

Response to Reviewers

Manuscript number: egusphere-2025-6470

Title: Elucidating CO₂ accumulation and dispersion in a semi-enclosed bay industrial park using Lidar and WRF-GHG modelling

We sincerely thank Reviewer #2 for the thorough, constructive, and candid assessment of our manuscript. We have revised the manuscript carefully. The point-by-point responses are given below. All changes made in the revised manuscript are marked in blue.

Reviewer #2:

Major comments:

This paper presents an interesting demonstration of scanning lidar technology and joint CO₂ and wind measurements. The interesting measurements are the most compelling element of the manuscript. The quality of the graphics and writing is good. There is potential for making this work publishable. In my opinion, however, the atmospheric modeling effort is not very useful given the limits of the modeling configuration that was used and the primary results of the paper are not highly compelling. I cannot support publication of this work in its current form.

1. The primary results of this study have been known for decades and do not warrant a new publication. 1) Pollutants accumulate near the surface when the atmospheric boundary layer is stable and they disperse during the day. 2) High winds disperse pollutants. 3) Coasts can cause sea breezes that move pollutants. 4) Mountains can trap pollutants. In my opinion these are the primary results of the study. The first two are clearly illustrated in Figure 2. The third point is demonstrated with examples and a wind rose, but the cause of the SE winds is not proven. The last point is not shown definitively but is plausible. None of these, however, is a new finding. The results obtained here are only unique in that they have been obtained with an innovative observational system. I do not think that reproducing very well established understanding of air pollution meteorology with a new instrument is worthy of publication unless the point is to demonstrate the limits of the instruments. This manuscript does not focus at all on the limits of the instruments.

We thank the reviewer for this important and constructive comment. We agree that the diurnal accumulation–dispersion cycle, the negative relationship between CO₂ concentration and wind speed, and the general effects of meteorological conditions on pollutant transport are well-established and should not be presented as novel scientific findings.

In the revised manuscript, the diurnal CO₂ cycle and the CO₂–wind-speed relationship are therefore described only as site-specific observational characteristics. The manuscript now places greater emphasis on the high-resolution plume structures resolved by the lidar, the quantitative evaluation of three emission-inventory configurations, and the joint lidar–WRF-GHG analysis. Statements that were not sufficiently supported by the observations have also been removed or weakened.

The Discussion and Conclusions have been reorganised accordingly, with observational results, interpretations, and future outlooks presented separately.

In response to the reviewer's suggestion regarding lidar limitations, we have also added

analyses of lidar observations under different wind-speed bins and of wind fields at different heights in Sect. 3.1. A more comprehensive multi-scenario assessment of lidar applicability is beyond the scope of the present revision and is now identified as a future research direction.

2. The WRF configuration is poorly suited to the problem at hand, and the authors use this to make unjustified claims about general failures of the WRF-GHG simulation system.

2A. Meteorology. Simulating the atmospheric dynamics of a small, mountainous bay is certainly challenging. Attempting to simulate this system and illustrating the limits of a model to reproduce these dynamics is a worthwhile topic of study. This WRF configuration, however, has relatively coarse vertical resolution, coarse horizontal resolution given the scale of the problem at hand, and the surface data appear fundamentally flawed with the lidar site sitting in the ocean and, according to the authors, a poor representation of the local topography. It is not surprising that this configuration is unable to simulate the complex flows observed by the lidar. Further, there is very little direct focus on model evaluation save for some general results in Table 2 and Figure 3. Figure 4-6 include both lidar data and model output, but at very different spatial resolutions and will little direct comparison. Figure 7 avoids model-data comparison. A serious attempt to simulate the bay meteorology and compared to lidar observations would be interesting. This manuscript does not include such an effort.

We thank the reviewer for this careful assessment. We agree that the original manuscript made overly broad statements regarding the limitations of the WRF-GHG modelling system. In the revised manuscript, these statements have been weakened, and the conclusions are now restricted to the performance of the present model configuration rather than being interpreted as general limitations of WRF-GHG.

We have also reorganised the meteorological evaluation and separated it from the CO₂ simulation analysis. The meteorological results are now presented as model-evaluation context for the subsequent CO₂ analysis rather than as novel scientific findings.

We agree that sensitivity experiments involving boundary-layer parameterisation, terrain resolution, and VPRM parameters would provide further insight into the performance of WRF in this complex coastal environment. However, in the present revision we prioritised the treatment of the anthropogenic emission inventory. The above WRF sensitivity experiments are therefore identified in the outlook section as important directions for future work.

2B. CO₂. The model clearly entirely fails to simulate the CO₂ concentrations observed at very fine resolution as observed by the lidar. The coarse horizontal resolution of the model almost guarantees this outcome. The extremely coarse resolution (in space and time) anthropogenic flux model (EDGAR) completely guarantees this failure. EDGAR fluxes are never shown. The model has the capacity to track different components of CO₂, as noted by the authors, but this is never shown. The model, as configured by the authors, is certain to fail in reproducing the observed CO₂ concentrations. This is not an interesting result.

We agree that the original anthropogenic emission treatment was too coarse to adequately represent the local industrial CO₂ emissions observed by the lidar. We have therefore substantially revised the emission-inventory processing and the quantitative observation–model comparison.

As described in Sect. 2.3.3, three anthropogenic-emission configurations are now considered: raw EDGAR as the baseline, EDGAR_S with industrial-source spatial redistribution and land–sea

mixed-grid correction, and EDGAR_P with an additional SNAP3 vertical emission profile over 0–710 m. A lidar-derived diurnal time-scale factor is further applied to convert the monthly-static inventory into an intra-day varying emission field.

The effects of these three configurations are now quantitatively evaluated against the lidar observations in Sect. 3.2.2. The WRF-GHG CO₂ concentrations used for the comparison are extracted over the 80–300 m layer to better match the lidar observation heights.

The revised manuscript therefore focuses on how emission-inventory refinement affects the discrepancy between the simulated and observed CO₂ concentrations, rather than treating the mismatch obtained with the original raw EDGAR configuration as evidence of a general limitation of WRF-GHG.

3. The paper relies primarily on the lidar observations, but the lidar data are not available. This is not acceptable for publication.

We thank the reviewer for raising this important point regarding data availability. We agree that the accessibility of the lidar-derived data underlying the results presented in this study should be clearly stated. In the revised manuscript, we have revised the Data Availability statement accordingly to make the accessibility of the data clearer in Lines 716-720:

“The coherent DIAL observational dataset used in this study, including the CO₂ concentration and synchronous wind-field products, is being prepared for deposition in a public data repository with a permanent DOI. The repository link and DOI will be provided upon completion of the public archive. In the meantime, the dataset, together with the relevant processing codes and model results, is available from the corresponding author upon reasonable request.”

We hope that this revision provides a clearer description of how the data used in this study can be accessed.

I would like to suggest some directions for revision that I believe could yield a publishable manuscript.

Focusing on the limits of the lidar technology seems the most obvious to me. The authors have shown only three high concentration cases. Are there conditions where the lidar can no longer detect the industrial emissions? This could be coupled with expanding the study to describe what is observed under a wider variety of meteorological conditions. This could be helpful for illustrating the conditions where this observing system could be used to quantify emissions.

We sincerely appreciate the reviewer’s constructive suggestion. We agree that examining the lidar observations under a wider range of meteorological conditions is important for better understanding the applicability and limitations of the observing system. In the revised manuscript, we have therefore expanded the analysis in Sect. 3.1 by including lidar observations under different wind-speed bins and wind fields retrieved by the lidar at different heights. These additional analyses provide a broader characterisation of the lidar observations under different meteorological conditions, rather than focusing only on the three high-CO₂ cases presented in the original manuscript. A more systematic assessment covering a broader range of atmospheric and emission conditions will be pursued in future work and has been noted accordingly in the outlook section.

Evaluating WRF meteorological performance with the lidar in this complex environment would also be worthwhile if the authors put more effort into experimenting with the WRF

configuration. It would be helpful to know if, with good vertical resolution, improved horizontal resolution and a reasonably good representation of the terrain of the bay and the temperature of the water, the atmospheric dynamics of this mountain bay could be simulated or not.

We thank the reviewer for this valuable suggestion. We agree that additional sensitivity experiments involving the WRF configuration would be useful for further evaluating the model performance in this complex coastal environment. In the present revision, however, we prioritised improvements to the anthropogenic emission inventory because the emission representation was identified as an important source of the model–observation discrepancy in the original analysis. Accordingly, the revised manuscript focuses on the refinement and quantitative evaluation of the emission inventory. We have therefore included these aspects in the outlook section as important directions for future research.

Estimating CO₂ emissions from the industrial park would, of course, be very interesting. This could potentially be done with the lidar winds, a simple guess as to the regional fluxes, Lagrangian particle modeling to infer the upwind areas influencing the lidar CO₂ concentrations and some simple variational experiments. I think the WRF simulation as presented is too coarse in resolution to handle the atmospheric transport in a useful way. This is just one idea.

We appreciate the reviewer’s constructive suggestion and agree that estimating CO₂ emissions from the industrial park would be a valuable extension of the present study. As suggested by the reviewer, such an analysis could be developed by combining the lidar-derived wind fields with Lagrangian particle modelling and a variational assimilation or inversion framework.

In the present study, we have focused on improving the representation of anthropogenic emissions and quantitatively evaluating the resulting model–observation agreement. Building on this framework, a dedicated emission-inversion analysis represents a natural next step and has now been highlighted in the outlook section as an important direction for future work.

I hope these ideas are helpful.

We thank the reviewer again for these constructive suggestions. They have helped us to more clearly define the scope of the present study, distinguish the directly supported results from future research directions, and refocus the manuscript on the high-resolution lidar observations, the emission-inventory refinement, and the quantitative observation–model evaluation. Our detailed point-by-point responses are provided below. We hope that the revisions made in response to the reviewer’s comments have substantially strengthened the manuscript and that the revised version is now suitable for publication.

Detailed Comments:

1. Lines 12-15. “Although satellite observations, ground-based measurements, and numerical modelling frameworks have been widely utilized, these approaches inherently struggle to simultaneously achieve high temporal and spatial resolution.” This statement is too broad to be helpful. Please make this more specific or delete it.

We thank the reviewer for pointing this out. We agree that the original statement was overly general and did not clearly convey the specific motivation of this study. We have therefore removed this statement and revised the opening to focus directly on the study context and approach (Lines 12–15):

“Industrial parks are important concentrated sources of anthropogenic CO₂ emissions, making accurate monitoring and modelling essential for understanding local CO₂ accumulation and dispersion and supporting effective carbon management. This study combines coherent differential absorption Lidar (DIAL) observations with the Weather Research and Forecasting model coupled with greenhouse gas fluxes (WRF-GHG) modelling simulations to examine CO₂ accumulation and dispersion in the semi-enclosed Luoyuan Bay industrial park.”

2. A daily cycle in CO₂ accumulation and a negative correlation between CO₂ concentration and wind speed are both patterns that have been understood for many decades of the study of pollutants in the atmosphere. If these are the main results of the study, I fear that this work may not yield any new understanding of atmospheric CO₂.

We thank the reviewer for this important comment and agree that the diurnal accumulation–dispersion cycle and the negative relationship between CO₂ concentration and wind speed are well-established features of atmospheric pollutant transport and should not be presented as novel findings. We have therefore revised the Results, Discussion, and Conclusions accordingly. These patterns are now treated only as basic site-specific observational characteristics that provide meteorological context for interpreting the observed CO₂ variability.

The revised manuscript instead emphasises the substantive contributions of this study: (1) the high-resolution DIAL observations that resolve the spatial and temporal structure of industrial CO₂ plumes in the semi-enclosed Luoyuan Bay industrial park; and (2) the quantitative observation–model evaluation of three emission-inventory configurations (EDGAR, EDGAR_S, and EDGAR_P), which demonstrates how spatial redistribution, land–sea mixed-grid correction, vertical allocation, and temporal scaling improve the representation of observed CO₂ variability. Thus, the new insight of this study lies not in identifying the general dependence of CO₂ on wind speed, but in using high-resolution observations to diagnose and reduce emission-related biases in WRF-GHG simulations over a complex industrial coastal environment.

In the Abstract, the negative correlation between CO₂ concentration and wind speed is retained only as an observational characteristic rather than as a claim of novelty, while the main contributions are now explicitly attributed to the resolved industrial-plume structure, the refined emission inventory, and the joint observation–model diagnosis.

3. Lines 22-25. “While the WRF-GHG model reproduces the overall temporal variation, it systematically underestimates CO₂ levels (420-460 ppm). The discrepancy is attributed to the limited spatial resolution of the emission inventory and the model’s inherent constraints in capturing terrain–wind field interactions within the bay.” The WRF model can be deployed in a nearly infinite array of configurations. The authors cannot argue that they have found “the model’s inherent constraints” unless they have explored the model’s potential configurations.

We thank the reviewer for this important comment. We agree that the phrase “the model’s inherent constraints” was too broad because the original study did not systematically evaluate alternative WRF configurations. We have therefore removed this statement and rewritten the corresponding part of the Abstract to report only the quantitatively evaluated performance of the three emission-inventory configurations under the present WRF-GHG setup. The revised text in Lines 22-25 reads:

“EDGAR_S redistributes industrial emissions and corrects land–sea grids, reducing the mean

bias from -165.4 to -115.9 ppm. EDGAR_P further adds a Selected Nomenclature for Air Pollutants (SNAP3) vertical profile, raising the correlation with the lidar from 0.60 to 0.69. With EDGAR_P, WRF-GHG produces a diurnal CO₂ range of 437.3–749.1 ppm, although a near-source underestimate remains.”

The remaining discrepancy is now discussed only in the context of the specific model and emission configurations evaluated in this study, rather than being attributed to an inherent limitation of the WRF model.

4. Lines 25-27. “This study highlights the unique capabilities of coherent differential absorption Lidar, elucidates the key limitations of current modelling approaches, and provides a robust scientific basis for refining carbon verification systems and enhancing the performance of regional carbon models.” This study highlights the capabilities of coherent DIAL, but does not provide the basis for refining carbon verification systems. I also am not really sure what a “carbon verification system” is.

We thank the reviewer for this comment. We agree that the original statement regarding the refinement of carbon-verification systems was not sufficiently supported, and that the term “carbon verification system” was not clearly defined. We have therefore removed this claim and the associated terminology. The concluding part of the Abstract (Lines 27-28) has been revised as follows:

“These results demonstrate the value of the Lidar–WRF joint analysis for CO₂ plume monitoring and model evaluation in complex coastal-industrial areas.”

5. Lines 37-39. “Currently, industrial CO₂ emissions monitoring technologies mainly include in situ monitoring and remote sensing monitoring methods. In situ instrumental monitoring conducts real-time measurement of atmospheric CO₂ concentrations via fixed-site online devices but struggles to effectively cover large-scale areas (Gurney et al., 2002).” Tracking industrial emissions relies primarily upon accounting methods. Atmospheric monitoring, at this point, is used only in research settings save for a few rare exceptions. In addition, the Gurney et al., (2002) paper is archaic. It is not a good example of current research to monitor CO₂ emissions using atmospheric methods.

We thank the reviewer for this important clarification. We agree that industrial CO₂ emissions are currently tracked primarily through accounting-based approaches, while atmospheric measurements are still used mainly in research applications for independently characterising and evaluating emission signals. We have therefore revised the paragraph to avoid presenting in-situ and remote sensing observations as the primary operational methods for industrial emission accounting. We have also removed the outdated Gurney et al. (2002) citation and replaced it with the recent study of Aigner et al. (2026), which illustrates the current capabilities and spatial-coverage limitations of atmospheric CO₂ sensor networks. The revised text in Lines 38-44 reads:

“Industrial CO₂ emissions are currently quantified primarily through bottom-up accounting approaches based on activity data and emission factors. Atmospheric CO₂ measurements, although still used mainly in research applications, provide an independent means of characterising concentration enhancements and evaluating emission patterns. High-precision in-situ instruments, such as CRDS analysers, provide accurate continuous CO₂ measurements but are generally deployed at a limited number of sites because of their cost and operational requirements, restricting

their ability to resolve fine-scale spatial variability (Aigner et al., 2026)."

6. Lines 78-79. "The organic combination of the two provides a new research approach."

The use of DIAL CO₂ measurements is innovative. It is an overstatement, however, to claim that this is "a new research approach." Many previous studies have used a combination of atmospheric mole fraction and atmospheric transport observations combined with mesoscale atmospheric models to study regional atmospheric composition and pollution sources.

We thank the reviewer for this helpful comment. We agree that combining atmospheric observations with mesoscale atmospheric models is an established research framework and should not be described as a new research approach. We have therefore removed the overstated claim and revised the text as follows (Lines 73–75):

"Combining high-resolution DIAL observations with process-based WRF-GHG transport modelling helps link detailed observations with atmospheric transport processes and improves the interpretation of CO₂ accumulation and dispersion."

7. Line 98. "This further worsens local air pollution." This statement implies a comparison to an earlier state of polluted air. Perhaps you mean to say, "The terrain can trap local emissions causing poor local air quality."

We thank the reviewer for this helpful suggestion. We agree that the original wording "further worsens local air pollution" was imprecise because it implied a comparison with a pre-existing polluted state. Following the reviewer's suggestion, we have revised the sentence to more directly describe the effect of the semi-enclosed terrain on atmospheric dispersion in Lines 100–102:

"The semi-enclosed terrain, surrounded by mountains on three sides and facing the sea on one side, can inhibit atmospheric dispersion and favour the local accumulation of industrial CO₂ emissions."

8. Lines 103-106. "a range resolution of 120 m, and a time resolution of 1 min. It can simultaneously measure CO₂ concentrations and wind fields. The CO₂ measurement accuracy has been confirmed in prior validation, measured at a background concentration of 463 ppm, showing a mean error of 2.05 ppm and a standard deviation of 7.18 ppm."

Concerns: 1) If there is "prior validation" then please include a citation for this work. 2) Does the CO₂ measurement performance apply to the range and time resolution noted - 1 minute and 120 meters? This is implied but not stated clearly. 3) Please explain what a "mean error" represents. If the measurement is always biased, it can be corrected. Is this the case?

We thank the reviewer for these helpful comments and have revised the text accordingly.

(1) The "prior validation" refers to the field intercomparison between the coherent DIAL and a tower-based cavity ring-down spectrometer (CRDS; Picarro G2401) reported in Yu et al. (2024), with the quantitative statistics reported for 1–6 June. We have now added the explicit citation "(Yu et al., 2024)" to the revised text.

(2) According to Yu et al. (2024), the CO₂ concentration product has a range resolution of 120 m and a temporal resolution of 1 min. The validation statistics were obtained at these same resolutions. Specifically, the DIAL data used for comparison were 1-min retrievals with a 120-m range resolution, using the average CO₂ concentration over the 1920–2040 m range gate,

corresponding to the approximately 2 km distance of the CRDS instrument on the meteorological tower. The native 1-s CRDS measurements were also averaged to 1 min. Therefore, the reported mean bias of 2.05 ppm and standard deviation of 7.18 ppm correspond to a temporal resolution of 1 min and a range resolution of 120 m at this validation range. These statistics characterize the DIAL performance under the specific conditions of the validation experiment rather than representing a range-independent instrument specification. This information has now been stated explicitly in the revised manuscript.

(3) The previously used term “mean error” refers to the mean of the DIAL–CRDS differences over the six-day validation period (1–6 June). It therefore represents the overall mean bias between the two measurements rather than a fixed instrumental offset. The daily mean differences varied in both sign and magnitude; for example, the mean difference was –0.33 ppm on 6 June. This indicates that the observed bias was not constant throughout the validation period. Such variability may reflect changes in carrier-to-noise ratio and atmospheric conditions, as well as the spatial sampling difference between the lidar range volume and the point in-situ measurement. The standard deviation of 7.18 ppm characterizes the scatter of the DIAL–CRDS differences around their mean. Because the observed bias was not constant, applying a single offset correction would not be appropriate. We have therefore replaced “mean error” with “mean bias” in the revised manuscript to avoid ambiguity.

9. Lines 106-107. “When compared to an optical cavity ring-down spectrometer, the correlation coefficient reaches 0.91 and the root-mean-square error (RMSE) is 5.24 ppm (Yu et al., 2024).” Apologies, but this is confusing. How is this related to the previous performance metrics, especially the quoted standard deviation of the CO₂ measurement? Is this the same work? If so, why are these quantities different? Is the time and range resolution for this comparison also 1 minute and 120 meters? Please be precise when stating instrument performance specifications.

We thank the reviewer for pointing out this ambiguity. We agree that the original text did not clearly explain how the different validation statistics reported in Yu et al. (2024) were calculated.

(1) These quantities were obtained from the same DIAL–CRDS field intercomparison during 1–6 June, but they represent different statistical measures and were summarized differently. The mean difference of 2.05 ppm and its standard deviation of 7.18 ppm characterize the DIAL–CRDS concentration differences over the six-day validation period. In contrast, the correlation coefficient and RMSE were first calculated separately for each day, and the reported values of and RMSE = 5.24 ppm are the arithmetic means of the six daily values. Specifically, the daily correlation coefficients were 0.91, 0.97, 0.92, 0.93, 0.90, and 0.81, while the corresponding daily RMSE values were 7.61, 6.79, 4.95, 3.57, 4.06, and 4.44 ppm. Their six-day averages are therefore 0.91 and 5.24 ppm, respectively. Because these statistics are based on different aggregation procedures, the 7.18 ppm standard deviation and the 5.24 ppm mean daily RMSE should not be interpreted as directly interchangeable measures.

(2) All daily comparisons were performed using the same CO₂ product resolution. The coherent DIAL retrievals had a temporal resolution of 1 min and a range resolution of 120 m, with the 1920–2040 m range gate used to match the approximately 2 km location of the tower-based CRDS. The native 1-s CRDS measurements were averaged to 1 min for temporal matching. We have now stated these temporal and spatial resolutions explicitly in the revised manuscript.

(3) To avoid further confusion, we have revised the manuscript to distinguish clearly between

the statistics of the DIAL–CRDS differences and the daily correlation/RMSE analysis. In particular, and RMSE = 5.24 ppm are now described explicitly as the mean daily correlation coefficient and mean daily RMSE, rather than as general instrument-performance specifications.

The revised text in Lines 102-104 reads:

“During the six-day validation, the DIAL–CRDS differences had an overall mean of 2.05 ppm and a standard deviation of 7.18 ppm, while the mean of the daily correlation coefficients and daily RMSE values was 0.91 and 5.24 ppm, respectively (Yu et al., 2024).”

10. Lines 107-109. “For wind field measurements, the RMSE of horizontal wind speed is less than 0.6 m s^{-1} , the RMSE of wind direction is less than 11° , and the vertical wind speed measurement accuracy is 0.2 m s^{-1} (Yuan et al., 2022)” As above, do these metrics refer to 1 minute and 120 m temporal and spatial resolution data? How have these metrics been determined? How is vertical wind speed accuracy determined? Finally, this sentence is missing a period.

We thank the reviewer for the careful reading and for pointing out that the original sentence did not sufficiently distinguish the resolution and uncertainty specifications of the different wind products. We have revised the text accordingly.

(1) Temporal and spatial resolution: The wind measurements do not have the same temporal resolution as the CO₂ product. According to Yu et al. (2024), the range resolutions of the CO₂ concentration and wind measurements are both 120 m, whereas their nominal temporal resolutions are 1 min and 1 s, respectively. We have now stated these resolutions explicitly in the revised manuscript.

(2) Horizontal wind uncertainty: The values of $<0.6 \text{ m s}^{-1}$ for wind speed and $<11^\circ$ for wind direction are reported in Yuan et al. (2022) as the RMS uncertainties of the retrieved two-dimensional horizontal wind field rather than RMSE values. The horizontal wind vector is reconstructed from the line-of-sight radial velocities obtained during PPI scanning. We have therefore corrected “RMSE” to “RMS” and clarified that these values characterize the horizontal-wind retrieval reported in the cited study rather than the CO₂ measurement resolution.

(3) Radial and vertical wind velocity: We agree that the original statement “vertical wind speed measurement accuracy is 0.2 m s^{-1} ” was insufficiently precise. The value of 0.2 m s^{-1} reported in Yuan et al. (2022) refers directly to the accuracy of the line-of-sight radial velocity at the adopted CNR threshold of -27 dB . Vertical wind velocity in the present system is subsequently retrieved from multiple radial-velocity measurements through VAD scanning, and therefore the radial-velocity accuracy should not be interpreted directly as an independently validated accuracy of the retrieved vertical-wind component. We have revised the manuscript accordingly and now report the 0.2 m s^{-1} value as the radial-velocity accuracy rather than the vertical-wind-speed accuracy.

(4) Missing period: The missing period has been added.

11. Figure 1. Please explain the color legend in 1(b). Is this elevation of the terrain?

We thank the reviewer for pointing this out. The revised Figure 1 has been redrawn and reorganised into four panels. The colour scale in the original panel (b), now shown as panel (d), represents terrain elevation in metres above sea level (m a.s.l.). This information has now been explicitly stated in the figure caption.

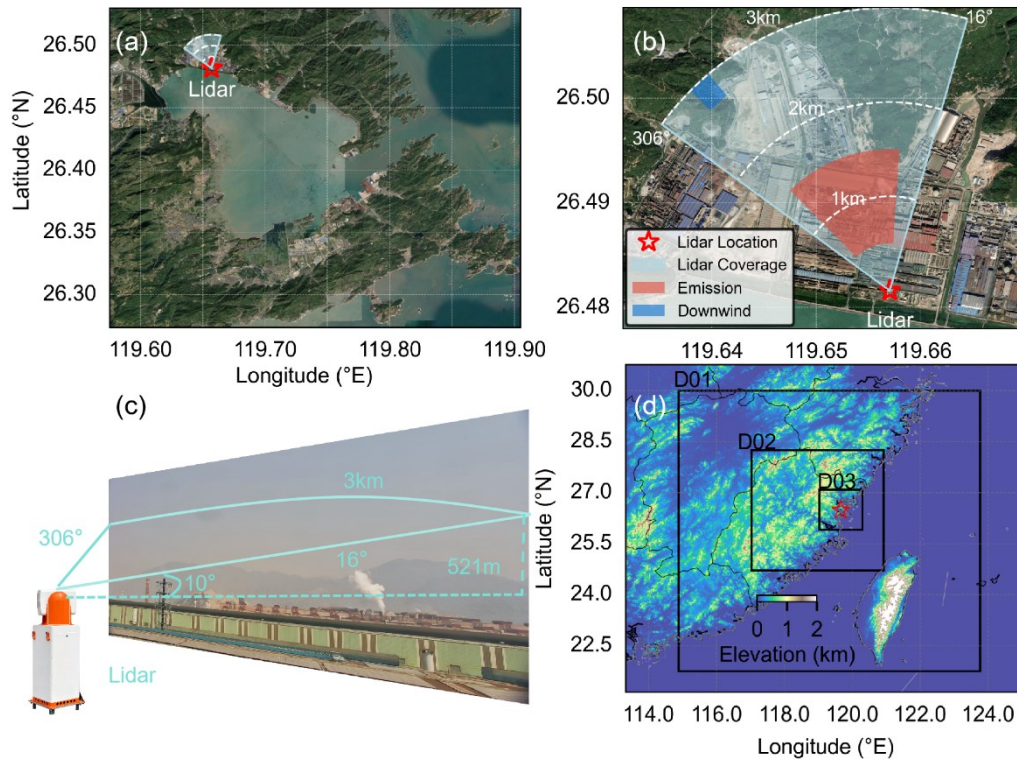


Figure 1. (a) Satellite map of Luoyuan Bay showing the lidar scanning coverage; the red five-pointed star indicates the lidar installation site. (b) The lidar's azimuth scanning range spans 306° to 16°. (c) Schematic of lidar scanning at a fixed elevation angle of 10°, with a vertical height difference of 521 m between the farthest detection position and the lidar platform. (d) Schematic diagram of the three-layer nested simulation domains of the WRF-GHG model. The red five-pointed star marks the lidar location. Map source: Microsoft Bing Maps (2026).

12. Figure 1. It would help to have a better view of the innermost domain and the bay. At present it is difficult to see how the lidar observational domain sits within the bay, and the size of the bay compared to the innermost WRF domain.

We agree and have redrawn Figure 1 accordingly. Panel (a) now includes a satellite map of Luoyuan Bay, which more clearly shows the location of the Lidar observational domain within the bay. The revised figure also provides a clearer view of the innermost WRF domain, making the spatial relationship and relative scale between the bay and the innermost modelling domain easier to identify.

13. Line 138-139. “The Copernicus Atmosphere Monitoring Service (CAMS)”. GHG boundary conditions are a critical element of the simulation and a new topic. This should be part of a new paragraph.

We thank the reviewer for this helpful suggestion. We agree that the GHG boundary conditions are an important and distinct component of the WRF-GHG configuration and therefore warrant a separate description. We have moved the CAMS GHG boundary-condition description into a separate paragraph in the Methods section. The revised text clarifies that the 3-hourly CAMS forecast fields are used only as lateral boundary conditions and explicitly states their spatial and vertical resolution (0.1° and 137 levels, respectively) and update interval (Lines 179-180):

“The initial conditions were read only once, at 00:00 UTC on 5 March, whereas the CAMS 3-hourly fields were used exclusively to update the lateral boundary conditions continuously at every 3 h interval throughout the run.”

14. Lines 142-143. “We selected the three-hourly interval products from the first day of each daily forecast, with a spatial resolution of 0.1° and 137 vertical model levels.” I don’t understand. What are these forecasts used for? The WRF simulation only needs one set of initial conditions but it needs continuous boundary conditions for the entire period of study. Does “each daily forecast” suggest that WRF is being run independently for each day of the study, with new boundary and initial conditions? Is any spin up time allowed? CAMS and WRF GHG fields could be quite different in equilibrium depending on both fluxes and transport. No spin up time could result in a simulation that is dominated by the modeled GHG fields adjusting to the differences between these two modeling systems.

We apologize for the lack of clarity and thank the reviewer for raising this important point. We have rewritten the paragraph to make the use of the CAMS fields and the WRF-GHG integration procedure explicit.

(1) Use of the CAMS forecasts and continuity of the WRF-GHG simulation: The WRF-GHG simulation was performed as a single continuous integration from 5 to 31 March 2025 and was not restarted each day. The CAMS GHG field at 00:00 UTC on 5 March was used to initialise the GHG tracers, and the subsequent CAMS fields were used to provide continuous lateral boundary conditions throughout the simulation. “Each daily forecast” in the original text referred to the fact that CAMS issues a new global GHG forecast each day, rather than to daily reinitialisation of WRF-GHG. From each CAMS forecast cycle, we selected the first 24 h of 3-hourly forecast fields and concatenated these fields to construct continuous 3-hourly lateral boundary conditions for WRF-GHG. We have revised the text to eliminate this ambiguity.

(2) Selection of the first-day CAMS forecasts: CAMS provides a new five-day global GHG forecast each day, with 3-hourly output. We used the first 24 h from each daily forecast cycle so that the lateral boundary conditions were consistently based on the shortest available forecast lead times, rather than using progressively longer-lead forecasts. This procedure provides a continuous sequence of 3-hourly GHG lateral boundary fields throughout the WRF-GHG integration.

(3) Spin-up: A 13-day spin-up period from 5 to 17 March was included and excluded from all subsequent analyses. This spin-up allows the WRF-GHG tracer fields to adjust from the imposed initial GHG distribution to the model transport and surface-flux configuration and substantially reduces the influence of initial-condition mismatch between CAMS and WRF-GHG. The analysis therefore begins on 18 March, and the WRF-GHG results discussed in this study cover 18–29 March. We have now stated the spin-up period and the analysed period explicitly in the revised manuscript.

15. Line 150. There is no discussion of anthropogenic emissions. Is WRF run without any anthropogenic CO₂ emissions estimate? Why? How is this comparable to the observations over this industrial area?

We apologise for the insufficient description in the original manuscript. Anthropogenic CO₂ emissions were included in all WRF-GHG simulations; however, we agree that their treatment was not adequately described. We have therefore added a dedicated subsection (Section 2.3.3, “*Anthropogenic emission settings*”) detailing the emission dataset, preprocessing procedure, and

localised corrections.

(1) Anthropogenic emission input: The original WRF-GHG simulations included anthropogenic CO₂ emissions using the EDGAR inventory, with all eight emission sectors introduced through the CO_{2,ant} tracer. In the original manuscript, only the data source was briefly mentioned, which understandably gave the impression that the anthropogenic-emission treatment had not been discussed. The original simulations used the 2023 emissions from EDGAR_2024_GHG. During the revision, EDGAR_2025_GHG (Crippa et al., 2025) became available, and we therefore updated the analysis using the latest available 2024 anthropogenic emissions.

(2) Refinement of the anthropogenic inventory: In the original analysis, the monthly EDGAR fields were mapped to the WRF domain using simple bilinear interpolation, without sub-grid spatial redistribution, vertical allocation, or intra-diurnal variation. The revised analysis introduces four refinements: (i) redistribution of industrial emissions to actual industrial parcels using the 2025 China industrial-land vector dataset; (ii) correction of land–sea mixed grids to remove spurious marine emissions while retaining traffic-related emissions; (iii) vertical allocation of industrial combustion emissions following the SNAP3 profile of Brunner et al. (2019) and Wang et al. (2026); and (iv) an observation-informed diurnal scaling factor derived from the lidar measurements to represent sub-daily variability. These treatments define the EDGAR_S configuration (spatially corrected) and EDGAR_P configuration (EDGAR_S plus vertical allocation).

(3) Comparison with the lidar observations: The revised manuscript evaluates three emission configurations quantitatively against the DIAL observations. Revised Figure 2 compares the raw EDGAR and EDGAR_S emission fields at 08:00 on 25 March 2025 and shows that the land–sea correction removes unrealistic marine emissions while the spatial redistribution concentrates industrial emissions over the actual plant areas. The subsequent observation–model comparison is used to quantify how these refinements affect the simulated CO₂ field over the industrial region. Because the diurnal scaling factor is derived from the lidar observations themselves, the corresponding experiment is interpreted as an observation-constrained sensitivity test rather than a fully independent validation.

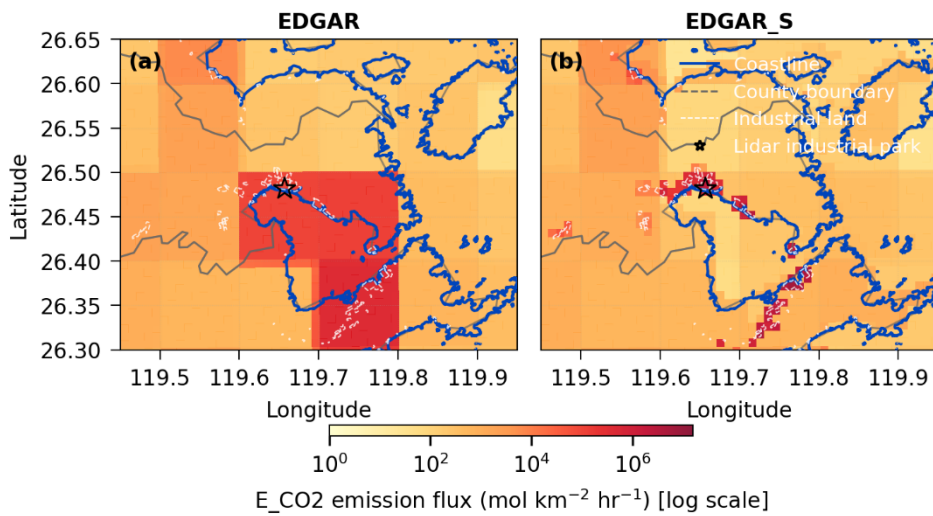


Figure 2. Comparison between raw EDGAR and EDGAR_S emission fields at 08:00 on 25 March 2025. EDGAR_S denotes the emission inventory after industrial-source spatial redistribution and land-sea mixed-grid correction.

Relevant references:

Aigner, P., Chen, J., Böhm, F., Chariot, M., Emmenegger, L., Frölich, L., Grange, S., Kühbacher, D., Kürzinger, K., Laurent, O., Makowski, M., Rubli, P., Schmitt, A., and Wenzel, A.: ACROPOLIS: Munich urban CO₂ sensor network, *Atmos. Meas. Tech.*, 19, 745–773, <https://doi.org/10.5194/amt-19-745-2026>, 2026.

Brunner, D., Kuhlmann, G., Marshall, J., Clément, V., Fuhrer, O., Broquet, G., Löschner, A., and Meijer, Y.: Accounting for the vertical distribution of emissions in atmospheric CO₂ simulations, *Atmos. Chem. Phys.*, 19, 4541–4559, <https://doi.org/10.5194/acp-19-4541-2019>, 2019.

Wang, X., Jiang, F., Wang, H., Zhang, Z., Wu, M., Wang, J., He, W., Ju, W., and Chen, J. M.: The role of OCO-3 XCO₂ retrievals in estimating global terrestrial net ecosystem exchanges, *Atmos. Chem. Phys.*, 25, 867–880, <https://doi.org/10.5194/acp-25-867-2025>, 2025.

16. Lines 158-161. “To accurately assess the impact of industrial park emissions on the surrounding environment, a forested area within the Lidar’s azimuth range of 315°–320° and radial distance of 2.7–3 km was selected as the downwind site. Dominantly composed of evergreen broad leaved forests, this area maintains a high photosynthetic rate year-round. With a vertical elevation difference of approximately 450 m from the Lidar installation, it effectively reflects the concentration level of pollutants after dispersion.” I am quite confused by this text. Why does a location 3km away and at a higher elevation “reflect the concentration level of pollutants after dispersion.” This is not nearly far enough downwind (if it is downwind) for emissions to be well mixed in the atmospheric boundary layer. Please explain what is meant by “after dispersion.” The altitude is also confusing. Is this ‘downwind site’ 450m above sea level? Or is 450 m the elevation in the atmosphere / above ground? as measured by the lidar?

We appreciate the reviewer’s careful reading and the opportunity to clarify this point. We agree that the original description was imprecise and could incorrectly imply that the plume was fully mixed after 3 km of transport. We have therefore rewritten this section.

(1) Clarification of the elevation: The reported elevation difference of approximately 450 m refers to the difference in terrain elevation between the forested hillside sector and the lidar site. It is neither the absolute elevation above sea level nor the atmospheric sampling height measured by the lidar. The lidar is installed approximately 30 m above the local ground and scans at a fixed elevation angle of 10°. At slant ranges of 2.7 and 3.0 km, the laser beam is approximately 469 and 521 m above the lidar optical level, respectively. Thus, the beam is approximately 499–551 m above the ground level at the lidar site. Because the forested terrain in this sector is approximately 450 m higher than the lidar-site ground level, the sampled air volume is only approximately 50–100 m above the local forested terrain. We have revised the manuscript and Figure 1(b–c) to make this geometry explicit.

(2) Clarification of “after dispersion”: We agree that a transport distance of approximately 3 km is insufficient to assume that an industrial plume is fully mixed throughout the atmospheric boundary layer. The phrase “after dispersion” in the original manuscript was therefore inappropriate and has been removed. Our intended meaning was that this sector samples CO₂ after downwind transport and partial dilution from the industrial area. It is not regarded as representative of a fully mixed boundary-layer background. We have revised the text accordingly to avoid implying

complete atmospheric mixing.

17. Lines 161-163. “Emission source data were obtained from the park’s emission area within the Lidar’s detection range, i.e., the industrial production zone with an azimuth of 315° – 4° and a radial distance of 500–1500 m.” What is “emission source data”?

We thank the reviewer for asking for this clarification. The term “emission source data” in the original manuscript was imprecise. It refers to the lidar-observed CO₂ concentration averaged over the defined industrial source region, rather than to an independent emission inventory or source-strength estimate. Within the lidar detection domain, we defined the industrial production area at azimuths of 315° – 4° and radial distances of 500–1500 m as the source-region observation sector (Fig. 1b). For each lidar volume scan, the CO₂ concentrations within this sector were spatially averaged to generate a time series representing the near-source CO₂ concentration. This time series is subsequently compared with the CO₂ concentration measured over the forested hillside receptor region to examine changes associated with transport and dispersion from the industrial area. We have replaced the ambiguous term “emission source data” with “near-source CO₂ concentration” and revised the corresponding text accordingly.

18. Lines 165-167. “During the entire study period, the average CO₂ concentration at the emission source was 578 ppm, while the average concentration at the downwind site was 487 ppm, with a difference of 91 ppm (Fig. 2a). Meanwhile, the mean wind speed was 2.7 m s^{-1} .” I don’t think averages over several weeks with variable meteorology and emissions are very useful.

We agree with the reviewer that a single average over the entire observation period, during which meteorological conditions and emissions varied substantially, has limited value for interpreting CO₂ transport and dispersion. In the revised manuscript, the full-period mean concentrations are retained only as brief descriptive background and are no longer used as the primary basis for physical interpretation. We have instead added grouped statistics for daytime and nighttime conditions and for low- and high-wind-speed periods to better characterize the dependence of the observed CO₂ concentrations on meteorological conditions. The corresponding discussion has been revised accordingly (Lines 267–274).

19. Lines 168-169. “the daily maximum concentration of the emission source exceeded 900 ppm on each day.” What is “the emission source”? Is this a location that the authors claim is the primary source of CO₂ emissions?

We thank the reviewer for the careful reading. We agree that the term “emission source” was ambiguous. It does not refer to a single location identified as the primary CO₂ emission source. Rather, it denotes the predefined industrial production region within the lidar observation domain, covering azimuths of 315° – 4° and radial distances of 500–1500 m, as shown in Fig. 1b. The reported concentration therefore refers to the lidar-observed CO₂ concentration within this near-source region. We have revised the terminology accordingly to avoid this ambiguity.

20. Given the complex spatial setting, a view of the “emission source” and the “downwind site” in a horizontal / vertical cross section view along the rough direction of the lidar beam might help to explain how the authors have chosen to subselect their observations.

We thank the reviewer for the thorough and constructive suggestion. In the revised manuscript, we have added subfigure (b) to Figure 1, which provides a horizontal view of the lidar scanning range and clearly delineates the near-source region and the forested receptor region. We have also added subfigure (c), which shows a schematic vertical view of the lidar beam geometry at a fixed elevation angle of 10° . These additions are intended to clarify how the two observation sectors were selected under the complex terrain setting. Corresponding descriptions of the selection method have been added to the relevant part of the text (Lines 255–259):

“The industrial production zone within the lidar detection coverage, defined by an azimuth of 315° – 4° and a radial distance of 500–1500 m, was designated as the emission source region (Fig. 1b). The emission source data used in this study were derived from spatially averaged CO_2 concentrations over the source region, calculated from each full lidar volume scan. For quality control, only lidar measurements with a carrier-to-noise ratio (CNR) greater than -27 dB within the scanning domain were retained for subsequent statistical analysis.”

21. Figure 2. What is the time and space resolution of the data shown in Figure 2(a)? Is that the data resolution used to compute the statistics in Figure 2(b), and shown in the data points of Figure 2(c)?

We thank the reviewer for these precise questions. We agree that the temporal and spatial sampling used in the different panels of Figure 2 (Figure 3 in the revise manuscript) was not sufficiently described in the original manuscript, and we have now clarified this information.

(1) Resolution of the data shown in Fig. 3(a): The lidar performs scanning measurements at a fixed elevation angle of 10° , over azimuths of 306° – 16° with a 2° azimuth step. The native CO_2 retrieval has a range resolution of 120 m, and the acquisition time at each CO_2 azimuth is 1 min. A complete scan cycle requires approximately 52 min, including the alternating CO_2 concentration and horizontal wind-field measurements. For the analysis shown in Fig. 3(a), the CO_2 retrievals were spatially averaged over the predefined near-source and forested receptor regions rather than using a single 120-m range gate. After quality control, these regionally averaged time series were interpolated onto a common 1-h time grid using Gaussian-weighted interpolation.

(2) Consistency among Fig. 3(a–c): Yes. The statistics shown in Fig. 3(b) and the paired data points shown in Fig. 3(c) were calculated from the same quality-controlled, spatially averaged dataset used in Fig. 3(a). After processing, the near-source and forested receptor-region datasets contained 1088 and 1023 valid samples, respectively. The processed time series and the corresponding revised comparison are shown in Fig. 3 below. We have now stated explicitly in the revised manuscript that Fig. 3(a–c) are based on the same processed dataset.

(3) Revision of the figure: We have also reorganized the original Figure 3 to improve the presentation of the wind-field characteristics. The wind-rose panel in the original Fig. 3(c) has been replaced by a diurnal frequency distribution of wind direction in the near-source region. In addition, a new Figure 4 has been added to present wind-rose diagrams at different altitude layers. These revised wind-direction analyses are shown in Fig. 4 below and provide a clearer representation of both the temporal and vertical variability of the wind field.

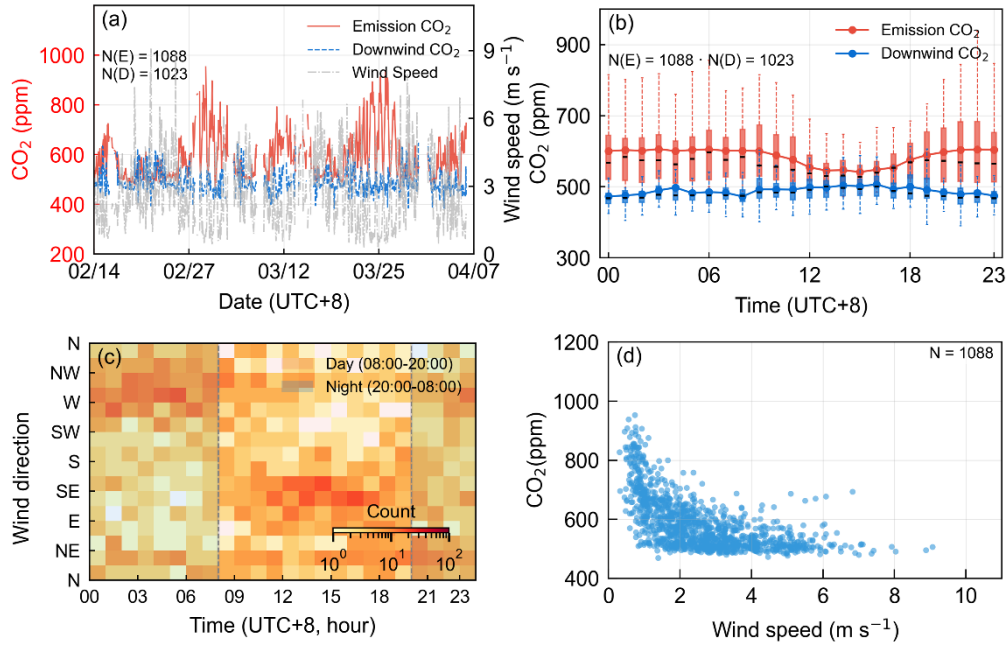


Figure 3. (a) Time series of CO₂ concentrations at the emission source and downwind point, along with wind speed measurements, from 14 February to 6 April 2025, as detected by the Lidar; (b) diurnal variations in CO₂ concentrations at the emission source and downwind point; (c) diurnal frequency statistics of wind direction at the emission source; (d) relationship between wind speed and CO₂ concentration within the industrial park area.

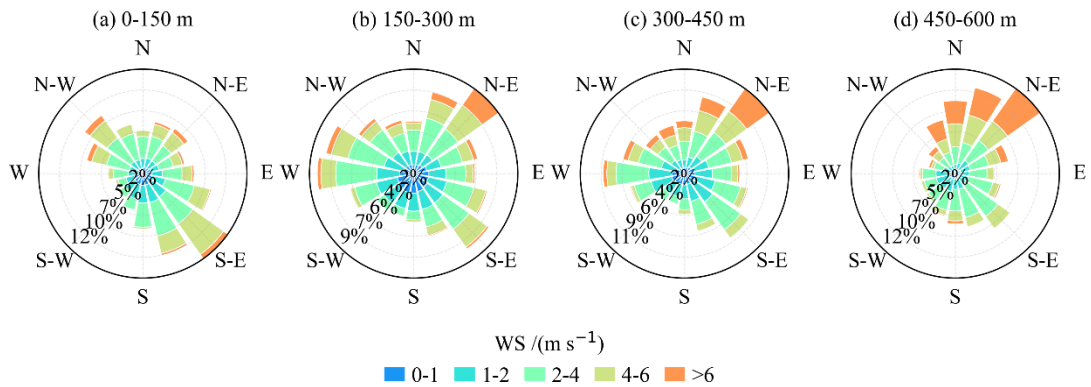


Figure 4. Wind-rose plots at different altitudes derived from Lidar measurements. (a) 0–150 m altitude; (b) 150–300 m altitude; (c) 300–450 m altitude; (d) 450–600 m altitude.

22. Lines 175-176. “with afternoon concentrations slightly higher than nighttime levels. This difference is closely related to changes in land-sea breeze circulation and turbulence.” This might be true, but it is not at all evident from the observations shown. Please either show this relationship or revise this statement.

We thank the reviewer for this important comment. We agree that the original statement was too strong because the relationship between the afternoon CO₂ enhancement and the local wind field was not sufficiently demonstrated. To address this issue, we added an analysis of the diurnal variations in wind speed and wind direction in the near-source region. The results show that during the afternoon the wind speed is generally higher and the prevailing wind is southeasterly. Because the forested receptor region lies northwest of the industrial source region, these southeasterly winds

are consistent with transport from the industrial area toward the receptor region.

The revised Figure 3(c) now shows the diurnal frequency distribution of wind direction in the near-source region, and the new Figure 4 presents wind-rose statistics at different altitude layers. Together with the daytime/nighttime CO₂ statistics, these results support the interpretation that the slightly higher afternoon CO₂ concentrations observed in the receptor region are associated with more effective southeast-to-northwest transport under stronger daytime winds. We have therefore revised the original statement to avoid implying a directly demonstrated causal relationship with land–sea-breeze circulation and turbulence.

23. Lines 178-183. “Meanwhile, as a semi-enclosed bay, Luoyuan Bay exhibits a prominent land-sea breeze effect: prevailing southeasterly winds transport high concentration of CO₂ from industrial emissions to the downwind site in the north-west. However, the surrounding mountainous terrain impedes horizontal dispersion, forcing part of the CO₂ to lift vertically, which further promotes pollutant dilution. In addition, the photosynthesis of forests enters an active period during the daytime, and the absorption of CO₂ partially offsets the increment of industrial emissions, preventing a sharp rise in concentrations at the downwind site.” This all might be true, but nothing in this manuscript demonstrates these hypothesized relationships. Speculative reasoning without a basis in observation might be suitable for the discussion, but these are not results.

We thank the reviewer for this important comment. We agree that several statements in the original Results section went beyond what could be directly supported by the observations. We have therefore revised this section to distinguish clearly between observational results and possible physical interpretations.

Specifically, the wind-rose diagram in the original Fig. 2(c) has been replaced by a diurnal frequency distribution of wind direction in the near-source region (Fig. 3), which directly shows the prevailing wind-direction pattern during the study period. We have also added a new Fig. 4, which presents wind-rose statistics for four altitude layers (0–150 m, 150–300 m, 300–450 m, and 450–600 m). These observations show clear vertical differences in wind speed and direction, with the strongest low-level variation occurring below approximately 300 m. The new analysis provides observational context for discussing the possible influence of the surrounding terrain on the local wind field, but we no longer present this influence as a directly demonstrated result.

In addition, statements lacking direct observational support have been removed from the Results section. For example, the previous claim that “the photosynthesis of forests enters an active period during the daytime, and the absorption of CO₂ partially offsets the increment of industrial emissions” has been deleted. Other mechanism-based interpretations have either been removed or, where retained, have been reformulated as possible explanations in the Discussion rather than as observational results.

24. Figure 2(d) is nice to see, but this relationship has been shown many, many times for a wide variety of pollutants.

We agree with the reviewer and have reframed Figure 2(d) as a baseline site-specific characterisation rather than a novel finding. The CO₂–wind-speed relationship is retained only to provide meteorological context for the subsequent transport analysis. The novelty of the manuscript lies instead in the high-resolution lidar-resolved plume structure and the quantitative WRF–lidar evaluation of three emission-inventory configurations, rather than in the CO₂–wind-speed

correlation per se.

25. Table 2 lacks units, it does not describe the domain studied and it does not describe the resolution of the data. The sources of data are also not clear. This is not a useful table. This certainly does not present “validation” of the WRF simulation. Figure 3 is more useful.

We thank the reviewer for this careful and rigorous comment. We agree that the original Table 2 did not provide sufficient information on units, spatial domain, data resolution, and data sources, and that it should not have been presented as a validation of the WRF simulation. We have therefore removed Table 2. The relevant quantitative metrics are now displayed directly in the revised Figure 5 (formerly Figure 3), where the observation–model comparison can be interpreted together with the corresponding data distributions. In addition, we have separated the evaluation of the meteorological fields and the CO₂ simulations into two dedicated subsections (Sections 3.2.1 and 3.2.2), respectively. Throughout the revised manuscript, we now use “evaluation” or “observation–model comparison” rather than “validation” where appropriate.

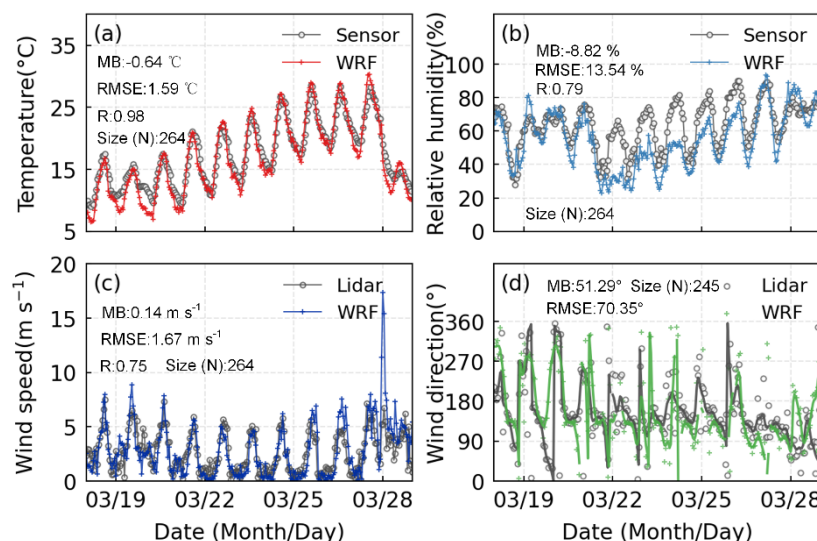


Figure 5. Evaluation of meteorological field simulations from WRF GHG; (a) Temperature; (b) Relative humidity; (c) Wind speed at the emission source; (d) Wind direction at the emission source.

26. Lines 200-202. “The WRF-GHG simulation results were compared and validated against Lidar detection data and sensor-collected data, as shown in Table 2. Figure 3d presents a comparison between the WRF simulated CO₂ concentrations and the Lidar measurements at the park’s emission source.” Sentences like these do not belong in the text of a manuscript. These should be in the captions for the table and figure. Please remove these from the text. Start each paragraph of the results with a topic sentence that presents the primary result that you wish to discuss in the following paragraph. Please follow this approach throughout the manuscript.

We thank the reviewer for this careful and rigorous comment. We agree that sentences serving only to direct the reader to a table or figure are better placed in the corresponding captions rather than in the main Results text. We have therefore removed such guiding phrases, including statements such as “as shown in Table 2” and “Fig. 3d presents...”, from the Results section and incorporated the relevant descriptive information into the table and figure captions where appropriate. In addition, we have systematically revised the Results section so that each paragraph now begins with a topic

sentence stating the primary result discussed in that paragraph, followed by the supporting quantitative evidence and interpretation.

27. Lines 202-203. “While the two datasets exhibit good consistency in temporal variability, substantial discrepancies remain in their absolute values.” I’m sorry, these two CO₂ values look utterly unrelated. There is not at all “good consistency in temporal variability.”

We thank the reviewer for this important comment. We agree that the original statement that the WRF-GHG simulation showed “good consistency in temporal variability” with the lidar observations was not supported by the original comparison and was therefore overly optimistic. We have removed this wording and substantially revised both the emission treatment and the observation–model evaluation.

As described in Sect. 2.3.3, three anthropogenic-emission configurations are now considered. The raw EDGAR inventory is used as the baseline. EDGAR_S applies industrial-source spatial redistribution and land–sea mixed-grid correction, while EDGAR_P further introduces the SNAP3 vertical profile for industrial emissions. A lidar-derived diurnal scaling factor is also applied to represent intra-day emission variability.

In Sect. 3.2.2, the three simulations are quantitatively compared with the lidar observations over the 80–300 m layer corresponding to the lidar-observed near-source region. The corresponding observation–model comparisons are shown in Figures 6 and 7. Over 18–28 March 2025 (N = 264 hourly samples), the correlation coefficient increases from 0.60 for raw EDGAR to 0.67 for EDGAR_S and 0.69 for EDGAR_P, while the mean bias changes from –165.4 ppm to –115.9 ppm and –122.1 ppm, respectively. These results show that the spatial correction substantially reduces the negative concentration bias, whereas the additional vertical allocation provides a modest improvement in temporal correlation but does not uniformly reduce the mean bias.

The daytime/nighttime analysis presented in Figure 7 further illustrates this distinction. During daytime, EDGAR_P gives the highest correlation ($R = 0.72$), whereas EDGAR_S gives the smallest mean bias (–101.5 ppm). At night, both EDGAR_S (–130.4 ppm) and EDGAR_P (–140.2 ppm) substantially reduce the negative bias relative to raw EDGAR (–201.4 ppm).

Because the diurnal scaling factor is derived from the lidar observations themselves, we interpret these experiments as an observation-constrained sensitivity evaluation rather than as a fully independent validation. Accordingly, the revised manuscript no longer claims “good consistency” between the original WRF-GHG simulation and the lidar observations.

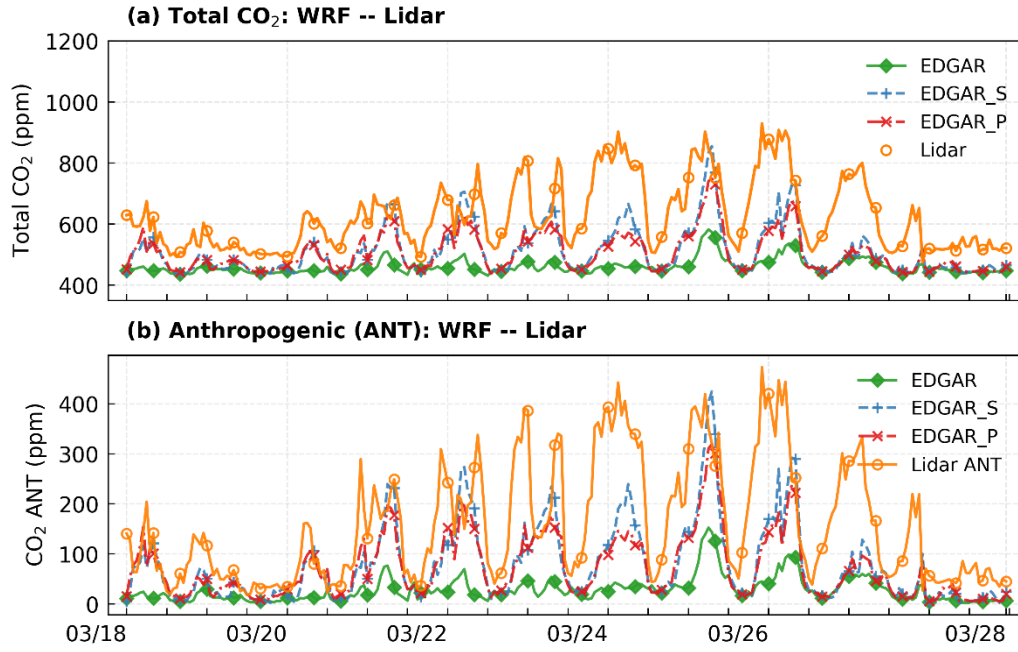


Figure 6. Comparison between in-source Lidar observations and simulations from three emission inventories, 18–28 March 2025. (a) Simulated CO₂ concentrations from three inventories vs. Lidar measurements at the industrial emission source; (b) Anthropogenic CO₂ component (CO₂_ant) from model outputs vs. industry-induced CO₂ concentration enhancements derived from Lidar data.

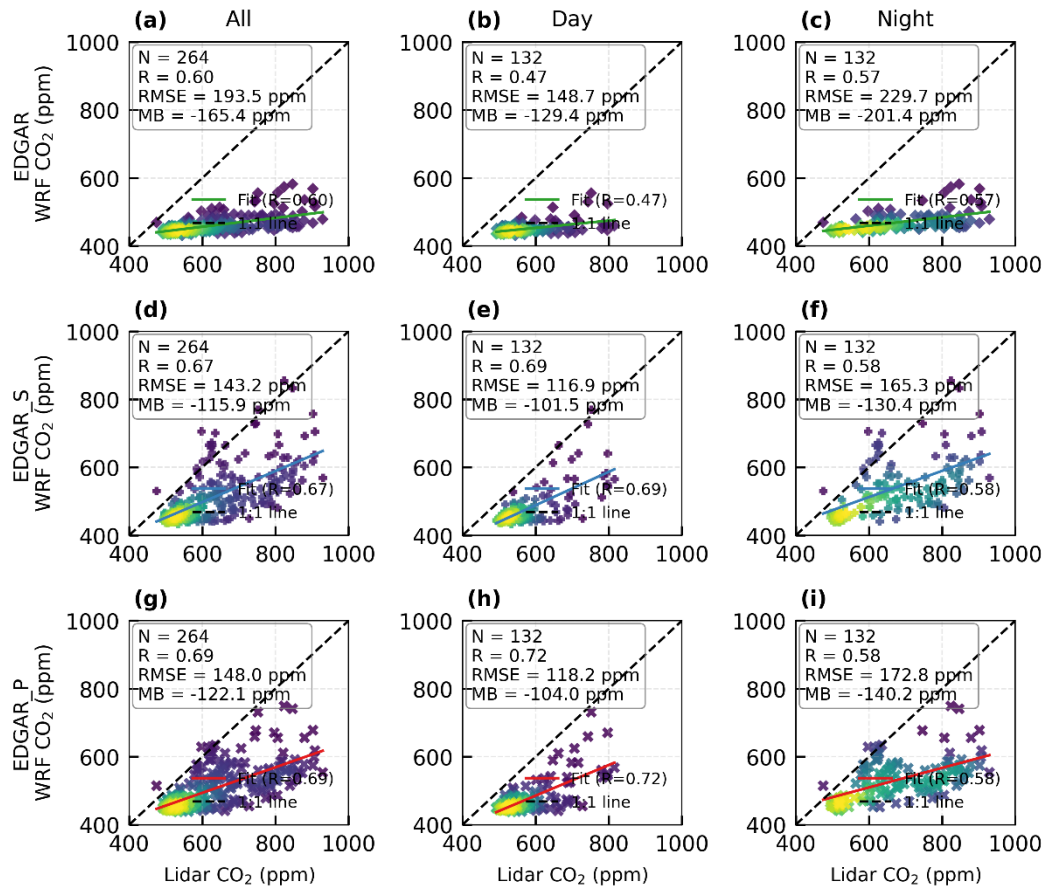


Figure 7. Scatter density plots of in-source Lidar CO₂ against simulations from three emission inventories. (a–c) Raw EDGAR; (d–f) EDGAR_S; (g–i) EDGAR_P, sorted by all hours, daytime, nighttime.

28. Lines 205-206. “The EDGAR_2024_GHG dataset used in the simulation has a latitudinal and longitudinal resolution of only 0.1° and a temporal resolution of monthly data.” This should be described in the methods.

We thank the reviewer for this careful comment. We agree that the spatial and temporal specifications of the EDGAR inventory are more appropriately described in the Methods section. We have therefore moved the EDGAR inventory specifications (0.1° horizontal resolution and monthly temporal resolution) from the Results to Section 2.3.3 (“Anthropogenic emission settings”), where the three emission configurations and their corresponding processing procedures are now described together.

29. Lines 203-205. “Additionally, the northern Luoyuan Bay where the park is situated is surrounded by mountains on three sides and adjacent to the sea on one side; this unique terrain poses a significant challenge for the terrain data inherently used by WRF.” As noted previously, it would be very helpful to see a distance / altitude cross section of the data locations and the model grids and altitude levels that are being evaluated in this study. It would help a great deal to show the actual terrain and the terrain data used in WRF. Finally, a visualization of the CO₂ fluxes from EDGAR would be exceptionally helpful. Without this information this text is not very interpretable. What, for example, is “the terrain data inherently used by WRF” and why does this cause a problem?

We thank the reviewer for this important and detailed comment. We agree that the original statement concerning “the terrain data inherently used by WRF” was vague and insufficiently supported by the analyses presented in the manuscript. We have therefore removed this wording and substantially weakened the associated interpretation.

To clarify the spatial relationship among the lidar observations, the model grids, and the surrounding terrain, the revised Figure 1 now includes the lidar observation sectors, the innermost WRF domain, a satellite map of Luoyuan Bay, and horizontal/vertical schematics of the lidar sampling geometry. These additions make the relative locations of the near-source region, receptor region, surrounding mountains, and model grid more explicit.

In addition, we have added Figure R1, which shows the land-cover classification and land–sea mask derived from the MODIS 30-arc-second dataset used in the WRF configuration, together with the actual coastline and the location of the industrial facility. This comparison reveals that part of the industrial area containing the lidar site is represented as Evergreen Needleleaf Forest in the model land-cover dataset, and that discrepancies also exist between the WRF land–sea mask and the actual coastline. We emphasise that these are land-cover and land–sea-mask representation issues rather than terrain-elevation errors. They may affect the surface properties and VPRM biospheric fluxes, but their quantitative impact has not been isolated in the present study.

For the anthropogenic emissions, the revised manuscript also adds visualisations of the raw EDGAR and spatially corrected EDGAR_S CO₂ emission fields. These figures show the coarse-grid structure of the original inventory, the removal of spurious marine emissions in mixed land–sea grids, and the redistribution of industrial emissions toward the actual plant areas.

In the present revision, we consider the anthropogenic emission inventory to be the dominant source of the CO₂ simulation bias and therefore focus the quantitative sensitivity analysis on emission-inventory refinement. We have not attempted to quantitatively attribute the remaining bias to terrain representation, because this would require dedicated experiments involving terrain resolution and model configuration. Accordingly, the possible effects of terrain representation, land-

cover classification, and VPRM parameters are now identified as directions for future work rather than being presented as demonstrated causes of the model–observation discrepancy.

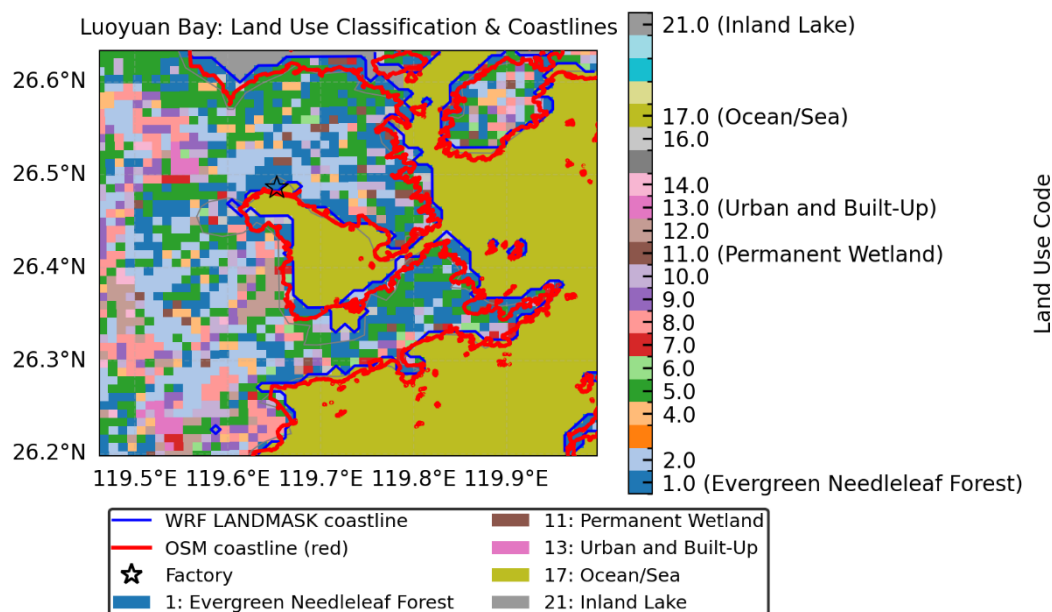


Figure R1. The land cover classification and land-sea mask derived from the MODIS 30-second data over the Luoyuan Bay area.

30. Lines 210-213. Correlation coefficients spanning a few weeks show that WRF simulates the daily cycle and weather patterns reasonably well. That is not surprising.

We thank the reviewer for this careful and rigorous comment. We agree that the ability of WRF to reproduce the basic diurnal cycle and short-term meteorological variability over a multi-week period is expected and should not be presented as a novel result. We have therefore reframed this analysis as model-evaluation context rather than as a substantive finding. Its purpose is to assess whether the simulated meteorological driving fields are sufficiently representative for the subsequent CO₂ transport analysis. The revised manuscript now foregrounds the emission-inventory refinement and the quantitative lidar–model comparison as the main contributions of the study.

31. Lines 214-216. “Among these, the correlation coefficient (R) for 10m wind speed is 0.7 with RMSE of 1.45 m s⁻¹, indicating that WRF underestimates wind speed to a certain extent;” RMSE and R values say nothing about biases. These data do not indicate that WRF underestimates wind speed.

We thank the reviewer for this careful and rigorous comment. We agree that the correlation coefficient and RMSE quantify the correspondence and overall error magnitude between the simulated and observed wind speeds, but do not indicate the direction of the bias. The original statement that WRF underestimated wind speed based on these two metrics was therefore incorrect and has been removed.

In the revised simulation, after excluding the spin-up period, the 10 m wind speed has a correlation coefficient of and an RMSE of 1.67 m s⁻¹, indicating that the simulation captures the main temporal variability of the observed wind speed. We have additionally calculated the mean bias (MB), which is +0.14 m s⁻¹. This positive MB indicates a slight overestimation of wind speed in the present simulation. We emphasise that this bias is specific to the current model configuration

and study period and should not be interpreted as a systematic bias inherent to the WRF model. The corresponding statistics are shown in Fig. 5.

32. Figure 4. The comparisons between the CO₂ lidar and the WRF CO₂ simulations are not useful. The observations are on an entirely different spatial scale than the model and represent a complex set of distances above ground. The large area WRF simulations are the surface layer where, especially in stable conditions but truly at any time of day close to a strong source, will be quite different than concentrations aloft. The authors have already noted that EDGAR has roughly 10km resolution hence cannot possibly represent emissions from this microscale domain (lidar) effectively. And if I understand the WRF maps, the model terrain puts the lidar in the water! I do not think anything useful can be learned from this simulation.

We thank the reviewer for this important and rigorous comment. We agree that the original comparison did not sufficiently account for the differences in spatial scale and sampling height between the lidar observations and the WRF-GHG simulations. We have therefore revised both the emission treatment and the observation–model comparison.

First, the WRF-GHG CO₂ concentrations used for comparison are no longer taken directly from the model surface layer. Based on the lidar observation heights in the near-source region, the simulated CO₂ concentrations are now averaged over the 80–300 m layer to provide a more appropriate comparison with the lidar observations.

Second, to address the coarse representation of local industrial emissions in the original EDGAR inventory, we quantitatively evaluated three emission configurations: raw EDGAR, EDGAR_S, and EDGAR_P (Sect. 3.2.2). The EDGAR_P simulation results were then used to regenerate Figures 8–10.

Third, the apparent placement of the industrial park over water in the original figures was partly caused by the insufficient spatial resolution of the default Cartopy coastline used for plotting. We have therefore updated the coastline using a higher-resolution OSM shapefile and overlaid hill-shaded terrain from the ASTER GDEM v3 30 m dataset.

With these revisions, Figures 8–10 now provide a more appropriate comparison between the lidar observations and the corresponding WRF-GHG simulations over Luoyuan Bay.

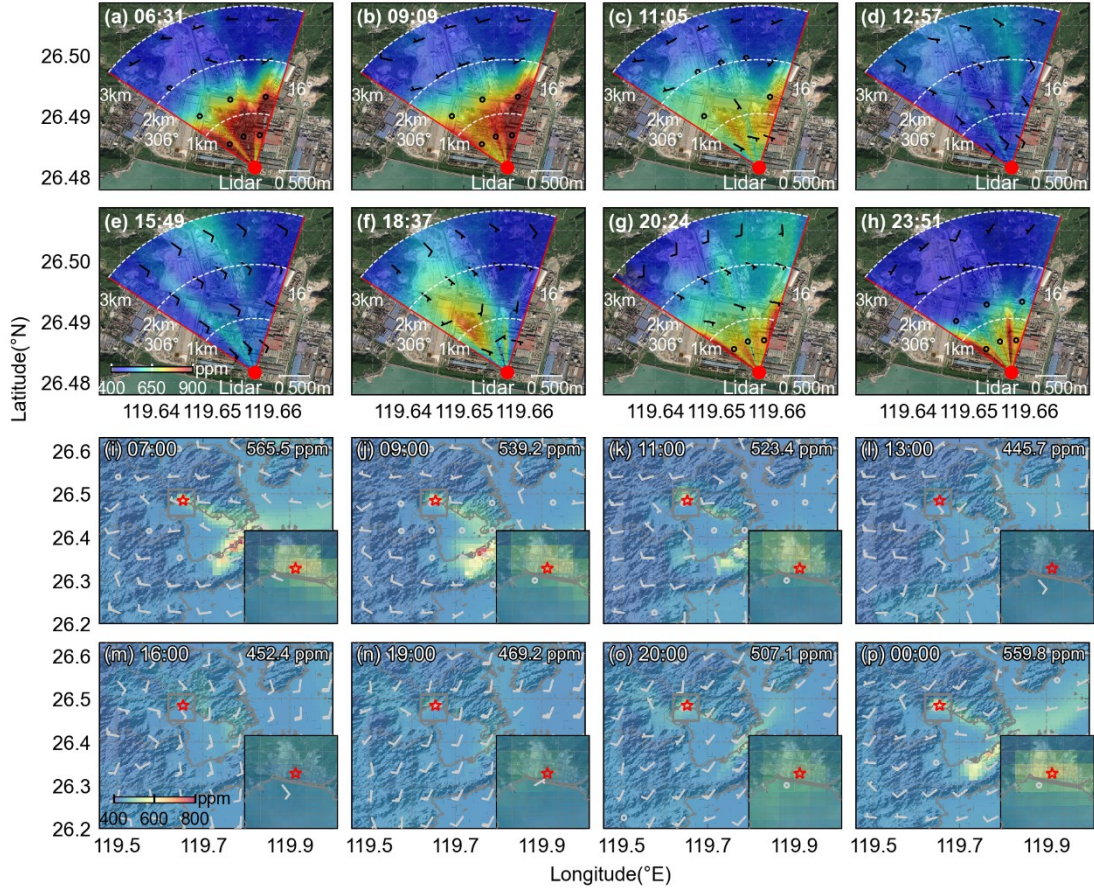


Figure 8. Partial detection results of the CO₂ Lidar and simulation results of WRF-GHG at corresponding times on 24 March 2025 (LT; UTC+8); (a)–(h) are the CO₂ Lidar observation results (panel time denotes the start of a about 52-min volume scan); (i)–(p) are the WRF-GHG simulated CO₂ concentrations (layer-averaged over 80–300 m) over Luoyuan Bay, with the magnified industrial park shown in the lower-right corner. The main panels overlay the simulated field on a hill-shaded DEM, and the inset overlays the Microsoft Bing satellite imagery of the park. The value in the upper-right of each panel (i)–(p) denotes the simulated CO₂ concentration (ppm) at the source region (80–300 m layer mean). Wind speed is uniformly represented by short dashes and small triangles perpendicular to the right side of the end of the wind direction bar (in the Northern Hemisphere), where long dashes represent 4 m s⁻¹, short dashes represent 2 m s⁻¹, triangles represent 20 m s⁻¹, and wind speeds less than 2 m s⁻¹ are represented as open circles. Map source: Microsoft Bing Maps 2026.

33. Figure 4. What is the time and space resolution of the lidar data displayed. This is never mentioned but is quite important. If these are snapshots in time, why have these been selected?

We thank the reviewer for this important comment. We agree that the temporal and spatial sampling of the lidar data, as well as the rationale for selecting the displayed scans, should be stated explicitly.

The lidar operates at a fixed elevation angle of 10°. For each CO₂ scan, the azimuth is stepped clockwise from 306° to 16° at 2° intervals, with 1 min of CO₂ acquisition at each azimuth. The native CO₂ retrieval has a range resolution of 120 m. After the CO₂ scan, the lidar performs a continuous wind-field sweep from 16° back to 306°. A complete CO₂/wind measurement cycle requires approximately 52 min. The time labelled in each lidar panel denotes the start time of the corresponding scan. Because the CO₂ field is assembled sequentially from measurements acquired

at different azimuths, each panel represents one scanning period rather than an instantaneous snapshot.

The displayed scans were selected to provide representative coverage of the diurnal evolution of the CO₂ field and to allow comparison with WRF-GHG simulations at corresponding times. They were not selected according to the magnitude of the observed CO₂ concentration. We have now added these details to Sect. 2.2, Sect. 3.3, and the figure caption.

34. Lines 243-246. “Only low-intensity emissions were detected at the sources, and CO₂ was transported from the southeast to the northwest. At 18:37 LT, the sea breeze weakened, and the wind field in the park became turbulent (wind speed: 2–4 m s⁻¹). The emission intensity increased at the source located at 1 km radial distance with 330°–350° azimuth, with CO₂ concentration reaching around 750 ppm and dispersing to the far field.” One minor concern. Winds can be turbulent regardless of the mean wind speed. Mean winds becoming lower(?) do not mean that atmospheric turbulence has increased. The opposite is often true. A more significant concern: The text appears to associate the strength of CO₂ sources with the observed CO₂ concentrations. Concentrations are the result of both the emissions strength and ventilation. Source strength cannot be judged solely via atmospheric concentrations.

We thank the reviewer for these important comments. We have revised the relevant text accordingly. Specifically, we no longer associate changes in mean wind speed with turbulence intensity, and we recognise that source strength cannot be inferred solely from atmospheric CO₂ concentrations. The observed CO₂ concentration reflects the combined effects of emissions and atmospheric ventilation; therefore, it is used here only to characterise the observed concentration enhancement rather than to infer changes in emission strength. The original sentences have been revised in Lines 421–427 as follows:

“A CO₂ enhancement associated with the industrial source region was detected and transported from southeast to northwest under southeasterly winds. At 18:37 LT, the wind speed in the park was 2–4 m s⁻¹. Within the radial distance of 0–1 km, the wind was easterly, whereas within 1–2 km it turned southwesterly. Within the near-source region at a radial distance of approximately 1 km and azimuths of 330°–350°, the CO₂ concentration increased to about 750 ppm, with the enhanced CO₂ extending toward the far field. In contrast, the region 2–3 km away at azimuths of 0°–16° showed no comparable CO₂ enhancement associated with this plume, and the concentration remained at approximately 450 ppm.”

Quantitative estimation of the emission rate would require a dedicated emission-inversion approach and is identified in the outlook section as a direction for future work.

35. Lines 252-253. “Combined with mountain obstruction and a decrease in Boundary Layer Height (BLH),” I have not seen any measurements of BLH. these would be useful and could be obtained by a lidar. If this is an assumption, that should be clearly stated.

We thank the reviewer for pointing this out. No direct boundary-layer height (BLH) measurements were available during the campaign. The decrease in BLH referred to in the original text was an inference rather than a directly observed quantity. We have therefore removed these statements from the revised discussion of Figures 8–10.

In Figure 11, we now explicitly clarify that both the BLH and the inversion layer are WRF-diagnosed quantities. Because the lidar operated in a fixed-elevation (10°) PPI scanning mode

during the campaign, it did not directly retrieve the BLH. These WRF-diagnosed quantities are therefore used only to provide meteorological context for interpreting the simulated vertical CO₂ distribution, rather than being treated as observational measurements. We also clarify that these diagnostics are specific to the current YSU configuration and have not been independently validated in this study.

36. I am afraid that I do not see much scientific content in Figure 4 and the associated discussion, save for a demonstration of the lidar technology. I have the same concerns for Figures 5 and 6.

We thank the reviewer for the careful and rigorous review. As mentioned in our response to Comment 32 above, we have regenerated Figures 8–10 (formerly Figures 4–6) using the EDGAR_P simulation results, which incorporate industrial-source spatial redistribution, land–sea mixed-grid correction, and vertical emission allocation. The revised figures now provide a more direct spatial comparison between the lidar observations and the WRF-GHG simulations. Accordingly, their purpose is no longer limited to demonstrating the lidar measurements, but also to evaluate how the refined emission treatment affects the simulated spatial and temporal CO₂ distributions within the bay.

37. Lines 299–304. “Joint analysis of the high concentration period (24–26 March) reveals that both the Lidar and WRF model captured the core dynamic of CO₂: “nighttime accumulation and daytime dispersion”. The dispersion and accumulation processes are jointly influenced by wind fields (sea-land breeze transition, wind speed) and terrain (mountain obstruction). The Lidar accurately identifies local emission plumes and small scale dispersion details, while the model reproduces the overall regional variation trend. These two approaches complement each other to verify the complex dispersion mechanism of CO₂ in the semi-enclosed bay.” The day/night cycle of CO₂ accumulation close to the surface where sources are located is not a significant or publishable finding. The variations in wind related to the bay location are interesting but it is not clear that the WRF simulation reproduces these patterns. I’m afraid that I do not agree that these are complementary approaches for this set of observations.

We thank the reviewer for these critical and constructive comments. We fully agree that the near-surface diurnal cycle of CO₂ in the source region is a well-established phenomenon and does not constitute a novel finding. In the revised manuscript, this feature is therefore described only as a site-specific observational characteristic rather than as a key contribution.

We have instead refocused the analysis on three aspects: (1) the quantitative evaluation of the three emission-inventory configurations (raw EDGAR, EDGAR_S, and EDGAR_P); (2) the assessment of how the emission-inventory refinements affect the model–observation discrepancy; and (3) a more direct comparison between WRF-GHG and lidar observations over comparable spatial and vertical sampling ranges. In particular, the WRF-GHG CO₂ fields are now extracted over the 80–300 m layer corresponding approximately to the lidar sampling heights in the near-source region.

Accordingly, the original Lines 299–304 have been substantially revised. The revised text no longer states that the lidar and WRF-GHG “complement each other to verify” the CO₂ dispersion mechanism. Instead, the lidar observations are used to characterise the local high-resolution CO₂ variability, while the WRF-GHG simulations are evaluated against these observations to examine how the different emission treatments influence the simulated temporal and spatial CO₂ distributions.

The remaining near-source underestimate is discussed only as a configuration-specific discrepancy that may partly reflect the use of a generic vertical emission profile.

Thus, the revised discussion no longer treats the day–night accumulation–dispersion cycle itself as a novel result. The main contribution is the quantitative, observation-constrained evaluation of the emission-inventory treatments and the resulting improvement in the lidar–WRF-GHG comparison.

38. Lines 329-334 should be in the figure caption, not the text.

We thank the reviewer for the careful reading. Since the relevant details were already provided in the original figure caption, we have removed the redundant description from the main text while retaining the information in the caption.

39. Figure 7(f). What is purpose of “the inversion layer”?

We thank the reviewer for the careful reading. The inversion layer shown in Figure 11 is a WRF-diagnosed feature derived from the simulated thermal structure. Its purpose is to provide additional meteorological context for interpreting the simulated vertical CO₂ distribution under different stability conditions. In the revised manuscript, it is considered together with the WRF-diagnosed boundary-layer height and vertical velocity, but is not treated as an independently validated quantity or as evidence of a new boundary-layer mechanism. We have clarified this role in the revised text and explicitly noted that both the inversion layer and BLH are diagnostics from the current YSU configuration. A quantitative evaluation of these diagnostics using different boundary-layer parameterisation schemes is identified as future work.

40. Why does Figure 7 show mostly model output, and mostly at altitude beyond the range of the lidar system? What is the significance of showing these modeled fields? A comparison to the lidar winds and ABL depth estimate with this type of resolution and view might be valuable. It would be helpful to know if WRF was simulating similar BLH mixing and winds in this complex region. The CO₂ fields from the model are not useful.

We thank the reviewer for this helpful suggestion. During the field campaign, the lidar operated in a fixed-elevation (10°) PPI scanning mode, which does not allow a direct retrieval of the boundary-layer height. Therefore, a direct lidar–WRF comparison of BLH cannot be provided from the present measurements.

We have revised Figure 11 accordingly. The simulated fields are now restricted to the 0–1 km altitude range to improve consistency with the lidar sampling range, rather than emphasising model output at altitudes beyond the lidar observations. Within this range, the revised figure includes a comparison of the CO₂ simulations obtained with the three emission-inventory configurations. In addition, the lidar-derived horizontal wind fields at different heights are now directly compared with the corresponding WRF simulations. These revisions place greater emphasis on the observation–model comparison within the altitude range relevant to the lidar measurements.

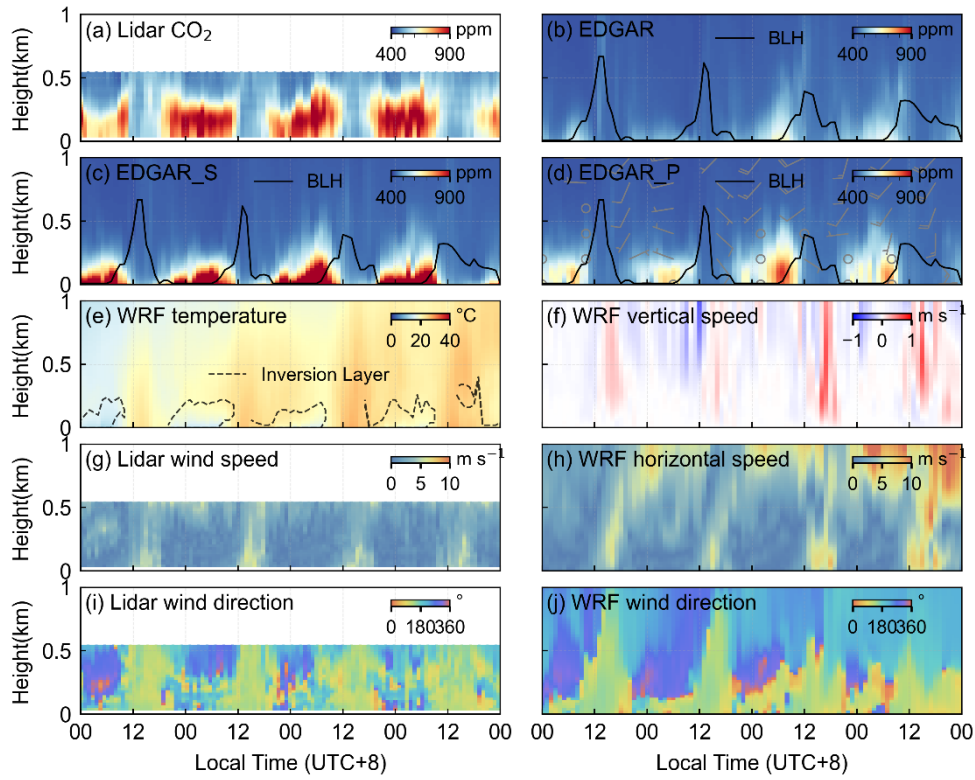


Figure 11. Time-series heatmaps of CO₂ Lidar observations and WRF simulation results from 00:00 local time (LT; UTC+8) on 23 March to 00:00 LT on 27 March 2025; (a) shows the distribution results of CO₂ detected by the CO₂ Lidar converted to vertical height, where the portion above 500 m (white dashed line) is invalid data; (b) presents the WRF-simulated CO₂ concentrations at the park location, with the black solid line representing the model-output BLH; (c)–(e) are the model-output vertical wind speed, horizontal wind speed, and horizontal wind direction, respectively; (f) is the model-output air temperature, where the black dashed line denotes the strongest temperature inversion layer with a temperature gradient exceeding 0.5 °C 100 m⁻¹.

41. Paragraph starting at line 335. This is reasonable but not at all new or unique. And this discussion neglects the impact of wind speed. The accumulation of pollutants close to the ground in stable conditions is not a new finding worthy of publication.

We thank the reviewer for this critical and constructive comment. We fully agree that the physical mechanisms discussed in the original paragraph, including stable boundary-layer conditions, inversion-related suppression of vertical mixing, and near-surface pollutant accumulation, are well-established and do not constitute novel findings by themselves. We also acknowledge that the original discussion did not adequately consider the role of wind speed in regulating pollutant dispersion.

We have therefore refocused the revised discussion on the quantitative evaluation of the three emission-inventory configurations (raw EDGAR, EDGAR_S, and EDGAR_P). The boundary-layer height, inversion layer, and wind conditions are now used only as contextual information for interpreting the differences among the simulated CO₂ distributions, rather than being presented as novel mechanisms or as independently demonstrated causes.

Specifically, the revised vertical cross-section analysis (Fig. 11 and associated text) compares the three emission configurations. EDGAR_P, which distributes industrial emissions over 0–710 m, improves the representation of the lidar-observed vertical CO₂ structure and increases the overall

temporal correlation with the lidar observations. The WRF-diagnosed BLH, inversion layer, and vertical velocity are used only to provide physical context for the differences in the simulated vertical distributions. In addition, the revised discussion explicitly incorporates wind speed when interpreting the observed CO₂ variability, considering the stronger daytime winds and weaker nighttime winds together with the corresponding atmospheric stability conditions.

Thus, the revised discussion no longer presents boundary-layer accumulation or topographic trapping as principal findings. Instead, these established meteorological processes are used only to support the interpretation of how different emission treatments affect the simulated CO₂ distribution. The main contribution of this analysis is therefore the quantitative, observation-constrained evaluation of the emission-inventory treatments rather than a restatement of established boundary-layer theory.

42. Paragraph starting with line 342. Mountainous terrain trapping local pollution sources is also not a new discovery.

We thank the reviewer for this comment. We agree that the general mechanism of topographic trapping of local pollutants is well established and does not by itself constitute a novel finding.

We have therefore refocused the revised analysis. First, the cross-section in Figure 12 has been regenerated using the EDGAR_P emission configuration, which includes spatial redistribution and SNAP3 vertical allocation. Second, the analysis is now restricted to the morning transition from nighttime accumulation to daytime ventilation (02:00–11:00 LT on 25 March), rather than emphasising the generic effect of topographic trapping. The revised discussion focuses on the evolution of the simulated CO₂ distribution and wind field during this transition period and uses the terrain-related flow structure only as context for interpreting the corresponding transport patterns.

Thus, the revised cross-section analysis no longer presents topographic trapping as a stand-alone conclusion or a novel mechanism. Instead, it is used to support the process-oriented interpretation of the EDGAR_P simulation during the morning transition. The main contribution of this analysis is the evaluation of the refined emission representation in relation to the observed and simulated CO₂ evolution, rather than the restatement of well-known topographic or boundary-layer effects.

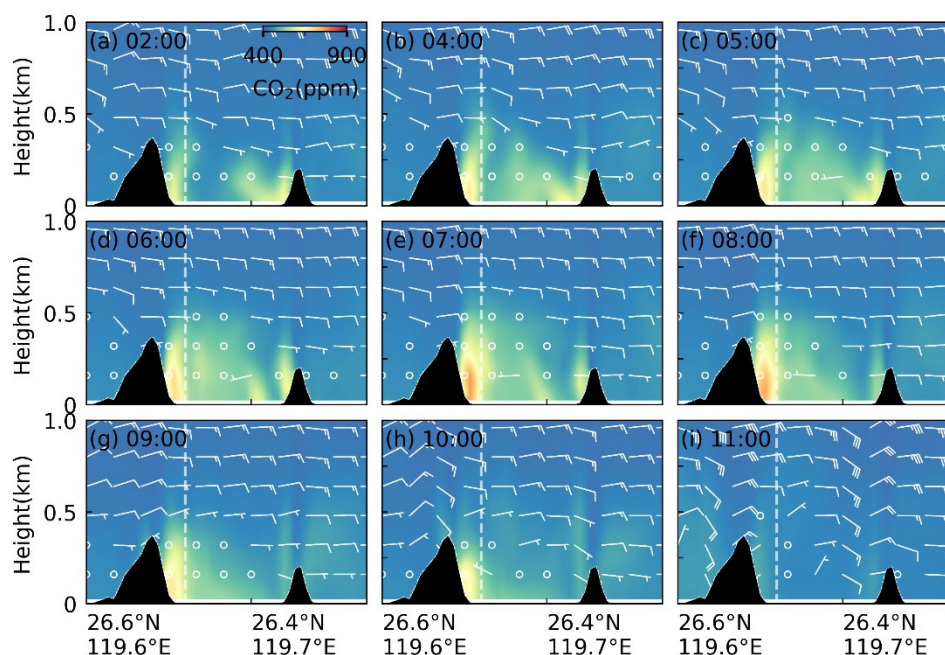


Figure 12. Vertical distributions of CO₂ concentrations and wind fields along the cross-section from (26.65°N, 119.65°E) to (26.18°N, 119.67°E) based on WRF simulation results on 25 March 2025 (LT; UTC+8); shaded areas represent mountainous terrain; the white dashed line denotes the location of the studied industrial park.

43. Lines 353-360. The source for this text is not clear. More importantly, the purpose of this text is not clear. I presume this is a discussion of the simulated meteorology and CO₂ transport, but this is not at all clear. The text speaks of CO₂ emitted by the industrial park accumulating in the boundary layer, but the authors have shown that the simulation does not at all realistically capture the local emissions. A comparison of simulated and observed meteorology and distributions of CO₂ in space and time would be worthwhile, but that is not the content of this text. A simple description of the environment in this bay is interesting but not surprising or new in terms of general meteorological properties.

We thank the reviewer for this comment. As also explained in our response to Comment 42, we agree that the general meteorological characteristics of a semi-enclosed bay are well established and should not be presented as a novel finding. We have therefore clarified that the CO₂ concentrations and wind fields discussed in this section are derived from the WRF-GHG simulation using the EDGAR_P emission configuration. The revised analysis focuses on the morning transition from nighttime accumulation to daytime ventilation and uses the simulated wind fields only to interpret the corresponding transport pathways. In addition, the WRF-GHG results are now quantitatively compared with the lidar observations over the same 80–300 m sampling layer. Thus, this paragraph is now presented as a process-oriented interpretation of the observation–model comparison rather than as a general description of the bay environment.

44. Lines 373-380 are a summary of what has already been presented in the manuscript. This is not suitable text for discussion or conclusions. The rest of the paragraph introduces a new figure, but it is a summary figure rather than conclusions. Conclusions should focus on the message that readers should take away from the new findings. That is not a summary.

We agree with the reviewer that a repetition of the results is not an appropriate conclusion. We have therefore rewritten the Conclusions to focus on the principal messages arising from the quantitative evaluation of the three emission-inventory configurations. The revised Conclusions emphasise that spatial redistribution reduces the overall mean bias from -165.4 ppm for raw EDGAR to -115.9 ppm for EDGAR_S, while the addition of the SNAP3 vertical profile in EDGAR_P increases the overall correlation with the lidar from 0.60 to 0.69 and improves the representation of the vertical CO₂ structure. The remaining near-source underestimate is discussed as a configuration-specific issue that may be related to differences between the generic SNAP3 release profile and the effective release heights of the studied industrial park. Observational findings, their interpretation, and future research directions are now presented separately so that the Conclusions focus on the principal take-home messages rather than repeating the Results.

45. If the manuscript were focused on proving the existence of the day/night pattern shown in figure 9 using both the WRF simulation and the lidar data, and comparing these when appropriate (e.g. atmospheric winds, BLH) this would be a better manuscript. Bringing up Figure 9 in the conclusions isn't appropriate. Further, the manuscript does not bring together lidar and WRF throughout to illustrate the validity of the pattern of events shown in Figure 9.

We agree that the conceptual day–night pattern should be supported by direct observation–model evidence and that a schematic figure should not substitute for such evidence. We have therefore quantitatively compared the lidar observations with the WRF-GHG results using the EDGAR_P emission configuration over comparable spatial and vertical sampling ranges. The revised manuscript compares CO₂ concentrations and synchronous wind fields in both time and space. Because the fixed-elevation lidar scan cannot directly retrieve the boundary-layer height, the WRF-diagnosed BLH is used only as a diagnostic quantity rather than as a directly validated variable.

In addition, we have added a new Figure 13 showing the EDGAR_P-simulated CO₂ and wind fields at the surface and at 100, 200, 300, and 500 m under representative nighttime, morning, and daytime conditions. This figure illustrates the three-dimensional accumulation and dispersion patterns reproduced by WRF and provides additional model-based support for the conceptual day–night pattern shown in the original Figure 9, now Figure 14. Figure 14 is therefore retained only as a schematic synthesis of the preceding observation–model analyses, rather than being introduced in the Conclusions as independent evidence.

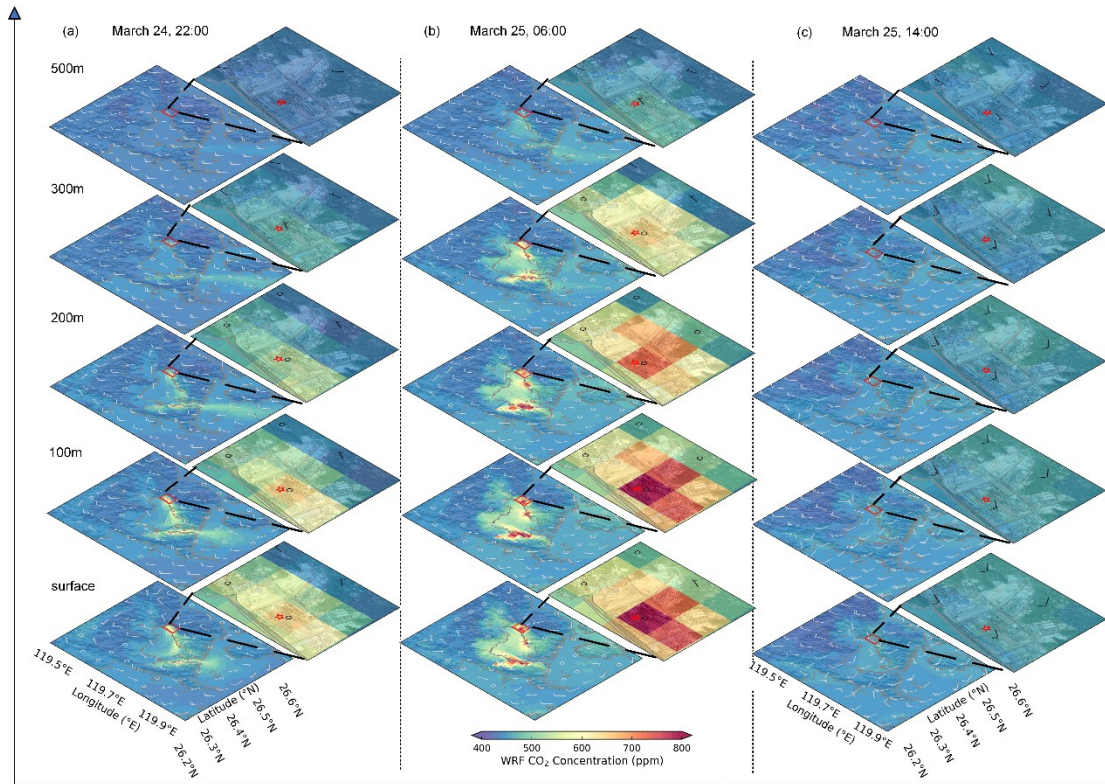


Figure 1. Vertical distributions of EDGAR_P-simulated CO₂ mixing ratios (shading) and wind vectors (arrows) at the surface and at 100, 200, 300, and 500 m above sea level, shown for three representative times: (a) 22:00 LT on 24 March, (b) 06:00 LT on 25 March, and (c) 14:00 LT on 25 March 2025. The main panels overlay the hill-shaded DEM to display the spatial pattern over Luoyuan Bay; the lower-right insets overlay Microsoft Bing satellite imagery and zoom in on the studied industrial park.

46. Line 390. This paragraph could be a conclusion about the limitations of the authors' implementation of WRF-GHG, but none of these issues would be new understanding for the research community. Groups that work with numerical models to simulate atmospheric flow in complex terrain would not expect this configuration to be successful.

We agree that the general challenges of applying mesoscale models in complex terrain are already well known and do not constitute a new scientific conclusion. We have therefore rewritten this part of the Conclusions and removed the broad claim concerning general limitations of WRF-GHG. The revised Conclusions focus on the performance of the specific model and emission configurations evaluated in this study. In particular, they report the quantitative improvements obtained with EDGAR_S and EDGAR_P, while the remaining near-source underestimate is discussed as a configuration-specific issue that may be related to the use of a generic vertical emission profile.

47. Lines 395-399. This discussion of the boundary layer parameterization in WRF is speculation. The authors have not proven anything about the WRF boundary layer parameterization and its response to complex terrain.

We agree that the original discussion of the WRF boundary-layer parameterisation was speculative because no sensitivity experiments involving different boundary-layer schemes were

conducted. We have therefore removed the unsupported causal statement. In the revised manuscript, the boundary-layer height and inversion layer are explicitly treated only as WRF-diagnosed quantities under the current YSU configuration and are used to provide context for the simulated meteorological and CO₂ fields, rather than to evaluate the performance of the boundary-layer parameterisation itself. A quantitative comparison of different boundary-layer parameterisations over complex terrain is now identified as future work rather than presented as a conclusion of this study.

48. The paragraph starting with line 400 is a valid summary of the observational findings of the manuscript. The demonstration of the lidar system is worthwhile. These are results, not conclusions.

We agree that the detailed lidar observational findings are results rather than conclusions. These findings are therefore presented in detail in the Results section, particularly Sects. 3.1 and 3.3. In the rewritten Conclusions, they are summarised only briefly, while greater emphasis is placed on the quantitative evaluation of the three emission-inventory configurations and the observation-constrained assessment of the corresponding WRF-GHG simulations.

49. Line 412 is the start of conclusions that I believe are suitable for this work.

We thank the reviewer for this positive assessment. We used the content beginning at the original Line 412, which the reviewer considered suitable for this work, as the basis for restructuring the Conclusions. We nevertheless revised the wording to incorporate the new quantitative comparison of the three emission inventories, distinguish the demonstrated capabilities of the lidar from the performance of the specific WRF-GHG configuration, and avoid unsupported general claims.

50. Line 412-413. “This study demonstrates the advantages of Lidar in monitoring industrial CO₂ emissions with high spatiotemporal resolution and highlights the limitations of the WRF-GHG model in characterizing complex terrain and refined emission sources.” I disagree with this statement in two ways. First, this manuscript does not monitoring industrial CO₂ emissions. It detects CO₂ from industrial emissions. That is quite different from monitoring CO₂ emissions. If these data were used to quantify CO₂ emissions from the industrial part, this would be a reasonable conclusion. Second, this manuscript does not highlight any general limitations of WRF-GHG in simulating this environment. It does show that the authors have chosen a configuration of WRF-GHG that is not suited to simulating this environment. More work on the model is needed before any broad conclusions can be drawn about the limits of this modeling system in general. It is not clear to me that showing that this configuration of WRF cannot reproduce the meteorology of the system is helpful. It is not useful to show that this implementation of WRF-GHG fails to simulated atmospheric CO₂ in this environment given the extremely inappropriate use of EDGAR to try to simulate CO₂ emissions in this very local environment.

We agree with both points raised by the reviewer. First, the manuscript does not quantify industrial CO₂ emission rates. We have therefore replaced the statement that the lidar “monitors industrial CO₂ emissions” with the more accurate description that it detects and characterises CO₂ plumes associated with industrial emissions at high spatiotemporal resolution.

Second, we no longer draw broad conclusions about the general limitations of WRF-GHG. The revised conclusion refers only to the performance of the configurations evaluated in this study. The

quantitative comparison shows that the coarse raw EDGAR representation was insufficient to adequately represent the local industrial sources at this scale. The spatially corrected EDGAR_S configuration substantially reduces the negative model bias, while the additional vertical allocation in EDGAR_P further improves the temporal correlation with the lidar observations. The revised manuscript therefore focuses on the effects of emission-inventory treatment under the present model configuration, rather than attributing the remaining discrepancy to general limitations of WRF-GHG.

51. This final paragraph contains some useful thoughts about research needs but it is quite long.

We agree that the original final paragraph was overly long. We have therefore substantially shortened and reorganised it around three research priorities arising from the present results: (1) developing an industrial-park-specific vertical emission profile and deriving observation-based emission estimates; (2) conducting sensitivity experiments involving the emission inventory, boundary-layer parameterisation, terrain resolution, and VPRM parameters; and (3) applying finer-resolution models, such as LES, to better resolve sub-kilometre terrain-forced flows. The broader and repetitive policy-related statements in the original paragraph have also been removed.

52. Code and data availability. There is nothing in this statement about the observations save that it can be obtained by contacting the corresponding author. That is not sufficient data access for a peer-reviewed publication relying primarily on new observations.

We appreciate the reviewer's concern regarding access to the observational data. We agree that more direct access to the observational dataset is appropriate for a study that relies primarily on new lidar measurements. In response, we are preparing to deposit the coherent DIAL dataset used in this study, including the CO₂ concentration and synchronous wind-field products, in a public data repository with a permanent DOI. The Data Availability statement will be updated with the repository link and DOI once the dataset has been archived. In the meantime, the dataset, together with the relevant processing codes and model results, is available from the corresponding author upon reasonable request. The links to the publicly available WRF-GHG input datasets and preprocessing codes have also been retained.

The corresponding text in Lines 716–720 has been revised as follows:

“The coherent DIAL observational dataset used in this study, including the CO₂ concentration and synchronous wind-field products, is being prepared for deposition in a public data repository with a permanent DOI. The repository link and DOI will be provided upon completion of the public archive. In the meantime, the dataset, together with the relevant processing codes and model results, is available from the corresponding author upon reasonable request.”