#Reviewer 1

This manuscript presents a valuable effort to model global-scale trace metal PM_{10} concentrations and quantify their changes and the associated health impacts during the COVID-19 period. My primary concerns pertain to the emission inventory compilation and the modelling framework of this study. While the statistical indicators (R, RMSE, and the slopes) in Figure 1 suggest the robustness of the modelling approach, I believe the manuscript could still be strengthened in methodology and the clarity of presentation in related sections to enhance its comprehensiveness and utility to the broader research community:

Response: Thank for reviewer's suggestions. I significantly revised the manuscript based on the reviewer's suggestions.

Comment 1: The model-observation comparison appears to rely predominantly on data from industrialized regions (e.g., China, the United States, Europe). However, as this is a global-scale assessment, and since the manuscript discusses PM trends in regions such as South America, Sub-Saharan Africa, Russia, and Australia, it would be beneficial to incorporate observational sites from these less-represented regions into the evaluation to see how the model-obs statistics could be affected. Many of these areas are less industrialized compared to China/U.S./Europe- it would be interesting to see how the model performs in such regions where emissions are dominated by natural instead of anthropogenic sources. Doing so would help address potential biases stemming from a mid-latitude Northern Hemisphere focus, making the model more convincing. While I understand the scarcity of observational sites in these regions, incorporating even a small number of additional locations might offer valuable insights and enhance the credibility of the study's global-scale conclusions.

Response: I agree with reviewer's suggestions. Apart from the regular monitoring sites in China, the United States, and Europe, we also collected some scattered data in some other regions such South America and Sub-Saharan Africa from previous references (Table S14). We re-evaluated the predictive accuracy of particle-bound trace metals based on the new ground-level observation dataset. The detailed results are depicted in Figure 1.

Comment 2: While the primary focus of the manuscript is on quantifying changes in ambient trace metal PM levels before and during the COVID-19 period, the spatial distribution and source attribution of trace metals on a global scale is itself a critical contribution. I feel like that the current presentation, which primarily includes global mean concentrations and spatial maps, could be expanded to provide more in-depth diagnostics and benefit the broader atmospheric chemistry & biogeochemistry community. For example, what percentage of total emissions for each trace metal is attributable to anthropogenic sources versus natural sources? How do these proportions relate to the observed changes between the two study years? Such analysis would offer a valuable complement to the discussion of meteorological drivers. Including a breakdown of anthropogenic vs. natural contributions could also clarify the degree to which observed changes are emission-driven.

Response: Thank for reviewer's suggestions. Indeed, the isolation of natural and anthropogenic sources facilitates an in-depth discussion in this study (Line 241-285), and provides critical support for the meteorological emission separation debate. We have added the figures (Figure S14-S22) about the isolation of natural and anthropogenic sources in the revised version. Besides, we also have added some discussions to explain the spatiotemporal variations of ambient trace metal concentrations at the global scale. Our results suggested that the natural-derived trace metal concentrations in WE showed marked decreases (e.g., Cr (-16%), Cu (-18%), and Mn (-18%) during COVID-19 period, which was also beneficial to the decreases of trace metal concentrations in WE. In addition, the ambient trace metals in Australia were mainly sourced from the natural emissions (e.g., dust emissions), which was closely associated with the meteorological conditions. Therefore, the trace metal concentrations in Australia were not sensitive to the anthropogenic emission change during COVID-19 period.

Comment 3: Natural emissions, especially from soil dust, may constitute a significant fraction of total emissions for some trace metals. In relation to Supplementary Text 2, could the authors clarify the basis of the "average mass concentration of each element in soil"? Was this value derived from global crustal averages, or did it incorporate land-use and soil type variability? This is important, as trace metal concentrations (e.g., Pb, As) can vary substantially depending on local conditions, particularly in urban and industrialized settings where historical contamination is present. To increase transparency and robustness, I suggest: a) Listing the specific mass concentrations used for each element in the inventory; b) performing a sensitivity analysis to test how variation in these values (e.g., within plausible upper/lower bounds for different soil types) affects modelled concentrations. This would provide confidence that the results are not overly sensitive to potentially uncertain parameters.

Response: Thank for reviewer's suggestions. We have listed the specific mass concentrations used for each element in the inventory (Table S10). In our study, we only classified the dust into Asian dust and non-Asian dust because the trace element concentrations in different land use types were not available (the high-resolution soil trace element dataset was not available). In addition, we also performed the sensitivity analysis to test how variation in these values (e.g., within plausible upper/lower bounds for different soil types) affects modelled concentrations.

As shown in Table S11, we have assessed the sensitivity of the global average ratios of soil-derived ambient trace element concentrations and total concentrations (%) to soil trace element concentrations. The result suggested the contribution ratios of soil-derived trace elements increased by less than 9% when the soil element concentrations increased by 40%. The results confirmed the dust emission estimates are not overly sensitive to potentially uncertain parameters.

 # Reviewer 2

The manuscript "Different response characteristics of ambient hazardous trace metals and health impacts to global emission reduction" presents a model-based estimate of the 2020 minus 2019 health impacts of various trace elements, such as lead and arsenic. The authors develop emission inventories of these trace elements for 2019 and 2020, simulate their transport in a global chemical transport model, and estimate the health benefits due to atmospheric trace element concentration reduction due to the COVID-19 pandemic. This manuscript presents model-based evidence on the wide-ranging side benefits of emission reductions, especially in regions with already higher emission levels of toxic trace elements. The main findings of this work are that lead and arsenic reductions may have caused the largest health benefits among the trace elements considered, and sources such as coal combustion and smelting may have contributed the largest to their emissions reduction during the 3 months of the pandemic. However, the presentation in the manuscript could be highly improved. The methodology and its assumptions were not clear. The authors use only 3 months of data from 2019 and 2020 to estimate health effects. A short period like this may not be representative of long-term health effects and trends. Inventory, model, and health effect calculations were not presented clearly. Overall, I believe the manuscript's findings are useful in highlighting the positive impacts of trace element emissions reductions in different parts of the world, and I recommend the manuscript for publication after major revisions.

Response: Thank for reviewer's suggestions. We have significantly revised the manuscript based on reviewer's suggestions.

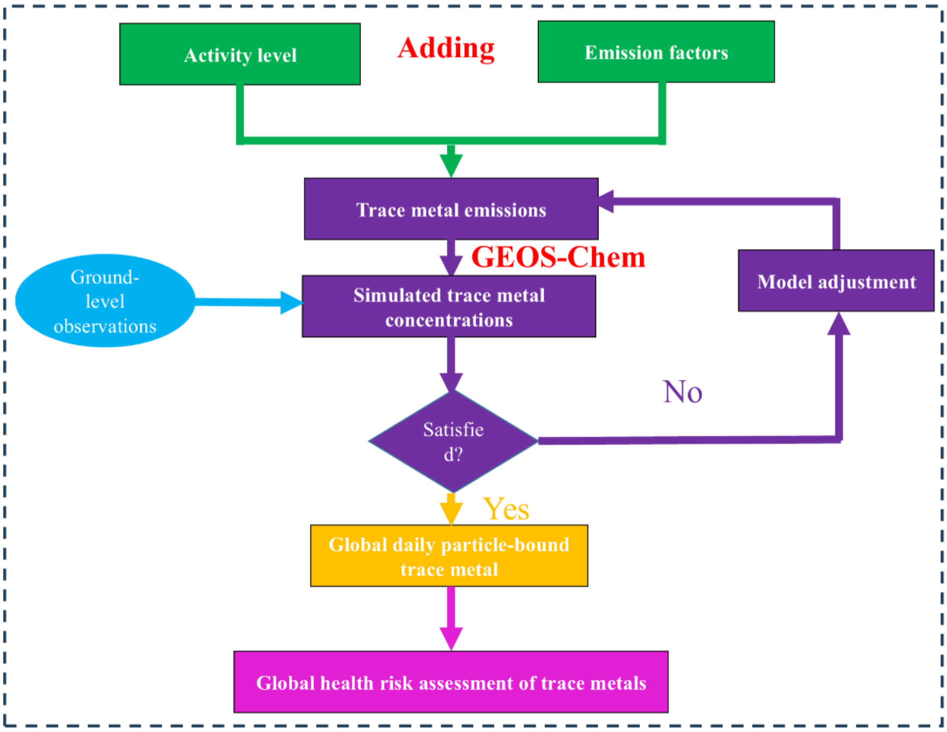
High-level Comments

Comment 1: Tighten the Introduction. The gaps and motivation for the design of the study were unclear. "What motivated you to perform this work the way you did?"

Response: Thank for reviewer's suggestions. Currently, some of studies only investigated the spatiotemporal variations of ambient trace metals in some isolated sites or at the regional scale. However, the global spatiotemporal variations of trace metals still remained unknown. Moreover, the response of ambient trace elements to changes in emissions is also unknown. COVID-19 case give us an unprecedented chance to investigate the emission reduction thresholds of trace metals in different regions around the world. From a health perspective, it remains unclear how we should reduce emissions and which pollution sources should be prioritized for control. Bridging this knowledge gap is crucial for implementing effective control and prevention measures targeted at specific trace elements. In summary, the purpose of our research is to guide how we can reduce emissions of trace metals. In addition, some redundant sentences also have been deleted.

Comment 2: Consider writing an overview of the Methods, highlighting its flow.

Response: Thank for reviewer's suggestions. The workflow has been shown in Figure S2 and the overview has been added in the revised version (Line 95-101).



Comment 3: Sec. 2.1 Inventory development was unclear. Consider tabulating it in the manuscript or supplementary information, including values from Zhu et al., used in this work. Consider describing the spatio-temporal resolution of the inventory, exact sources/sectors, and emission factors.

Response: Thank for reviewer's suggestions. We have added the detailed calculation methods, spatiotemporal resolution of emission inventory, exact sources/sectors, and emission factors. In our study, the spatial and temporal resolutions of emission inventory are 0.5° and monthly, respectively. The anthropogenic trace metal emission inventory includes coal combustion, liquid fuel combustion, ferrous metal smelting, nonferrous metal smelting, non-metallic minerals manufacturing, municipal solid waste incineration, and brake wear. The natural emission inventory includes dust, biomass burning, and sea salt spray. Emission factors have been introduced in Table S2-S13.

Comment 4: Sec. 2.2. Consider adding a table of the observations segregated by macro regions, species, number of observations, and time-period of observations.

Response: Thank for reviewer's suggestions. We have added the observations collected from previous references in Table S14. We only showed the data obtained from previous references. Apart from these data, we also incorporated many regular monitoring data in China, Europe, and the United States to validate the predictive accuracy of this model.

Comment 5: It was unclear why if 2020 emissions were used, why low-emission scenarios were performed? Moreover, findings from the 20-80% reduction scenarios did not seem central to the paper. The basis for the reduction scenarios and amounts was not clear.

Response: Thank for reviewer's suggestions. Indeed, the main content of our study is to compare the spatiotemporal variations of ambient trace metal concentrations during the "business-as-usual" and COVID-19 periods. COVID-19 case only confirmed the lockdown measures certainly decreased the concentrations of some trace metals, while the case did not give us sufficient knowledge to guide us how to control the trace metal emission. We hope to obtain more information about how to mitigate the ambient trace metal pollution at global scale. Therefore, we not only analyzed the impact of COVID-19 lockdown, but also we further performed the sensitivity experiment to identify the response of trace metal concentrations to precursor emission control. Based on the results, we could determine the optimal emission strategies and priority species. For instance, the sensitivity analysis confirmed that trace metal emission reductions are most effective in mitigating health damages in regions with high baseline exposures, such as China and India. Future efforts to target emission reductions in these regions could yield substantial public health benefits.

Comment 6: Comparing 2020 emissions against only 2019 meteorology may not provide a robust assessment of emissions-driven changes. This is because 2019 could be meteorologically anomalous, which could bias particle suspension and transport, even if an additional sensitivity simulation using 2020 emissions with 2019 winds is included. To better isolate the effects of emissions from meteorology, I recommend the following approach: Run one set of simulations using 2020 emissions and 2015–2019 MERRA-2 winds, and another set using 2015–2019 emissions with the same 2015–2019 meteorology. Comparing the monthly averages from these two experiments would minimize the influence of any single-year meteorological anomaly and more clearly attribute differences to emissions changes. If computer time is a limitation, I recommend performing at least 3 years of simulations for robustness.

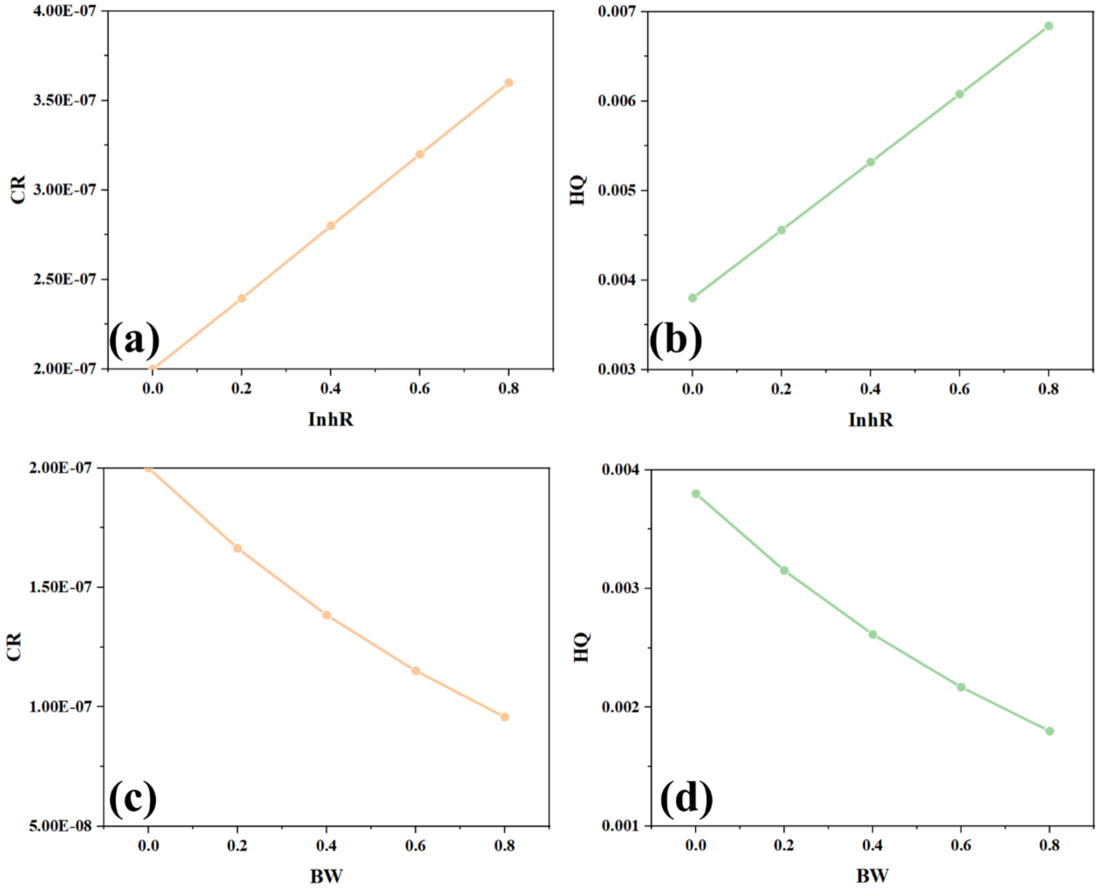
Response: I agree with reviewer's suggestions. (Section 2.4) We have performed three simulations including 2017–2019 meteorology+2017–2019 emission, 2017–2019 meteorology+2020 emission, and 2020 meteorology+2020 emission. Besides, all of the figures have been redrawn and the results were reanalyzed. The detailed revisions have been shown in the revised version.

Comment 7: The analysis focus in the Results section was not consistent – it fluctuated from microregions in one country to global macro regions in the same paragraphs.

Response: Thank for reviewer's suggestions. In fact, our study primarily focuses on the spatiotemporal variation analysis of trace metals across major global regions. Although China and India are individual countries, they actually represent vast regions due to their massive scale—comparable in magnitude to North America and Western Europe.

Comment 8: Sensitivity analysis is required to derive robust conclusions for health impacts. Consider compiling a broader set of risk values from the literature and using a lower and upper bound for the calculation. Then, a Monte Carlo-like sensitivity simulation could yield uncertainty bounds to health impacts. If computationally limited, I suggest using absolute lows and highs to obtain the lower and upper bounds.

Response: Thank for reviewer's suggestions. We have performed the sensitivity analysis to derive robust conclusions for health impacts. The results suggested that both of InhR and BW showed the approximately linear relationship with both of CR and HQ values (Figure S26). Overall, the results confirmed the health risk assessment model was robust because both of CR and HQ values did not show intense or irregular changes along with the linear change of InhR and BW. In addition, the lower and upper bounds for health risk assessment were also added in the revised version (Line 287-333).



Comment 9: Overall writing could be tightened.

Response: Thank for reviewer's suggestions. We have streamlined the content of the entire manuscript.

Specific Comments

Comment 10: L96: Detailed data of what? Please clarify.

Response: Thank for reviewer's suggestions. The detailed data including data sources of activity level and emission factors of multiple sources were shown in Table S1-S9.

Comment 11: L104: Refer to it as "Text S1".

Response: I agree with reviewer's suggestions. "Text S1" has been added in the revised version.

Comment 12: L114: Add exact reference and the table (or file) of the compiled observed dataset.

Response: Thank for reviewer's suggestions. We have added the observed dataset collected from references in Table S14. In addition, we also incorporated many regular monitoring data in China, Europe, and the United States to validate the predictive accuracy of this model. However, these regular observations were not added in Table S14 because the amount of the data is too large.

Comment 13: L59, 115, and others: Add the full form of the model names at least once.

Response: I agree with reviewer's suggestions. We have added the full form of the model when if appeared firstly. The full names of CMAQ and GEOS-Chem are Community Multiscale Air Quality and Global Modeling System for Chemistry, respectively.

Comment 14: L126: Are the natural sources different in 2019 and 2020 in the model? Clarify

Response: Thank for reviewer's suggestions. The natural sources in 2019 and 2020 are different in the model.

Comment 15: L127: Define "TE".

Response: Thank for reviewer's suggestions. (Line 143-144) "TE" has been replaced by "trace metal concentrations".

Comment 16: L164: Why is EF 365 days if the assessment is done only for a few months?

Response: Thank for reviewer's suggestions. In our study, we only calculated the total days in three months in modelling. Thus, we have changed d y^{-1} to d $period^{-1}$ in the revised version.

Comment 17: L178: Check superscript formatting (compared to L193).

Response: Thank for reviewer's suggestions. (Line 199) We have corrected the superscript formatting in Line 181.

Comment 18: L183-185: Explain this in detail. Why would simulating only one region make it underestimate the concentrations?

Response: Thank for reviewer's suggestions. In our study, we simulated the global trace metals in the atmosphere, while Liu et al. (2021) only simulated those in China. China often suffered from serious trace metal pollution and the predictive accuracy was not very high for most of trace metals because the heavy-pollution events were often difficult to capture. Therefore, the trace metal concentrations in China were often underestimated (simulations < observations).

Comment 19: L191: Shorten the section heading.

Response: Thank for reviewer's suggestions. We have shortened the section heading.

Comment 20: L201: It is unclear which exact regions are being analyzed in this manuscript. This can be clarified using a table or listing out the regions up front the methods.

Response: Thank for reviewer's suggestions. We have clarified all of the study regions in the methods (Line 153-156).

Comment 21: L228: It contradicts the previous statement. Most residential energy does not include coal combustion when globally averaged.

Response: Thank for reviewer's suggestions. In our study, coal combustion included coal-fired power plants, coal-fired industries, and coal-fired for residents. During lockdown period, the household energy consumption (e.g., coal combustion, natural gases combustion) even showed slight increase due to the stay-at-home order.

Comment 22: L261: The treatment of trace elements in the model should be clarified in the methods.
Response: Thank for reviewer's suggestions. (Section 2.4) In our study, we treated the deposition processes of trace metals similarly as aerosol particles because most (90% or more) atmospheric trace metals sorb onto aerosols especially fine-mode (i.e., $PM_{2.5}$) aerosols. Furthermore, the trace metals were generally considered to be inert, and thus the chemical reactions were not added in the trace metal modelling. Only physical processed such as emission, mixing, transport, and depositions were considered in the model. Wet deposition processes include sub-grid scavenging in convective updrafts, in-cloud rainout, and below-cloud washout (Liu et al., 2001). Dry deposition was calculated using a resistance-in-series model (Wesely, 2007). The model was driven by meteorological data assimilated from the MERRA2 reanalysis (Qiu et al., 2020). A global simulation at a 2 × 2.5° resolution was conducted to estimate the concentrations of hazardous trace metals on a global scale (Qiu et al., 2020).
Comment 23: L287: Why aggravation? Clarify.
Response: Thank for reviewer's suggestions. During COVID-19 period in 2020, unfavorable meteorological conditions (e.g., low wind speed and high relatively humidity) in North China Plain aggravated the air pollution and trace metal concentrations. Although the anthropogenic emission experienced rapid decrease during this period, the unfavorable meteorological offset the positive contribution of emission reduction.
Comment 24: L288-290: If so, I recommend performing health analyses for each of these regions to confirm this hypothesis. You can separate emissions and meteorology impacts using the framework already used in the manuscript.
Response: I agree with reviewer's suggestions. We have added the isolation of emission and meteorology to health risks (Table S17). (Line 324-331) "To confirm the assumption, we isolated the contributions of emission change and meteorology to the total changes of health risks. For instance, the emission-induced CR and HQ values of As, Cd, Cr, Cu, Mn, Ni, Pb, V, and Zn accounted for -111%, 95%, -121%, -113%, -112%, 129%, 60%, -111%, and 101% of the total changes after COVID-19 outbreak, respectively. The meteorology-induced CR and HQ values of As, Cd, Cr, Cu, Mn, Ni, Pb, V, and Zn accounted for 11%, 5%, 21%, 13%, 12%, -29%, 40%, 11%, and -1%, respectively. The results were in good agreement with our assumptions, indicating Chinese lockdown measures overcome the unfavorable meteorological conditions to decrease the health risks associated with the trace metal exposures (Table S17)." has been added in the revised version.
Comment 25: L293: It should be up front.
Response: Thank for reviewer's suggestions. (Line 169-171) This sentence has been placed on the methods in the revised version.
Comment 26: L313: Probably better to show emissions from individual sectors considered in this work and their 2020/2019 ratios for some representative regions.
Response: Thank for reviewer's suggestions. Your suggestions would indeed help provide a more comprehensive explanation of the spatiotemporal variations in atmospheric trace metals. However, as suggested by Reviewer #1, we have already quantified the proportional contributions of natural and anthropogenic sources of trace metals. This quantification, along with the separation of meteorological and emission influences, sufficiently explains the spatiotemporal differences in atmospheric trace metal concentrations. Further detailed categorization of anthropogenic source sectors is unnecessary. Currently, our manuscript already contains over 30 figures (figure + table are 48) (including SI). Adding further divisions of anthropogenic sources would significantly

increase its length and potentially dilute the key focus. We hope the reviewer can appreciate our considerations regarding the manuscript's structure and balance.

Comment 27: L322: This contradicts L313, which states that Pb increased globally, even if slightly. Please clarify.

Response: Thank for reviewer's suggestions. We have corrected this sentence to ensure the agreement of these sentences (Line 367 and 376). In fact, both of these sentences are not opposite. In our study, we found the global average ambient Pb concentration remained invariable (0.4%) during COVID-19 period indeed. It was assumed that the coal consumption for residential use did not significantly decrease during this period. The fact that global average Pb concentrations increased rather than decreased during the pandemic suggests that reducing Pb pollution is highly challenging—but this does not mean the policy recommendation to control Pb emissions is flawed. One of the key findings of this study is that prioritizing the control of Pb emissions particularly those from coal combustion is essential to effectively mitigate its health hazards.

Comment 28: Figure 1: What does the colorbar indicate? If there is no variability (all blue color), consider changing the colorbar scale to highlight any variability.

Response: Thank for reviewer's suggestions. Figure 1 is the Density Scatter Plot. The colorbar indicates the number of the sampling sites. In our study, different trace metals showed different samples. Some trace metals showed more ground-level observations, while other showed less observations. Some data may overlap, so a high-density scatter plot is needed for visualization.

Comment 29: Figure 3: What do the violin plots show?

Response: Thank for reviewer's suggestions. The violin plots show the spatial variations of the concentrations for particle-bound trace metals in eight major regions.

Comment 30: Figure 5: Why is there inter-regional disparity in CR and HQ risks in China and India? I understand it is described in the manuscript but more clarification is needed (see above comment).

Response: Thank for reviewer's suggestions. (Line 308-355) We have added more explanations about the spatial difference of CR and HQ risks in China and India. The trace metal concentrations in most of regions across China displayed decreases during COVID-19 period, while some unique regions such as Northeast China and Southern coastal regions even experienced slight increases of trace metal concentrations. This is primarily because these regions have relatively high humidity, which facilitates the aggregation of fine particulate matter. Additionally, low wind speeds and limited environmental capacity hinder the dispersion of pollutants, leading to elevated concentrations of particulate matter and high levels of trace metals adsorbed onto these particles (Huang et al., 2021; Li et al., 2023). Previous studies have highlighted that these regions experienced persistent air pollution or increased metal concentrations during the lockdown, primarily due to unfavorable meteorological conditions (Li et al., 2023b).

India also showed the similar reasons with China. The trace metal concentrations in most of regions experienced rapid decreases in India, while only several sites showed slight increases, which was linked with the unfavorable meteorological conditions.