Reply on RC2

General Comments

The authors of CRUX-1.0 performed data analyses of CO_2 and O_3 measurements from Antarctica. I read the manuscript with great interest, as this type of work is beneficial for advancing our understanding of atmospheric gas abundances and their trends. I recognize that this manuscript should primarily demonstrate the technical and operational capabilities of the measurement system, highlight the challenges of continuous measurements in harsh environments, and provide practical solutions to address those challenges.

The authors are encouraged to reduce extraneous information and focus on the core of their work. Furthermore, they should discuss in more detail how the calibration protocols can be improved and how the power supply can be enhanced to enable long-term observations.

Response: We extend our sincere gratitude for your valuable, detailed, and professional feedback on this paper. Your suggestions have significantly contributed to enhancing the paper's scientific rigor, technical accuracy, data reliability, and adherence to academic standards. We have carefully reviewed each comment and made comprehensive revisions and additions to the manuscript; all modifications are clearly indicated in the corresponding sections.

Major Comments

1. **Lengthy background information:** In several instances, additional information is provided that is not directly required or used in the manuscript. For example, the Introduction and

Section 2.2 (descriptions of other stations) are excessively lengthy and detract from the main technical focus.

Response: Thanks for your feedback. We substantially shortened the redundant background content.

Revision: Removed marine/mobile/terrestrial tower redundant cases; condensed 4 reference stations to brief core descriptions.

This study adopts four representative Global Atmosphere Watch (GAW) background stations covering polar, mid-latitude coastal and alpine regions for atmospheric comparative analysis. The Amundsen-Scott South Pole Station (SPO; 90 ° S, 24.8 ° W; 2837 m a.s.l.) is a NOAA-operated Antarctic background station, performing hourly in-situ CO_2 and O_3 observations with biweekly flask sampling and weekly NOAA-2016 scale calibration (Mahesh et al., 2003). It maintains a 98.8% data continuity rate and provides benchmark Antarctic atmospheric background data (Prinn et al., 2018). Barrow Station (BRW; 71.32°N, 156.61°W; 11 m a.s.l.), a key NOAA GMCC Arctic coastal station, conducts hourly CO_2 and O_3 monitoring and biweekly flask cross-validation. Influenced by North American emissions and Arctic atmospheric circulation, BRW captures both pristine Arctic background air masses and regional pollution episodes, with all data standardized to 10-min intervals via NOAA quality control protocols (Heintzenberg, 1989). The long-term NOAA-operated Mauna Loa Station (MLO; 19.54°N, 155.58°W; 3397 m a.s.l.), established in 1958, hosts the world's longest continuous greenhouse gas observation records. It implements hourly WMO GAW-calibrated CO_2 and O_3 monitoring and daily flask sampling, and its remote marine location minimizes anthropogenic interference to ensure reliable mid-latitude baseline observations (Keeling et al., 1976). As a high-altitude alpine background station under the ACTRIS network, Jungfraujoch Station (JFJ; 46.54°N, 7.96°E; 3580 m a.s.l.) monitors CO_2 and O_3 at 6-min intervals, with weekly flask sampling and monthly systematic calibration, effectively reflecting mid-latitude European atmospheric trace gas variations (Kubistin et al., 2014). The station maintains a long-term data completeness of approximately 92% under routine quality assurance (Sturm et al., 2012). Notably, systematic biases between NOAA and ACTRIS datasets are less than 1% for CO_2 and 3% for O_3, which are sufficiently small to be negligible for regional and global comparative analyses (Prinn et al., 2018).

2. **CO_2 calibration protocol:** CO_2 calibration was performed using a single standard gas. Is this approach sufficient for maintaining WMO-compliant accuracy? How could this protocol be improved (e.g., by incorporating a second standard or zero-air calibration)?

Response: Thanks for your feedback.We greatly appreciate the reviewer's professional and constructive comments on the CO_2 calibration strategy, data accuracy, and basic observational indicators. We have carefully supplemented and revised the relevant content in combination with our pre-experiment results, system improvement plan, and WMO/GAW observation specifications. The detailed responses are as follows:

1. Is single standard gas calibration for CO_2 sufficient for WMO accuracy requirements?

Response: Prior to the Antarctic expedition, we conducted a systematic precision comparison experiment of greenhouse gas analyzers using LICOR-830 (employed in this study) and Picarro G2301 in the laboratory (Figure R-1). We adopted three types of traceable CO_2 standard gases for measurements, and carried out repeatability tests, drift tests, and target standard gas tests (three-point cross-concentration calibration), followed by a quantitative precision comparison of the observed concentrations.

The key test results are:

(1) Repeatability test: The precision of LICOR-830 was ~0.4 ppm, while that of Picarro G2301 was ~0.04 ppm, showing a one-order-of-magnitude difference (Figure R-2);

(2) Drift test: The accuracy of LICOR-830 was −4.22%, and that of Picarro G2301 was 0.82%, with a five-fold difference in accuracy;

(3) Target standard gas test: After 1 hour of calibration using three standard gases of different concentrations (Figure R-3), the precision of LICOR-830 was improved by 15 times to 0.026 ppm (accuracy: −0.11%), and the precision of Picarro G2301 was slightly improved to 0.0341 ppm (accuracy: 0.03%) (Figure R-4).

Figure R-1: Demonstration of the testing procedures using the LICOR-830 and Picarro G2301 instruments in the laboratory

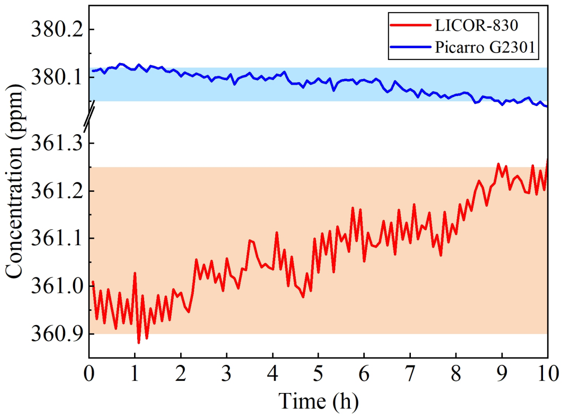

Figure R-2: The average concentration drift for the LICOR-830 and Picarro G2301 instruments occurs every 5 minutes.

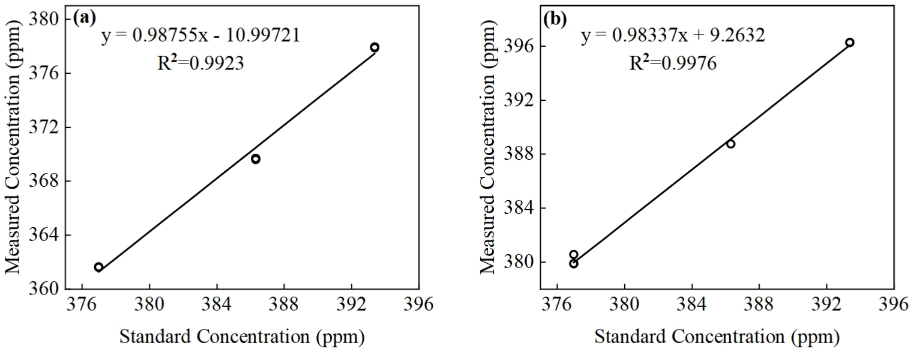

Figure R-3: Scatter plots of standard gas concentration versus measured concentration for the LICOR-830 (a) and Picarro G2301 (b) instruments

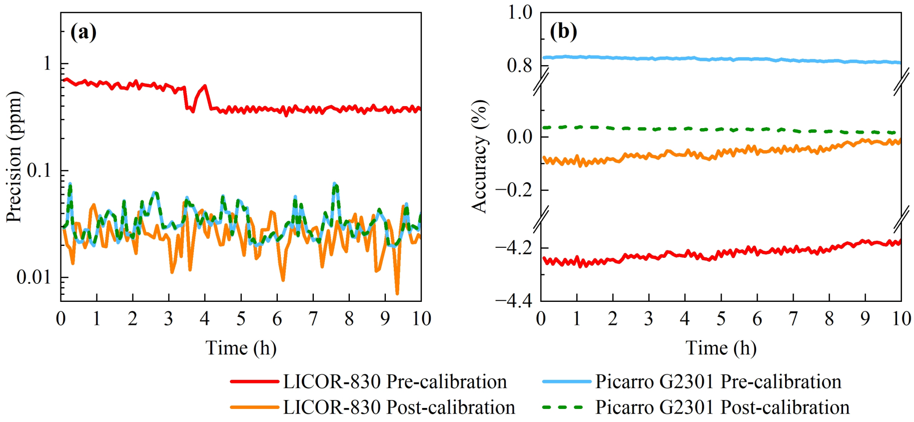

Figure R-4: Comparison chart of pre-correction and post-correction accuracy (a) and precision (b) for the LICOR-830 and Picarro G2301

These results confirm that the calibrated precision of LICOR-830 exhibits application potential. For the CRUX-1.0 system used in this study, it has no built-in sampling sequence control program; thus, we used a data logger to regularly trigger solenoid valves for periodic single-point calibration, which satisfies the data correction demand for short-term deployment. Regular single-point calibration triggered by the data logger, combined with the pre-expedition three-point cross-concentration calibration completed in the laboratory, can fully meet the data correction and precision requirements of this study and comply with WMO accuracy standards.

2. How to improve the calibration strategy for long-term observations?

Response: For long-term unattended high-precision observations, single-point calibration is insufficient, and we acknowledge that this is a major drawback of the current CRUX-1.0 system. To address this issue, we are currently carrying out system integration and field trials of the upgraded CRUX-2.0 system in Lenghu, Qinghai, and the trials are still ongoing. With the support of relevant instrument manufacturers, we conducted performance comparison tests on the low-power portable ABB GLA131-GGA analyzer and the low-power portable Picarro G4301 analyzer in the laboratory. The tests show that the accuracy of the ABB GLA131-GGA analyzer is significantly improved after multi-point concentration calibration, and it can meet the high-precision monitoring requirements of GAW after proper calibration. In addition, this instrument is equipped with an automatic sampling sequence control program, and when combined with a multi-channel solenoid valve box, it can realize autonomous timed calibration with multi-concentration standard gases. At present, the CRUX-2.0 system has been fully integrated and is undergoing field observation trials on Saishiteng Mountain, Qinghai. We expect to complete the trials in June and will continue to share the latest observation and test results of the CRUX-2.0 system with you. We also hope to deploy this system on the Antarctic ice sheet in the future.

For future long-term unattended high-precision observations, single-point calibration is insufficient, and we acknowledge that this is a limitation of the current CRUX-1.0 system. To address this issue, we are currently carrying out system integration and field trials of the upgraded CRUX-2.0 system in Lenghu, Qinghai, and the trials are still ongoing. We conducted a comprehensive performance comparison test between the low-power portable ABB GLA131-GGA analyzer and the low-power portable Picarro G4301 analyzer in the laboratory, covering repeatability, drift, and target standard gas calibration tests, with key results as follows:

Repeatability test: The ABB GLA131-GGA analyzer meets the WMO/GAW inter-site comparability targets for observation precision, achieving repeatability of $\pm 0.2\times10^{-6}$ for CO_2 and $\pm 2\times10^{-9}$ for CH_4 over a 300-second averaging period.

Drift test: The drift ranges of the GLA131-GGA analyzer in CO_2 and CH_4 measurements (0.28×10^{-6} and 0.72×10^{-9}, respectively) are superior to those of the Picarro G4301. The significant drift observed in the Picarro G4301 indicates limited reliability for long-term monitoring, requiring frequent calibration to maintain data accuracy.

Target standard gas test: After calibration, the relative error of CO_2 measurements by the GLA131-GGA analyzer decreased significantly from 1.86% to -0.03%, while the relative error of CH_4 measurements by the Picarro G4301 decreased from -0.39% to -0.15%.

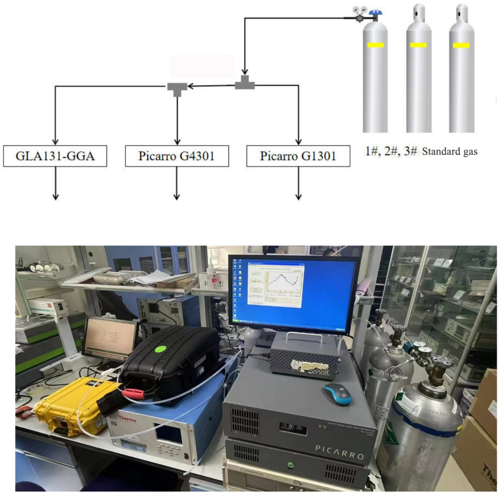

Figure R-5: Schematic diagram of the experimental gas circuit connection and demonstration of the testing procedure

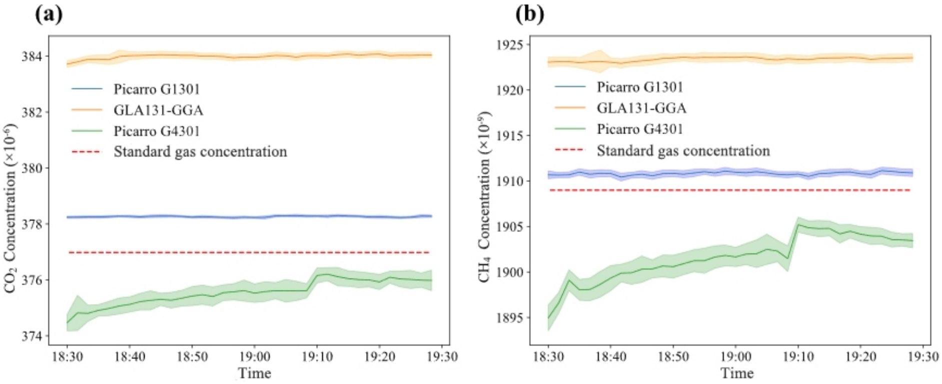

Figure R-6: Average concentration values of measured CO2 (a) and CH4 (b) concentrations over 100-second intervals for the portable instrument

These results confirm that the accuracy of the ABB GLA131-GGA analyzer is significantly improved after multi-point concentration calibration, demonstrating its capability to meet the high-precision monitoring requirements of GAW with appropriate calibration. In addition, this analyzer is equipped with an automatic sampling sequence control program, and when paired with a multi-channel solenoid valve box, it can realize autonomous timed calibration using multiple concentrations of standard gases. At present, the CRUX-2.0 system has been fully integrated and is undergoing field observation trials on Saishiteng Mountain, Qinghai. We expect to complete the trials in July and will share the latest observation results of the CRUX-2.0 system.

3. **O_3 calibration protocol:** Is it sufficient to perform O_3 calibration only before and after the measurement period? Specifically, for deployment periods longer than one month, what is the expected drift, and how should it be addressed? Please elaborate on this in the manuscript.

Response: We appreciate this valuable comment regarding ozone calibration strategy, and we have supplemented detailed descriptions about the rationality of the current pre/post-deployment calibration scheme, quantified long-term instrument drift in Section 4.1 (Calibration Protocol) of the revised manuscript, supported by our previous laboratory intercomparison experiment of Model 205 vs Thermo Model 49i (Wang et al., 2017) and official instrument specification of 2B Technologies. Details are specified as follows:

Calibration of the ozone monitor requires an ozone calibration instrument that meets international traceability standards. Therefore, cross-concentration calibrations are performed every three months according to observation protocols. However, the calibration procedures for Antarctic ice sheet instruments differ from those at manned stations. For instance, in 2011, after the BAS team retrieved all instruments from the observation network to Halley Station, they validated them using the newly calibrated TE-Model 49C. Upon returning to Cambridge, the instruments were re-calibrated using the NPL-certified TE-Model 49iPS, confirming that the calibration parameters of the 10 instruments remained stable throughout the year's deployment, with data deviations not exceeding the initial accuracy range of 2% (Bauguitte et al., 2011). In 2016, testing at Antarctic Dome A confirmed the stable operation of the instrument in the field, and reliable observations can be obtained by conducting calibrations twice per year (Ding et al., 2020a).

Our research group previously conducted a 22-hour synchronous side-by-side intercomparison between the 2B Model 205 ozone analyzer (equipped on CRUX-1.0) and WMO/GAW reference-grade Thermo Model 49i analyzer under ambient atmospheric conditions The comparison results show high consistency with (R^2=0.99419), and 75% of absolute measurement deviations between the two instruments are less than 4 ppb. The Model 205 presents regular systematic bias: slightly lower readings at high O_3 concentration (>100 ppb) and slightly higher readings under low-concentration conditions, which can be well corrected via linear regression based on the five-point pre/post calibration data.

Combining manufacturer’s factory specifications and our long-term tests, the monthly zero drift of Model 205 under constant-temperature laboratory conditions is less than 0.6 ppb. After superimposing the extra thermal drift induced by ± 2.5 °C periodic cabinet temperature fluctuation in Antarctic field, the total monthly comprehensive drift is controlled below 0.7 ppb, still lower than the WMO/GAW permissible total measurement uncertainty of ±0.8 ppb for surface ozone. The ozone raw datasets were calibrated using the five-point cross-concentration formula ($y = 0.9944 * x - 0.4949$, $R^2 = 0.9998$). This calibration function was established in laboratory before instrument shipment via systematic calibration on the Model 205 ozone analyzer with a Thermo 49ips standard ozone calibrator. Upon completion of the QA/QC procedures, the daily-averaged calibrated surface ozone concentration at Taishan Station was 20.1±0.8 ppb.

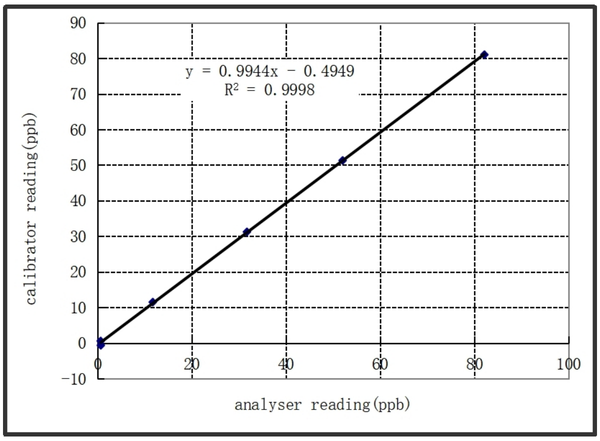

Figure S-5: Dynamic calibration of the linearequation about Model 205

4. **Deployment duration and environmental challenges:** While this study provides a good

demonstration of system capability, a one-month measurement period is insufficient to

validate long-term reliability in such a challenging environment. How can power supply

improvements enable longer deployments? Would it be feasible to extend future deployments to at least three months or longer? Please discuss these challenges and propose concrete solutions.

Response: We appreciate the reviewer ' s valuable comment regarding long-term field deployment limits and power optimization, and we have supplemented detailed descriptions of power configuration differences between CRUX - 1.0 and newly designed CRUX - 2.0 hybrid energy system in the revised manuscript, with the CRUX-2.0 power supply schematic added into supplementary materials (Figure R-7).

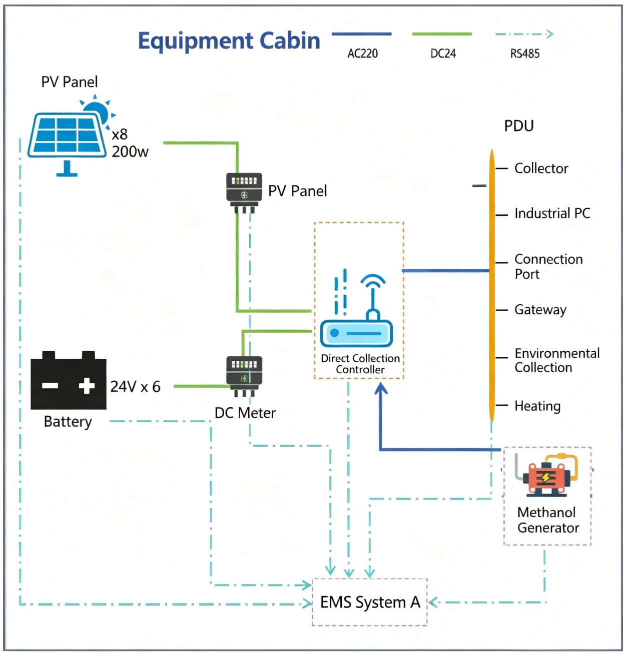

Figure R-7: Schematic Diagram of Standalone PV-Battery-Methanol Hybrid Power Supply for CRUX-2.0

CRUX - 1.0 relies on the fixed diesel power facility from Southeast University at Taishan Station. Its operation requires regular manual seasonal fuel refilling, accompanied by unstable power output and constrained energy reserve, which restricts long-duration observation. To overcome these drawbacks, CRUX-2.0 adopts a standalone PV-battery-methanol hybrid power system (Supplementary Figure R-7). Eight 200 W photovoltaic arrays together with 24 V lithium battery packs supply routine power and store surplus electricity, whereas the EMS automatically triggers the methanol generator to feed all system loads under overcast weather or polar-night darkness, removing frequent on-site refueling demands. This design is engineered to sustain no less than three months of uninterrupted unattended operation. The prototype is now undergoing field experiments at Saishiteng Mountain and will complete validation in July, with planned Antarctic deployment in future CHINARE expeditions to achieve extended multi-month monitoring.

5. **Power supply management:** Can the authors provide a technical solution for managing the station's power supply more effectively? Are there any practical strategies to further reduce power consumption without compromising data quality?

Response: We appreciate this valuable suggestion on optimized power management and energy-saving strategies for long-term unattended polar deployment, and we have added a concise summary of our practical technical roadmap in the revised manuscript.

1. Optimized station power management solution

We are developing an upgraded intelligent power management framework for the next-generation CRUX system. The core design adopts a hybrid configuration combining photovoltaic power, low-temperature lithium energy storage and dual backup generators (methanol + diesel). An upgraded EMS (Energy Management System) is deployed to dynamically switch operational modes following polar seasonal variation: PV dominates power supply during polar summer with surplus electricity stored in batteries; once battery SOC drops to the preset threshold under overcast weather or full polar night, backup generators are automatically activated to recharge batteries and sustain core loads.

Full detailed hardware specification, complete EMS control algorithm and dual-cabin (observation & energy cabin) microgrid topological parameters remain under ongoing technical verification, and these proprietary technical contents will be comprehensively reported in a dedicated follow-up manuscript; thus we only outline the overall management philosophy within the present paper.

2. Practical energy-saving approaches without compromising observational data quality

Field statistics confirm cabin heating accounts for the largest proportion of total power consumption of the current system. Three targeted practical energy-saving measures are being validated at our Qinghai field test site to cut unnecessary power draw while guaranteeing continuous valid measurement:

(1) Passive heat-loss reduction: Optimize cabin enclosure with enhanced multi-layer thermal insulation and structural cold-bridge blocking to minimize conductive heat dissipation under ultra-low ambient temperature, lowering baseline heating demand fundamentally.

(2) Waste heat recycling: Collect waste heat exhausted from running diesel/methanol

generators and redirect such residual heat into cabin interior for auxiliary thermal maintenance, substantially cutting electric heating power input.

(3) Hierarchical intelligent load control: Classify all onboard loads into three tiers via EMS logic. Tier-1 loads (gas analyzers, data acquisition and satellite communication) are always powered unconditionally to secure complete dataset; Tier-2 essential thermal loads run as required; non-critical Tier-3 loads (auxiliary lighting, redundant supplementary heating) will be automatically disconnected when battery reserve is insufficient, achieving rational power curtailment without any loss of observational information.

3. Current research progress

All above optimization schemes are undergoing continuous field verification at Saishiteng Mountain, Qinghai Province, and abundant real-world operational and energy consumption datasets have been accumulated throughout field trials. Quantitative energy-saving efficiency and full-set detailed system technical data will be systematically sorted out and published in our subsequent independent research article after finishing all validation tests.

Detailed Comments

- **Line 158:** Is the reported temperature variability (austral winter vs. summer) statistically or operationally significant for system performance? If so, please clarify.

Response: We appreciate this valuable suggestion, and we have supplemented quantitative analysis supported by the updated Figure 4 (the newly added synchronous ambient temperature time series) in the revised manuscript to clarify the temperature influence on gas measurements:

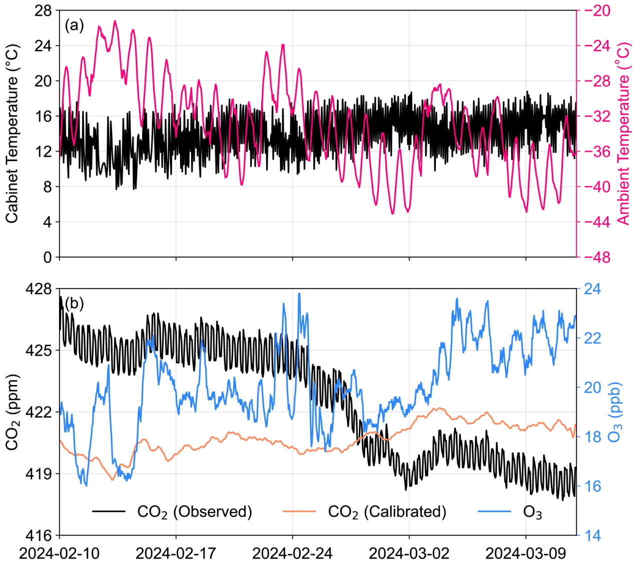

Fig. 4 Compares synchronous in-cabinet temperature (black, left axis) and outdoor ambient temperature (pink, right axis) (a), hourly average concentration of observed CO_2, hourly average concentration of calibrated CO_2 and O3 concentration transmitted from the unattended automated observation system at Taishan Station (b).

From Figure 4a: The pink curve denotes ambient outdoor temperature at Taishan Station (ranging from −45.0 °C to −20.5 °C with strong daily periodic oscillation under Antarctic late summer), while the black curve is the internal cabinet temperature controlled by the PTC active thermal system. Although the cabinet temperature shows a CV of 16%–20% (corresponding to ±2.5 °C short-term fluctuation), the integrated heating cabinet successfully confines instrument operating temperature within 8.0–19.0 °C, avoiding extreme cold interference from ambient air.

From Figure 4b: The raw observed CO_2 (black), calibrated CO_2 (orange), and O_3 (blue). No synchronous correlation is observed between cabinet temperature fluctuation and the variation trends of CO_2 and O_3 concentrations: the long-term evolutions of calibrated CO_2 and O_3 are dominated by natural atmospheric background change rather than cabinet temperature drift. Quantitatively, the ±2.5 °C cabinet fluctuation only induces maximum drift of ≤ 0.2 $ppm \cdot h^{-1}$ for CO_2 and ≤ 0.1 $ppb \cdot h^{-1}$ for O_3. These drift magnitudes are far below the WMO/GAW permissible accuracy thresholds (CO_2: ±1.5 ppm; O_3: ±0.8 ppb).

In conclusion, the 16%–20% CV of cabinet temperature fluctuation exerts negligible statistical impact on final CO_2 and O_3 observation results, and the active temperature control design effectively offsets the severe disturbance from drastic Antarctic ambient temperature swings. Relevant descriptive texts are added in Section 4.2.

Revision:

Original: 4.1 Observational results & 4.2 Accuracy evaluation

Revised:

4.1 Observational results

From February 10 to March 12, 2024, the observation system at Taishan Station operated autonomously under unattended low-temperature conditions, with good performance, and satellite remote data transmission was normal. By analyzing the transmitted data, we found the following: in the environment at Taishan Station, where the average temperature was below -25°C in February, the internal temperature of the equipment cabinet was able to be stably maintained at 13.9 ± 2.5°C (Figure 4a), meeting the environmental temperature requirements for both types of optical cavity equipment used in this experiment.

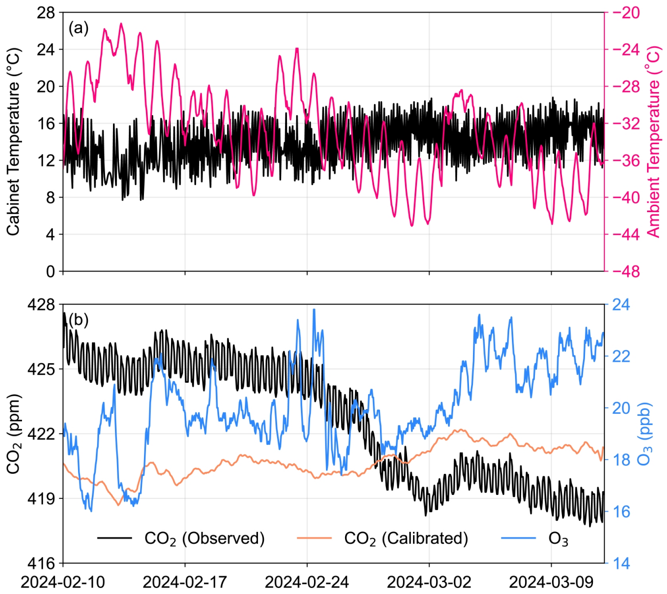

Fig. 4 Compares synchronous in-cabinet temperature (black, left axis) and outdoor ambient temperature (pink, right axis) (a), hourly average concentration of observed CO_2, hourly average

concentration of calibrated CO_2 and O3 concentration transmitted from the unattended automated observation system at Taishan Station (b).

The hourly average CO_2 concentration observed at Taishan Station was 422.6 ± 2.8 ppm (Figure 4b). To constrain instrumental drift under unattended field conditions, the system executes an automatic 10-minute standard gas calibration every 11 hours and 10 minutes; this customized calibration interval comprehensively balances drift correction efficiency, limited on-site power supply and automatic operation requirements. All CO_2 raw datasets underwent standardized multi-step QA/QC procedures sequentially: exhaust interference exclusion based on sampling layout and field wind rose statistics, wind-speed-dependent pollution screening, 3σ statistical outlier removal, quantitative concentration calibration, and final 48-hour moving average smoothing for noise reduction.the daily-averaged calibrated CO_2 concentration was determined as 420.7 ± 0.7 ppm. The ozone observed data were calibrated using the five-point cross-concentration formula ($y = 0.9944 * x - 0.4949$, $R^2 = 0.9998$). This calibration function was established in laboratory before instrument shipment via systematic calibration on the Model 205 ozone analyzer with a Thermo 49ips standard ozone calibrator. Upon completion of the QA/QC procedures, the daily-averaged calibrated surface ozone concentration at Taishan Station was 20.1±0.8 ppb.

4.2 Accuracy evaluation

We evaluated system performance using WMO/GAW required indicators: CO_2 (zero drift ≤0.1 ppm/month, span drift ≤0.2% FS, accuracy ±1.5 ppm); O_3 (zero drift ≤0.5 ppb/month, span drift ≤1% FS, accuracy ±0.8 ppb), plus standard deviation and coefficient of variation (CV). The standard deviation provides insight into the dispersion of measurement data, reflecting the precision and stability of the system. On the other hand, CV, which normalizes the standard deviation by the mean value, allows for a comparison of variability across different parameters and conditions, regardless of their magnitude. Using these two metrics, we evaluated the precision of O_3 and CO_2 concentrations, as well as the cabinet temperature, over time.

Prior to the Antarctic field campaign, we had been completed laboratory performance intercomparison between the LICOR-830 analyzer used in this work and Picarro G2301 by means of repeatability, long-term drift and three-point cross-concentration calibration with traceable CO_2 standard gases, with related test plots (Figures S-1 to S-4) attached in supplementary files. The test results indicated that LICOR-830 had a repeatability of 0.4 ppm versus 0.04 ppm for Picarro G2301, alongside drift biases of −4.22 % and 0.82 % respectively; after three-concentration standard gas calibration for 1 h, LICOR-830's precision was markedly optimized to 0.026 ppm (bias: −0.11 %), while Picarro G2301's precision changed slightly to 0.0341 ppm with an accuracy of 0.03 %. Among the aforementioned results, the accuracy of LICO-R-830 meets the WMO's precision requirements for CO_2 observations after calibration.

In addtion, for the ozone measurement module equipped with a 2B Model 205 dual-beam UV analyzer, dual-point calibration conducted before and after field deployment meets the data quality requirement for one-month Antarctic observation ((Bauguitte et al., 2011; Ding et al., 2020a). The early laboratory comparison test (Wang et al., 2017) between Model 205 and WMO reference Thermo Model 49i proved excellent measurement consistency (R^2=0.9942), with most absolute deviations below 4 ppb, and regular concentration-dependent bias can be fully corrected by linear fitting using pre- and post-campaign calibration results. Laboratory aging experiments demonstrate the intrinsic monthly zero drift of Model 205 is less than 0.6 ppb; after adding extra thermal drift from periodic cabinet temperature fluctuation, the total monthly drift remains below 0.7 ppb, within the WMO/GAW allowed total uncertainty of ±0.8 ppb for ozone.

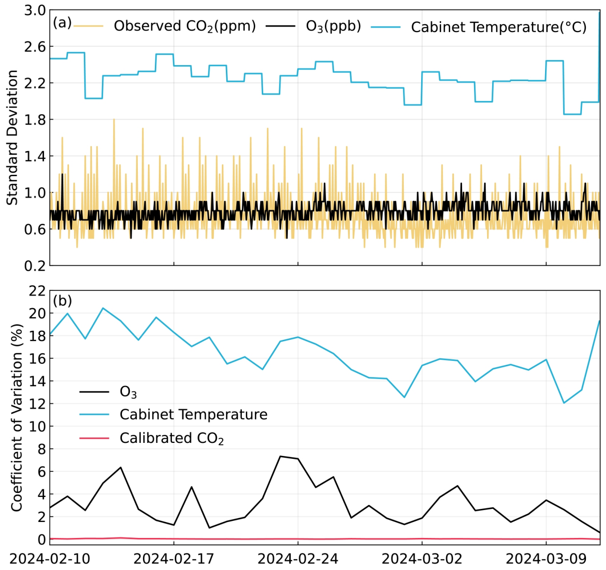

Fig. 5 Evaluation of the Precision of O_3 and CO_2 Concentrations and Cabinet Temperatures through Standard Deviation and Coefficient of Variation Analysis Standard Deviation (a) Coefficient of Variation (b)

The O_3 concentration (black line) consistently has a low standard deviation and minimal fluctuations, indicating high measurement stability. In contrast, the original CO_2 concentration (yellow line) displays substantial variation, with a peak standard deviation of 1.8, suggesting lower precision (Figure 5a). The cabinet temperature (blue line) maintains a stable standard deviation, although a stepwise decrease over time suggests some control, albeit with potential for further improvement. In terms of the CV, the O_3 concentration remains low (<6%) and continues to decrease over time, reflecting enhanced measurement consistency. As illustrated in Figure 4, the ambient surface temperature at Taishan Station fluctuates widely between −45.0 °C and −20.5 °C with prominent diurnal cycles, whereas the active thermal cabinet limits internal instrument temperature within 8.0–19.0 °C with a CV of 16%–20% (±2.5 °C fluctuation). Comparative time-series analysis of CO_2 and O_3 in Figure 4b confirms that such cabinet temperature variation only yields tiny measurement drift ($\leq$0.2 ppm h^{-1} for CO_2, $\leq$0.1 ppb h^{-1} for O_3),, with no significant impact on measurement results (Figure 5b). Postcalibration, the CO_2 concentration (red line) significantly decreases in CV, approaching 0%, demonstrating the effectiveness of calibration in improving measurement accuracy.

- **Lines 167–171:** Why is this wind pattern information included? Was it used in data analysis (e.g., to flag periods of potential contamination or blowing snow)? If not, consider removing or moving it to supplementary material.

Response: We appreciate the reviewer’s valuable suggestion. As the wind pattern information described in Lines 167–171 was not adopted in any subsequent data analysis, including contamination screening and blowing-snow-related data flagging, we have completely deleted the relevant text from the main manuscript in the revised version according to your advice.

- **Lines 179–180:** Did you observe any measurable impact from the generator cabin on your CO_2 or O_3 measurements? If so, how were affected data identified and treated?

Response: We appreciate this critical comment regarding potential generator cabin interference and data quality control. We have supplemented the complete multi-step data quality assurance (QA/QC) protocol in Section 4.2 of the revised manuscript, and verified no contamination from the generator cabin through spatial layout, meteorological analysis, and strict data processing.

1. Spatial layout to avoid exhaust pollution.

The sampling inlet is located 50 m due south of the generator cabin. Generator exhaust can only reach the inlet under northerly winds (337.5°–22.5°). Meteorological records show all wind directions during the campaign were 27°–119° (northeast/east/southeast), with no northerly wind observed. Thus, generator exhaust never reached the sampling system.

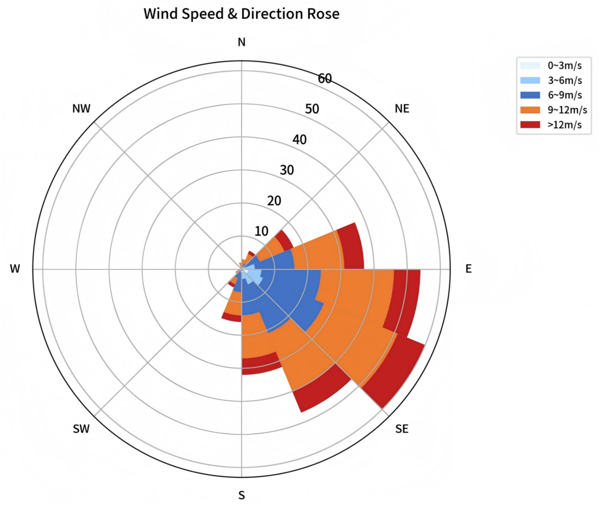

Fig. R-7 Rose diagram of wind speed and direction during the trial period

2. Meteorological screening for local pollution

All measured wind speeds during the entire observation period were consistently higher than 3 m/s. No low-wind stagnant conditions with poor atmospheric dispersion appeared at any time, so no records were removed via wind-speed-based filtering.

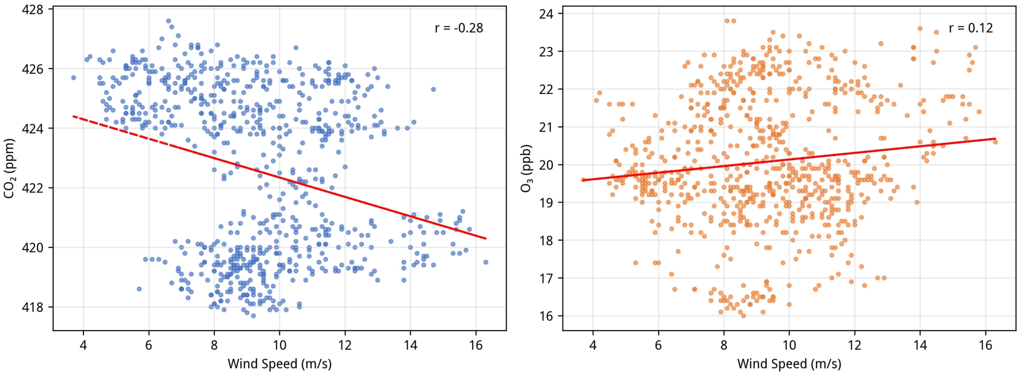

Fig. R-8 Correlation between wind speed and CO_2/O_3 ratios

3. Statistical outlier removal

A 3σ filter was applied to hourly CO_2 and O_3 means to remove anomalous spikes deviating more than 3 standard deviations from the average.

4. Instrument calibration & drift correction

After the field campaign, we further calibrated the dataset using the laboratory cross-concentration calibration formula established pre-deployment, and we corrected the instrument measurement drift using on-site standard gas data.ensuring the data met the original factory accuracy.

5. Temporal smoothing

A 48-hour moving average smoothing was performed to reduce high-frequency noise while preserving the natural variation of atmospheric background concentrations.

- **Section 2.2 (Lines 183–223):** The descriptions of other stations (South Pole, Barrow, Mauna Loa, Jungfraujoch) are lengthy. It is unclear whether all of this information is necessary for the manuscript. This section could be heavily shortened, and any essential comparisons could be moved to the Introduction or Discussion.

Response: Thanks for your feedback. We substantially shortened the redundant background content.

Revision: Removed marine/mobile/terrestrial tower redundant cases; condensed 4 reference stations to brief core descriptions.

Revised Line No.: Lines 178–200 (2.2 Other Observation Stations)

This study adopts four representative Global Atmosphere Watch (GAW) background stations covering polar, mid-latitude coastal and alpine regions for atmospheric comparative analysis. The Amundsen-Scott South Pole Station (SPO; 90 ° S, 24.8 ° W; 2837 m a.s.l.) is a NOAA-operated Antarctic background station, performing hourly in-situ CO_2 and O_3 observations with biweekly flask sampling and weekly NOAA-2016 scale calibration (Mahesh et al., 2003). It maintains a 98.8% data continuity rate and provides benchmark Antarctic atmospheric background data (Prinn et al., 2018). Barrow Station (BRW; 71.32°N, 156.61°W; 11 m a.s.l.), a key NOAA GMCC Arctic coastal station, conducts hourly CO_2 and O_3 monitoring and biweekly flask cross-validation. Influenced by North American emissions and Arctic atmospheric circulation, BRW captures both pristine Arctic background air masses and regional pollution episodes, with all data standardized to 10-min intervals via NOAA quality control protocols (Heintzenberg, 1989). The long-term NOAA-operated Mauna Loa Station (MLO; 19.54°N, 155.58°W; 3397 m a.s.l.), established in 1958, hosts the world's longest continuous greenhouse gas observation records. It implements hourly WMO GAW-calibrated CO_2 and O_3 monitoring and daily flask sampling, and its remote marine location minimizes anthropogenic interference to ensure reliable mid-latitude baseline observations (Keeling et al., 1976). As a high-altitude alpine background station under the ACTRIS network, Jungfraujoch Station (JFJ; 46.54°N, 7.96°E; 3580 m a.s.l.) monitors CO_2 and O_3 at 6-min intervals, with weekly flask sampling and monthly systematic calibration, effectively reflecting mid-latitude European atmospheric trace gas variations (Kubistin et al., 2014). The station maintains a long-term data completeness of approximately 92% under routine quality assurance (Sturm et al., 2012). Notably, systematic biases between NOAA and ACTRIS datasets are less than 1% for CO_2 and 3% for O_3, which are sufficiently small to be negligible for regional and global comparative analyses (Prinn et al., 2018).

- **Lines 269–272 and Table 2:** Is the current CO_2 calibration protocol (single standard gas,twice daily, 5 minutes each) sufficient? Would it be preferable to use a broader span (e.g., a low and a high standard)? What was the mixing ratio of the calibration gas used?

Response: Thanks for your feedback. We appreciate the reviewer's careful comment regarding our CO_2 calibration scheme. Detailed responses on calibration sufficiency, multi-standard improvement suggestion and calibration gas mixing ratio are provided below:

1. Mixing ratio of the employed calibration gas

The single-point reference gas applied for in-situ Antarctic calibration has a CO_2 molar mixing ratio of 420 ppm, which matches the typical background ambient CO_2 concentration of the Antarctic inland atmosphere.

2. Sufficiency of the present CO_2 calibration protocol

Our overall calibration strategy consists of two complementary parts: pre-campaign three-point cross-concentration calibration in laboratory before Antarctic deployment + field single-point calibration at the Antarctic site.

In field operation, single-point calibration runs twice per day with 10 min duration for each calibration cycle (we have corrected the description of 5 min to 10 min in revised manuscript and Table 2 accordingly).

Before the Antarctic fieldwork, we performed comprehensive laboratory characterization of the LICOR‐830 analyzer alongside Picarro G2301 with three independent traceable CO_2 standard gases via three-point cross-concentration calibration, with key test results summarized as below:

(1) Repeatability test: Raw precision of LICOR‐830 was ~0.4 ppm versus ~0.04 ppm for Picarro G2301 (one order-of-magnitude gap);

(2) Drift test: Uncalibrated accuracy of LICOR‐830 was −4.22% versus 0.82% for Picarro G2301;

(3) Three-point cross-calibration test: After 1-hour multi-concentration calibration in lab, LICOR-830's precision was improved 15-fold to 0.026 ppm with an accuracy of −0.11%, which fulfills the WMO Southern Hemisphere CO_2 precision criterion (±0.05 ppm).

Combining the well-constrained pre-expedition three-point laboratory calibration and twice-daily 10-min on-site single-point regular calibration controlled by data logger-driven solenoid valves, short-term (1-month) Antarctic observation can fully satisfy GAW accuracy requirements, so the existing protocol is sufficient for our short-duration field measurement in this manuscript.

- **Lines 328–331:** Please discuss the O_3 calibration protocol in greater detail, specifically addressing its adequacy for longer observation periods (e.g., multiple months). How was drift assessed and corrected?

Response: Thank you for this important comment on the O_3 calibration strategy and long-term drift control.

Sufficiency for the 1-month deployment: Pre- and post-deployment calibration is fully sufficient for this one-month field experiment at Taishan Station. The 2B Technologies Model 205 O_3 monitor used in CRUX-1.0 has an extremely low intrinsic drift: its baseline drift is < 3 ppb/year and sensitivity drift < 3%/year, as specified by the manufacturer. Our previous laboratory intercomparison (Wang et al., 2017; Tian et al., 2022) and year-round validation at Kunlun Station (Dome A, Antarctica) further confirmed that the drift over one month is less than 2%, well within the WMO/GAW accuracy requirement (±0.8 ppb for O_3).

Long-term drift mitigation (for ≥ 3-month deployment): For multi-month unattended operation, we propose a complete solution:

①Implement monthly automatic zero/span calibration using an integrated ozone generator and zero-air scrubber;

②Apply a dynamic drift correction algorithm based on internal temperature, pressure, and runtime data to compensate for slow sensor drift;

③Use the linear correction equation established in our laboratory calibration ($R^2 > 0.999$) to post-process the data;

④Maintain the active thermal control to reduce temperature-related drift.

These methods have been verified in polar ozone monitoring programs (Bauguitte et al., 2011; Ding et al., 2020a).

Revision:

Revised Line No.: Lines 305–313 (3.2.2 Ozone Monitor)

Original: Owing to the overall energy consumption limitations and unattended conditions at Taishan Station, we conducted two calibrations, one before and one after the observation experiment, with data deviations not exceeding the initial accuracy range of 2%.

Revised: Owing to the overall energy consumption limitations and unattended conditions at Taishan Station, we conducted calibrations before deployment and after retrieval. This strategy is fully sufficient for the one-month experiment, because the 2B Model 205 monitor features ultra-low intrinsic drift (<3 ppb/year for baseline drift, <3%/year for sensitivity drift). The drift over one month was verified to be <2% in laboratory tests and Antarctic field applications (Wang et al., 2017), meeting the WMO/GAW accuracy requirement.

Figures

- **Figure 2:** The numbers (axis labels and legends) are very small and difficult to read. Please increase font size for better legibility.

  Response: We enlarged font size of axis labels and legends in Figure 2.

  Revision: Redrawn Figure 2 with enlarged font size for all text elements.

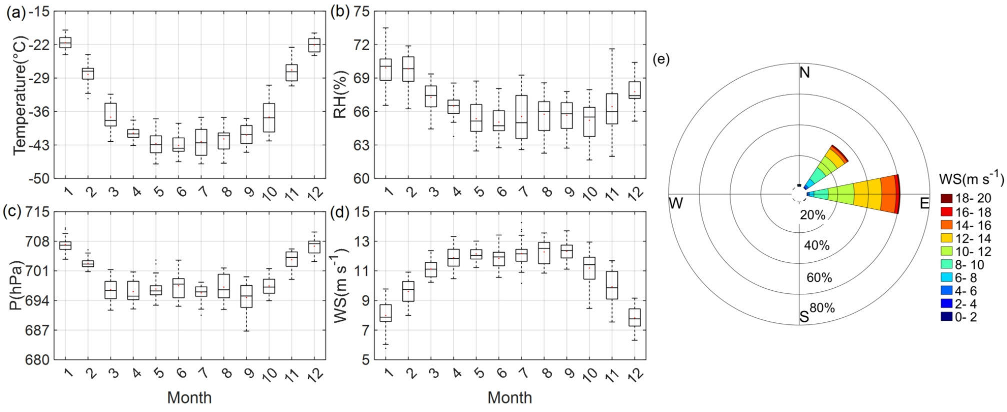

Fig. 2: Multiyear monthly average temperature (a), relative humidity (b), atmospheric pressure (c), wind speed, and wind direction (d, e) at Taishan Station.

- **Figure 6:** What specific insight does the comparison of continuous measurements from all five stations provide? The purpose of this figure should be more clearly stated in the caption and main text.

  Response: We greatly appreciate this constructive comment. We have revised both the main text (Section 5.2 Site Comparison) and the caption of Figure 6 to clearly state the specific scientific insights and the purpose of the five-station comparison.

  In the main text (Section 5.2), we have rewritten the opening paragraph to explicitly define the three core goals of this comparison:

  ①To validate the global representativeness of CO_2 and O_3 measurements from CRUX-1.0 at Taishan Station;

  ②To quantify regional differences in background greenhouse gas and ozone levels between polar and mid-latitude regions;

  ③ To verify the reliability of the unattended observation system in extreme Antarctic environments.

  We have also strengthened the interpretation of results to highlight key scientific insights, such as the consistent background levels between Taishan Station and the South Pole Station, the low-emission characteristics of the Antarctic inland, and the weak photochemical ozone production in polar regions.

  In the caption of Figure 6, we have added a clear statement of the figure's purpose: to validate the reliability and global representativeness of CRUX-1.0 measurements, reveal regional differences in atmospheric CO_2 and O_3, and confirm the role of the unattended system in complementing the global atmospheric background network.

Tables

- **Table 1:** I could not locate Table 1 in the manuscript. Please check the numbering or provide the missing table.

Response: We apologize sincerely for this table numbering error caused by our negligence. The original Table 1 was removed during the revision, but we failed to adjust the numbering of subsequent tables accordingly. We have now renumbered the original Table 2 as Table 1 and corrected all table citations and numbering throughout the manuscript to ensure consistency. The table numbering issue has been fully resolved.

- **Table 2:** Comparing power usage across stations shows that CRUX-1.0 consumes substantially more power (350 W) than the BAS Ozone Network (11–13 W) or the Kunlun Station Ozone Monitor (5 W). Is this increase solely due to the addition of CO_2 measurements and active temperature control? Are there practical ways to lower power usage without degrading data quality? Please discuss.

Response: We sincerely appreciate this valuable comment on power consumption optimization. We fully agree that CRUX-1.0 consumes more power (350 W) than the BAS Ozone Network and Kunlun Station Ozone Monitor, and we have supplemented a detailed discussion on the causes of power consumption, reasonable design considerations, and practical power reduction strategies in the revised manuscript (Section 5.1 and Table 1).

Root cause of higher power consumption

The higher power consumption is indeed mainly caused by two core designs: (1) Simultaneous in-situ measurements of both CO_2 and O_3, whereas the BAS and Kunlun systems only monitor O_3; (2) Active temperature control using a 250 W non-electrified surface PTC heating module with a forced air blower, which accounts for more than 70% of the total power consumption. This active heating cabinet is essential to stabilize the instrument working environment at 10–15 °C under extreme Antarctic low temperatures (down to -40 °C).

Rationale for choosing the current heating scheme

We initially considered a lower-power alternative: self-regulating temperature heating cable (10–35 W/m at 10 °C) wrapped around instruments and covered with aerogel insulation film. Although this combination provides good thermal insulation, it requires on-site winding, wrapping, and packing, which is extremely complicated to implement in the harsh Antarctic field environment. In addition, a 20 m heating cable (required for full coverage) yields a total power consumption similar to the current PTC cabinet, providing no significant power-saving benefit.

We also learned a critical lesson from the BAS Antarctic autonomous ozone network: its low-power passive design seriously underestimated the self-heating energy consumption during polar night, leading to severe battery depletion, frequent system shutdowns, and significant data loss in polar night. To ensure uninterrupted operation and high data continuity in the extreme Antarctic inland environment, we prioritized system stability and on-site maintainability, and adopted the integrated PTC heating cabinet that supports fast installation, convenient pipeline connection, and easy on-site inspection and maintenance.

In future work, we will continue to test different thermal insulation methods based on your suggestions, and strive to identify a solution with lower energy consumption and better thermal insulation performance.