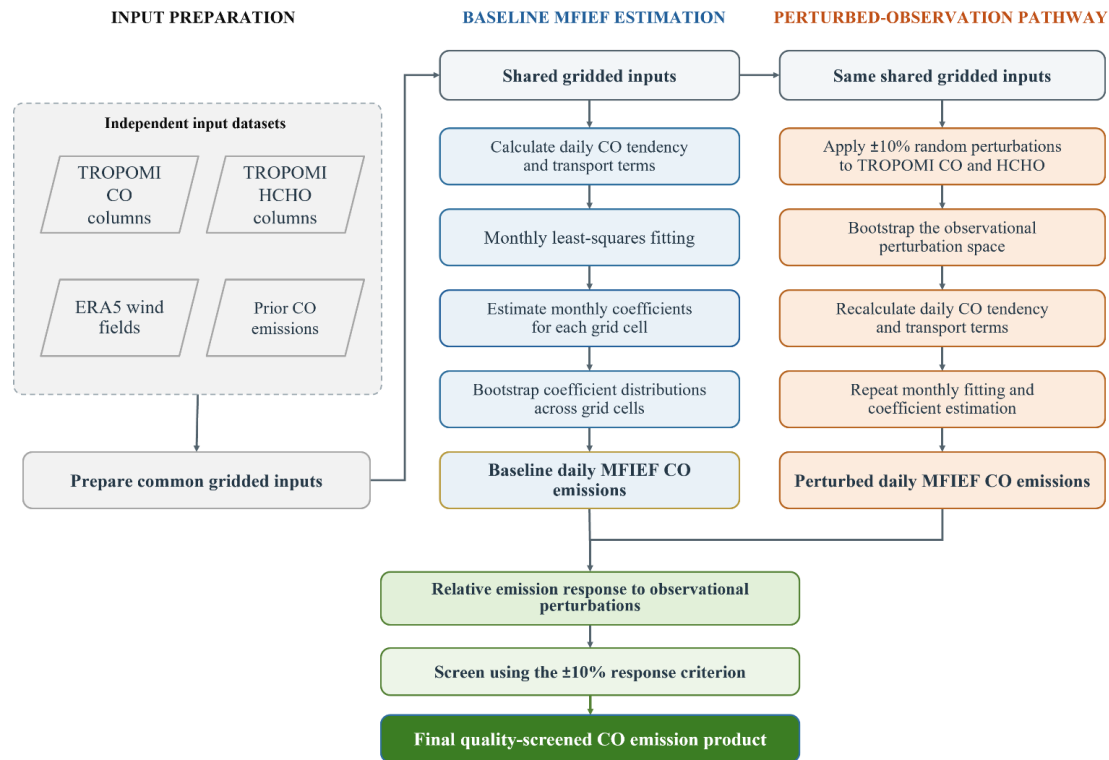


Figure 2. Workflow of MFIEF-based CO emission estimation and uncertainty-propagation analysis based on $\pm 10\%$ perturbations of TROPOMI CO and HCHO observations.

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.

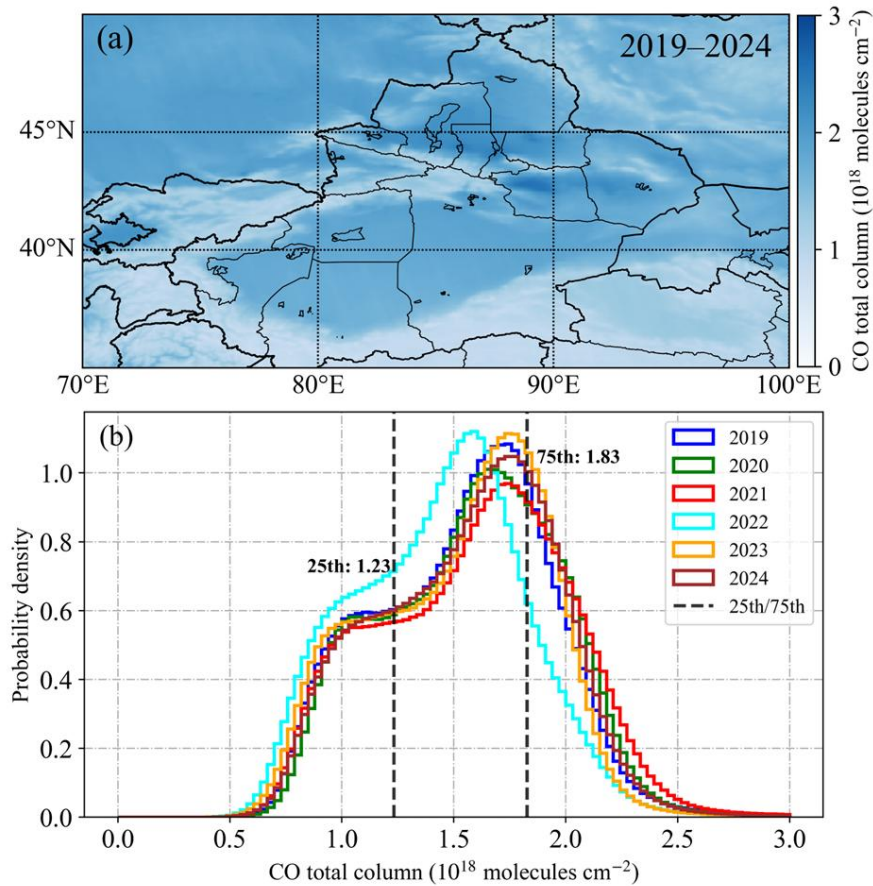


Figure 3. Spatial and statistical distributions of TROPOMI CO total columns over the study area during 2019–2024. (a) Mean spatial distribution over the six-year period. (b) Annual probability density distributions of CO total-column values over the study area.

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.

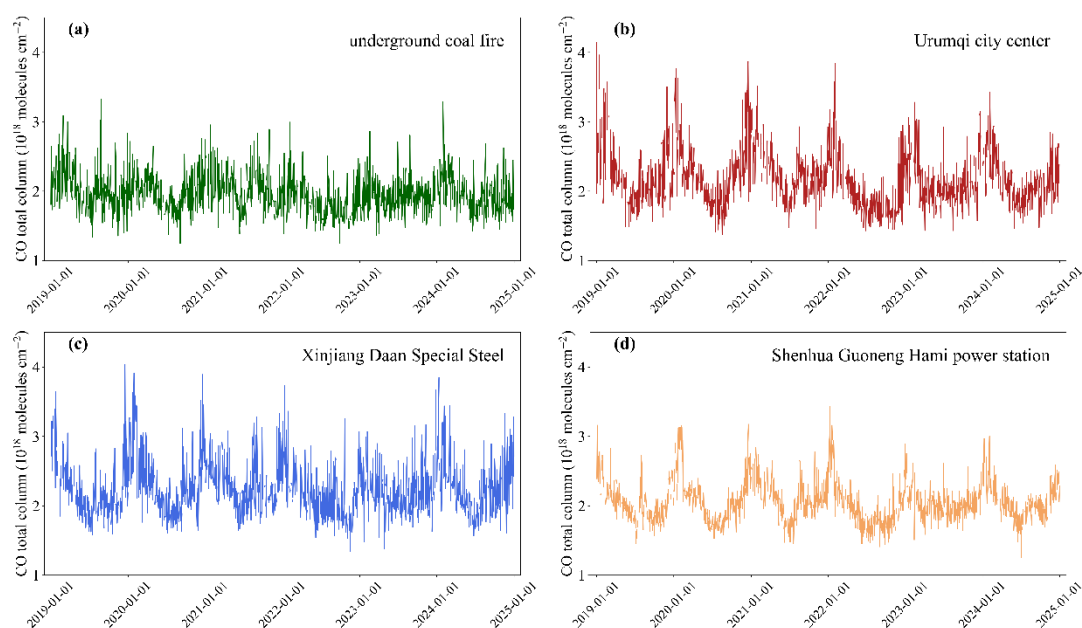


Figure 4. Time series of TROPOMI CO total columns in the $0.05^{\circ} \times 0.05^{\circ}$ grid cells containing (a) the Wugonggou underground coal-fire site, (b) the center of Urumqi, (c) Xinjiang Daan Special Steel, and (d) Shenhua Guoneng Hami Power Station during 2019–2024.

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.

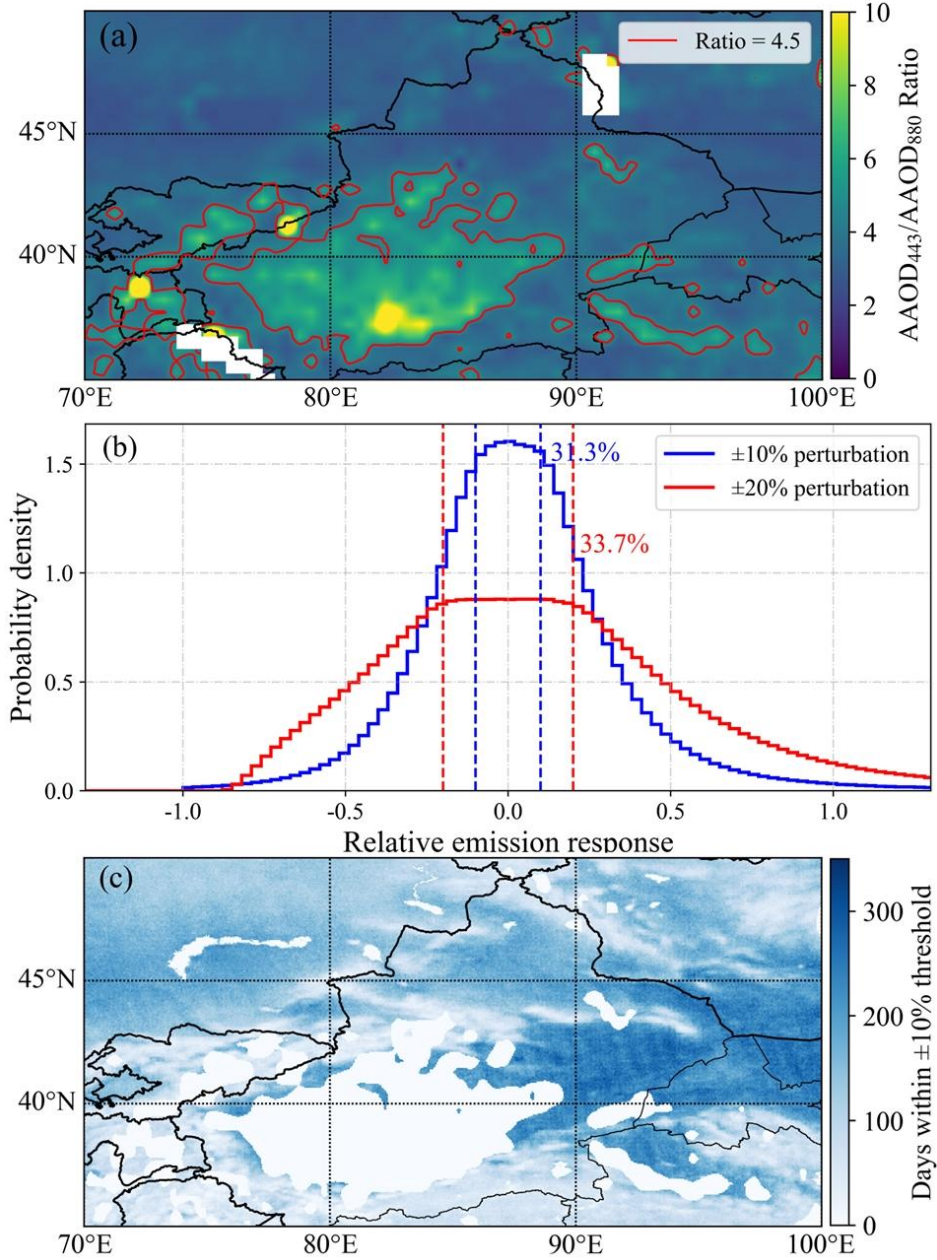


Figure 5: (a) Spatial distribution of the mean AAOD443/AAOD880 ratio during 2019–2024. Grid cells with ratios greater than 4.5 were identified as dust-dominated areas and excluded from the subsequent CO emission inversion. (b) Probability density distributions of the relative emission changes resulting from $\pm 10\%$ and $\pm 20\%$ 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 $\pm 10\%$ 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.

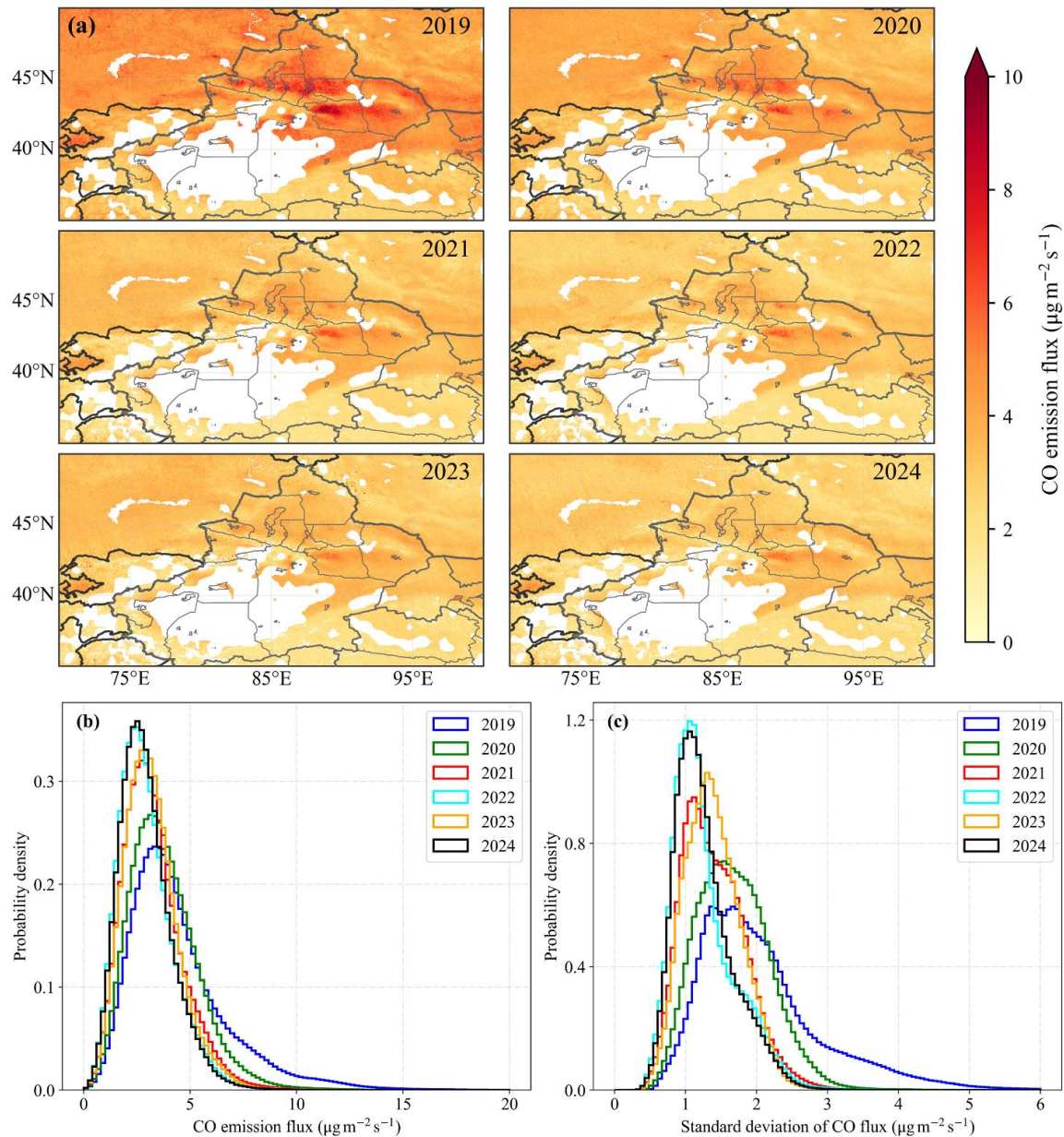


Figure 6. Annual characteristics of MFIEF-derived CO emissions from 2019 to 2024. (a) Annual mean CO emission flux; grid cells with fewer than 10 valid emission-estimate days in a year are shown as missing values. (b) Probability density distributions of annual mean CO emission fluxes. (c) Probability 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.

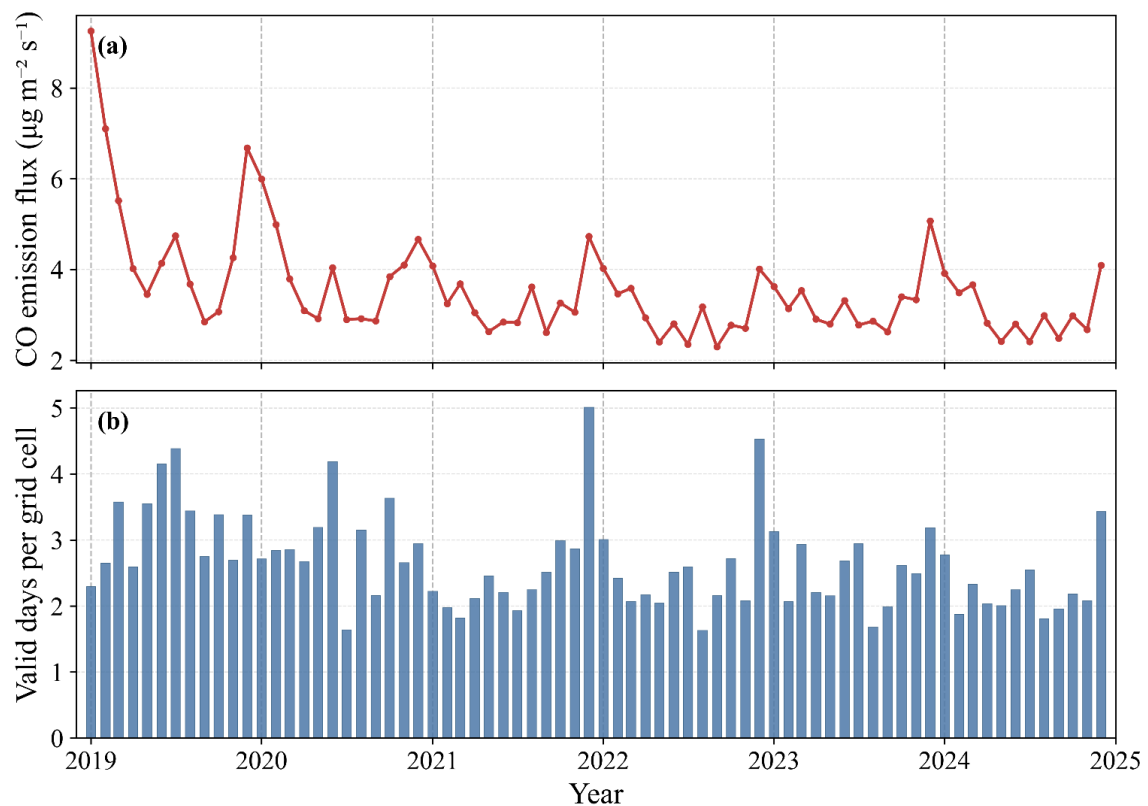


Figure 7. Monthly mean CO emission flux and temporal sampling coverage over the study region from January 2019 to December 2024. (a) Spatially averaged monthly CO emission flux derived from grid cells satisfying the observational-quality criteria. (b) Mean number of valid observation days per grid cell 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.

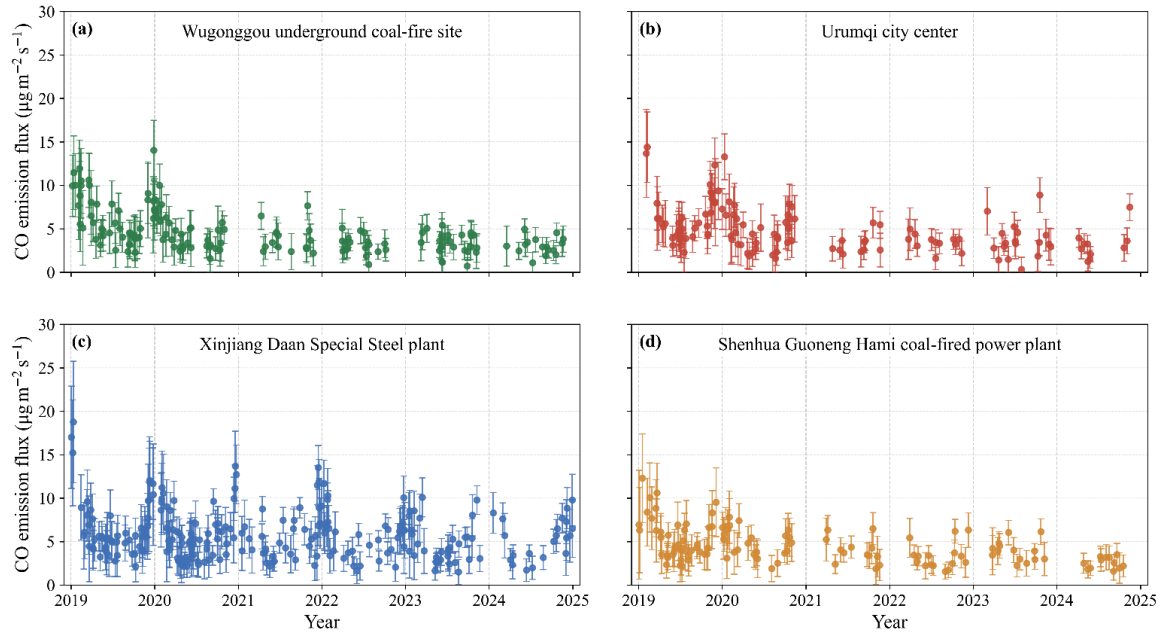


Figure 8. Temporal variations in MFIEF-derived CO emission fluxes at four representative locations from January 2019 to December 2024: (a) the Wugonggou underground coal-fire site, (b) Urumqi city center, (c) the Xinjiang Daan Special Steel plant, and (d) the Shenhua Guoneng Hami coal-fired power plant. 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.

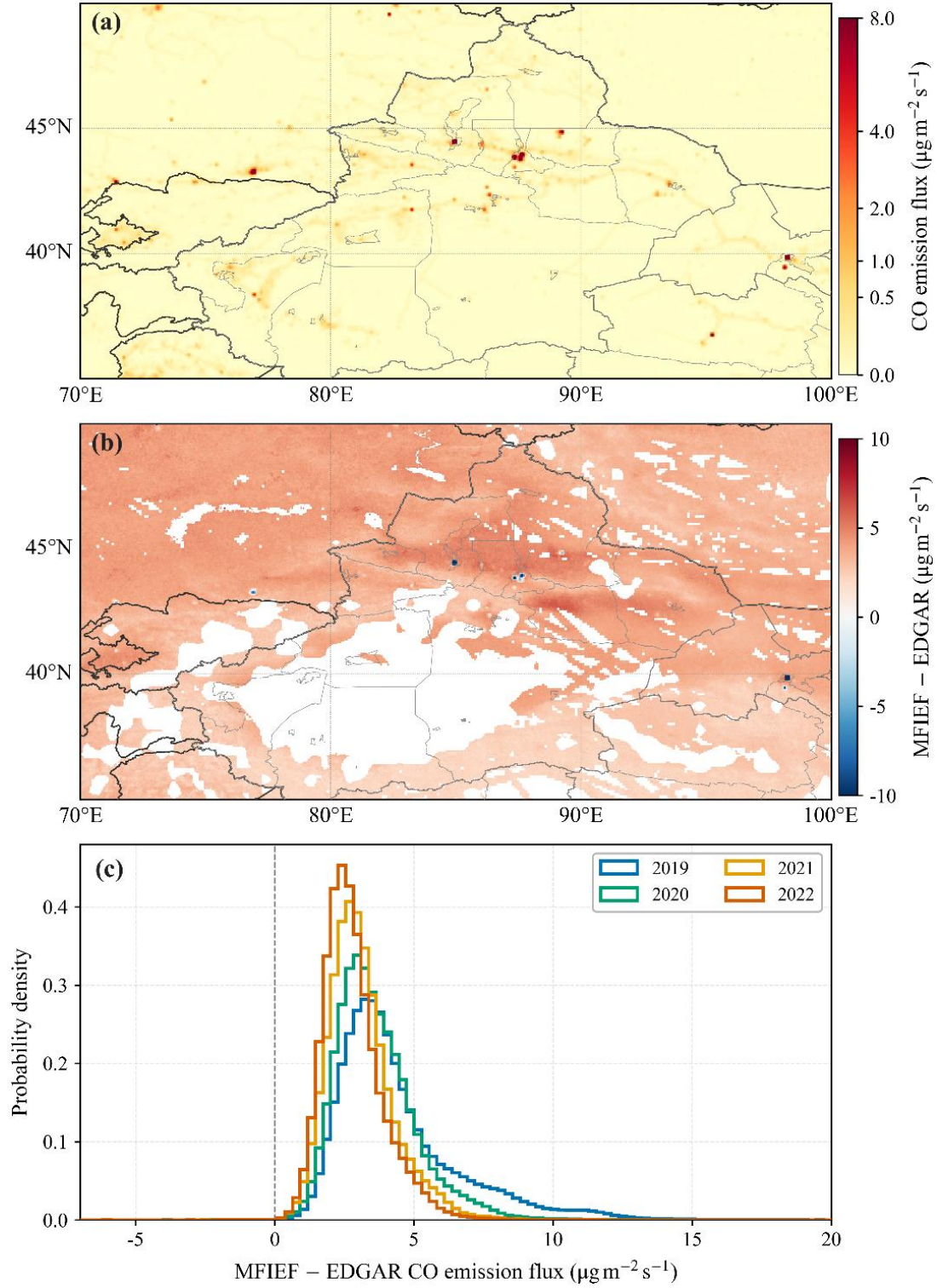


Figure 10. Comparison of CO emission estimates from MFIEF and EDGAR during their overlapping period from 2019 to 2022. (a) Mean EDGAR anthropogenic CO emission flux. (b) Mean grid-cell difference calculated as MFIEF minus EDGAR. Positive values indicate larger MFIEF-derived fluxes, whereas negative values 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.

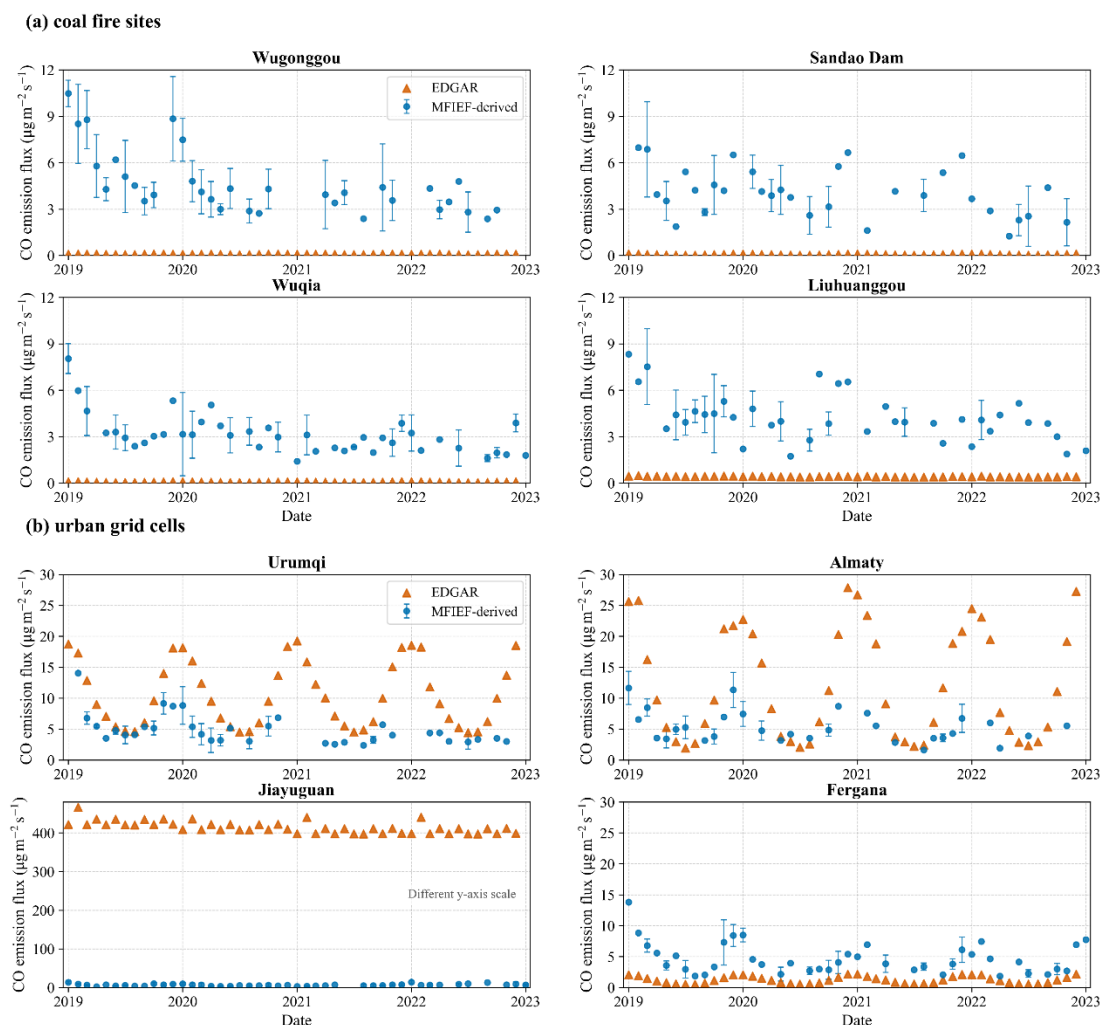


Figure 11. Comparison of MFIEF-derived grid-scale CO emission fluxes and EDGAR anthropogenic CO emissions at selected locations. (a) Persistent underground coal-fire sites, which are not explicitly represented as a separate source category in EDGAR. (b) Selected urban centers. MFIEF-derived emissions 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

Minor comments/corrections

Line	Type	Original	Correction	Comment	Response
12	Addition	marked interannual decline	—	Across the study period	Revised manuscript, Abstract, line 13. “Regional mean CO emissions decreased by 37.9% between 2019 and 2024, with a significant seasonally adjusted trend of $-0.33 \mu\text{g m}^{-2} \text{s}^{-1} \text{yr}^{-1}$ ”
22	Addition	significant environmental problems	—	Expand on this	Revised manuscript, Introduction, lines 26–27: “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	... major sink of the hydroxyl	—	Revised manuscript,

		radical (OH) affecting the concentrations of...	radical (OH), which in turn affects the concentrations of...		Introduction, lines 31–32.
29	Restructure	Once emitted into the atmosphere, CO undergoes chemistry in-situ including oxidation into carbon dioxide (CO ₂) and production from the oxidation of volatile organic compounds	—	CO can't be produced if it's already been emitted, needs rephrasing.	Revised manuscript, Introduction, lines 36–37.
31	Modify	In unpolluted atmosphere...	Under unpolluted atmospheric conditions...	—	Revised manuscript, Introduction, lines 38–39.
32	Remove	...in winter...	—	—	Revised manuscript, Introduction, lines 38–39.
33	Modify	...higher values during times controlled mostly by...	...elevated concentrations...	—	Revised manuscript, Introduction, lines 41–43.
37	Modify	...due to when and where...	...when...	—	Revised manuscript.
41	Add	China's MEIC system	—	Define acronym	Revised manuscript, Introduction, lines 49-50.
44	Add	WMO GAW	—	Define acronym	Revised manuscript, Introduction, line 54.

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 53–55.
49	Remove	integrate atmospheric transport and chemistry to infer emissions from concentration fields	—	duplicated	Revised manuscript
60	Modify	...the emissions of CO is both...	...the emission of CO is...	—	Revised manuscript
61	Modify	Though	thought	—	Revised manuscript
65	Modify	As related...	Those related...	—	Revised manuscript
65	Restructure	non-power generating industries, as well as for industry and power production	—	So... all forms of industry then?	Revised manuscript, Introduction, lines 72–75.
84	Restructure	extensive coal mining as well as many industries which use the coal and are powered by the coal.	—	Repeat coal too much	manuscript, Sect. 2.1, lines 93–94.
86	Remove	Also	—	—	Revised manuscript
94	Add	TROPOMI	TROPOspheric Monitoring	—	Revised manuscript

			Instrument (TROPOMI)		
95	Modify	Column loading	Total column	—	Revised manuscript
95	Modify	Retrieved daily	retrieved	—	Revised manuscript
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	Revised manuscript, Sect. 2.2, lines 116–117.
97	Remove	CH ₄ , and aerosols	—	Not relevant to this study	Revised manuscript, Sect. 2.2, lines 116–117.
100	Remove	And then resampled	—	repeated	Revised manuscript
101	Restructure	The data were then resampled with a resolution of $0.05^\circ \times 0.05^\circ$.	—	Is the resampling repeated? Confusingly worded here.	Revised manuscript, Sect. 2.2, lines 118–120.
103	Modify	HCHO is one of VOCs...	—	—	Revised manuscript, Sect. 2.2, lines 122–123.
105	Restructure	lifetime of HCHO columns is far shorter than those of CO columns,	—	Lifetime is for chemicals, not concentrations	Revised manuscript, Sect. 2.2, lines 125–127.
108	Modify	its uncertainty is larger than for CO	It is less sensitive to HCHO than CO	—	Revised manuscript, Sect. 2.2, lines 125–126.
109	Modify	In this work, concurrent data of	In this work, concurrent total column observations	—	Revised manuscript
110	Modify	GEE is utilized,	GEE is utilized and resampled by HARP	—	Revised manuscript

		resampled by HARP			
112	Restructure	wind:	—	I'm not a fan of these headers to the paragraphs. Wind should at least be capitalised.	Revised manuscript, Sect. 2.2, line 131
115	Modify	The 05:00 and 06:00 UTC times correspond to the observation periods of the TROPOMI satellite	—	Presumably this only applies over the region of interest in this study?	Revised manuscript, Sect. 2.2, lines 133-134.
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 $0.05^{\circ} \times 0.05^{\circ}$.	—	Revised manuscript
123	Add	nine distinct angles	Nine distinct viewing angles	—	Revised manuscript
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.	Revised manuscript. The resampling was performed only to match the spatial resolution of the emission estimates and other datasets, rather than to provide additional spatial information.
130	Modify	Equation (2)	—	Do you mean equation 1? If not, why reference eq 2 first?	Equation (2) was corrected to Equation (1). Revised manuscript, Sect. 2.3, line 161.
132	Modify	approximation of	approximate	—	Revised manuscript

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 terms do you mean the alpha terms?	Revised manuscript
142	Modify	these terms	Emission, production, and transport terms	—	Revised manuscript
142	Modify	Eq. (1)	Equation (1)	—	Revised manuscript
143	Modify	after converted the unit into	in units of	—	Revised manuscript
149	Restructure	The terms of this equation are initialized using this equation,	—	Repetition of this equation, reference the specific equation (eq 1 I think)	Revised manuscript
151	Modify	grid-by-grid	—	Unclear what this means. Gridcell-by-gridcell?	Yes, it means. grid cell-by-grid cell. Revised 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	—	Too long sentence	Revised manuscript

		each pixel specifically grouping all daily data from TROPOMI and the ECMWF reanalysis where they all occur.			
154	Add	weighting terms	weighting terms α_i	—	Revised manuscript
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.	Revised manuscript, Sect. 2.3, line 176.
159	Modify	The resulting equation	Equation (1)	—	Revised manuscript
159	Modify	Grids	gridcells	Across the paper too	Revised manuscript
213	Modify	biomass incomplete combustion	—	Specify wildfires, agricultural, both?	Revised manuscript, Sect. 3.1, lines 246–248.
214	Add	The statistical distribution of the CO total column	—	Averaged over what time period?	Revised manuscript, Sect. 3.1, lines 248–250.
224	Add	The CO total column on a daily basis is provided in Fig. 4 for the	—	Over what spatial location? Closest TROPOMI observation? Averaged over gridbox?	Revised manuscript, Sect. 3.1, lines 260–262.

		following sites...			
224	Add	Wugonggou Coal Fire in Fukang	—	Brief background on what this is. Natural? Long-term? Remote? Underground?	Revised manuscript, Sect. 3.1, lines 268–269.
225	Add	representative power plant, a representative steel plant	—	How are they determined to be representative? Which ones were chosen?	Revised manuscript, Sect. 3.1, lines 260–262.
225	Add	center of Urumqi	—	Presumably a city? Where is it located?	Revised manuscript, Sect. 3.1, lines 260–262.
248	Add	thereby minimizing potential biases in CO emission estimates associated with dust contamination	—	Provide a source and/or explain further	Revised manuscript, Sect. 3.2, lines 275–282.
260	Add	135 days	135 valid days	—	Revised manuscript
271	Remove	Been	—	—	Revised manuscript
272	Remove	with a reduction of approximately 38%	a reduction of 38%	—	Revised manuscript
273	Modify	which are defined as regions where the concentration is exceeded twice the study area's mean	which we define as regions where the concentration exceeded twice the study area's mean	—	Revised manuscript
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	—	Revised manuscript

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 EDGAR/MEIC emission-factor comparison was removed; a numerical magnitude comparison is provided in the revised manuscript, Sect. 4, lines 437–441.
354	Modify	A comparison between this emission and EDGAR	—	What emission?	Revised manuscript, Sect. 4, lines 407–408.
366	Restructure	In the figure, the blue areas of the figure are mainly distributed.	—	rephrase	Revised manuscript
375	Add	These statistics suggest that our inversion generally yields higher CO emissions than EDGAR	—	Why are emissions higher, why so distributed?	Revised manuscript, Sect. 4, lines 438–446.
395	Modify	we applied the MFIEF	—	Define MFIEF or at least mention it is an inversion algorithm	Revised manuscript
399	Remove	Both	—	—	Revised manuscript
401	Modify	all of which leads to	Both of which contribute to	—	Revised manuscript
402	Modify	The results reveal	We find	—	Revised manuscript, Sect. 3.2, lines 311–312.
403	Modify	that it is never the less	That CO is nevertheless	—	Revised manuscript

404	Add	with stronger emissions during winter and early spring	—	Add some statistics	Revised manuscript, Sect. 3.2, line 369.
408	Modify	urban agglomerations	Urban centers	—	Revised manuscript, Sect.4, line 423.
421	Modify	different directions	sources	—	Revised manuscript
421	Modify	on-the-ground	Ground-based	—	Revised manuscript
424	Remove	the issue of	—	—	Revised manuscript