Response to Referee #2

We provide below our point-by-point responses (in blue) to the comments from the reviewer, and the corresponding changes implemented in the revised manuscript.

General Comments

This manuscript addresses an important and practical problem in atmospheric trace gas measurements: the accurate calibration of methane isotopic composition ($\delta^{13}CH_4$) using cavity ring-down spectroscopy under variable conditions. The motivation is clear and relevant: isotopic measurements of CH_4 are a key tool for source attribution, but current calibration strategies can be biased by water vapor interference and instrument drift, particularly in moist air. The authors conduct laboratory and field evaluations of two calibration strategies, a typical delta-based calibration and an isotopologue-specific calibration, and apply empirical humidity corrections to quantify performance differences. While humid air can generally be dried, the study is particularly interesting because it reveals a CH_4-dependent bias in the delta-based calibration.

The study is well structured, with a logically ordered presentation of methods, calibration experiments, and field application. The topic fits the requirements of AMT, as it focuses on methodological improvements in high-precision trace gas analysis. The results, particularly the finding that $\delta^{13}CH_4$ can be biased under humid conditions and under varying methane concentrations, can be significant for the community. However, there are several important points that require clarification and justification. Some of these may constitute major issues before the manuscript is ready for publication. These are described below.

AR: Thank you for the constructive and encouraging assessment of our manuscript. We appreciate your positive evaluation of the manuscript's relevance, structure, and suitability for AMT, as well as your recognition that the observed inter-method differences under humid and variable-CH_4 conditions are of practical importance for the community. We also appreciate your request for further clarification and justification of several methodological aspects.

We have revised the manuscript extensively in response to your comments. Key updates include:

(i) an updated Jurong field analysis (including revised processing and supplementary correlation analyses in Appendix Figs. A1) to better separate humidity-linked effects from concentration-related behavior and to avoid over-interpreting field correlations;
(ii) an Allan deviation analysis to quantify short-term precision and justify the operational averaging time used for working-standard calibration;
(iii) a more explicit description of how assigned reference-gas values (and uncertainties) are used to anchor the calibration and how calibration parameters are applied in the field;
(iv) clarify what is corrected by the manufacturer's internal algorithms versus our offline empirical humidity corrections;
(v) consistent terminology and notation for reference gases and working standards, including an updated naming scheme to avoid ambiguity. We address each specific

point below and indicate where corresponding changes were made in the revised manuscript.

Major Comments

- The manuscript does not clearly define how accurately $\delta^{13}CH_4$ is intended to be measured and for which scientific purpose. Please clarify the target precision and accuracy (e.g. WMO compatibility goals) and how these relate to the different measurement environments. In particular, requirements may differ substantially between a background site and sites with strong variability and source influence such as Jurong.

AR: Thank you for this insightful comment. The WMO/GAW network compatibility targets for $\delta^{13}CH_4$ are ±0.02‰ for well-mixed background air and an extended target of ±0.2‰ for large-signal (source-influenced/urban) conditions (k=2), as summarized in IAEA-TECDOC-2066 (IAEA., 2024). We evaluate our target precision and accuracy mainly using target gases and report the uncertainties in the form of mean bias (μ) relative to the assigned $\delta^{13}CH_4$ value. The results of the comparison are presented in Table R1.

Table R1. Comparison of isotopologue-specific and delta-based approaches under dry and humidified target-gas conditions.

Condition	Approach	μ (‰)	σ (‰)	RMSE (‰)	MAE (‰)	Relative to compatibility goals (±0.2‰ or ±0.02‰)
Dry target-gas	isotopologue-specific (Fig. 4d)	+0.15	2.17	2.18	1.81	**Within ±0.2‰; +0.13‰ above ±0.02‰**
	delta-based (Fig. 4e)	+0.55	1.90	1.98	1.59	+0.35‰ above ±0.2‰
Humidified target-gas (uncorrected)	isotopologue-specific (Fig. 6d)	−0.84	0.30	0.89	0.84	0.64‰ outside ±0.2‰
Humidified target-gas (water-corrected)	isotopologue-specific (Fig. 6d)	−0.31	0.24	0.39	0.34	**0.11‰ above ±0.2‰**
Humidified target-gas (uncorrected)	delta-based (Fig. 6e)	+0.48	0.26	0.55	0.49	0.28‰ above ±0.2‰
Humidified target-gas (water-corrected)	delta-based (Fig. 6e)	+0.49	0.26	0.56	0.50	0.29‰ above ±0.2‰

Note: Reported metrics include mean bias (μ), standard deviation (σ, where available), RMSE, and MAE. For humidified target gas, isotopologue-specific results are shown before and after water-vapor correction; the delta-based results are essentially unchanged. Deviations relative to the ±0.2‰ compatibility goal (and the ±0.02‰ benchmark, where relevant) are indicated in the table.

The Jurong site is located in rice paddy fields and can be considered to represent large-signal conditions, where we set our measurement target according to the WMO/GAW extended target of ±0.2‰. The SORPES site provides both background and large-signal measurements, and the appropriate measurement accuracy target would be between ±0.02‰ and ±0.2‰, i.e., between the WMO/GAW target and the extended target. Our target-gas measurement results indicate that measurements using the isotopologue-specific approach under dry conditions meet the WMO/GAW extended target of ±0.2‰, while an effective water-vapor correction combined with the isotopologue-specific approach can bring performance close to this extended

target under humid conditions. However, in both cases, they do not meet the WMO/GAW network compatibility target of ±0.02‰ for well-mixed background air. Further reduction toward ±0.02‰ will require a more precise and stable analyzer, in combination with sufficiently frequent calibration against reference gases.

The added text in the revised manuscript of Section 3.4 reads as follows:

"Jurong is situated in rice paddy fields and is therefore best characterized as a large-signal, source-influenced environment, and we refer to the WMO/GAW extended compatibility target of ±0.2‰ for the target accuracy of our measurements (IAEA., 2024). In contrast, SORPES captures both well-mixed background periods and large-signal episodes, so the appropriate accuracy requirement depends on the specific application and is expected to fall between the WMO/GAW network target (±0.02‰, for well-mixed background air) and the extended target (±0.2‰, for large-signal (source-influenced/urban) conditions). Our target-gas tests suggest that the isotopologue-specific approach can satisfy the ±0.2‰ criterion under dry conditions, and that applying an effective water-vapor correction brings performance close to this extended target under humid conditions. However, in both cases, they do not meet the stringent ±0.02‰ compatibility level required for background-network applications. Achieving that level will likely require improved instrumental precision and stability, together with sufficiently frequent calibration to reference gases."

- The manuscript lacks a sufficiently detailed description of how the values of the reference gases were obtained. How were their CH_4 concentrations and $\delta^{13}CH_4$ values calibrated relative to the working standard linked to INSTAAR? Were the reference gases calibrated using a delta-based or isotopologue-specific approach, and could this choice influence the comparison of calibration strategies later in the manuscript? This point is particularly important to avoid potential circular reasoning when evaluating the performance of the calibration strategies.

  AR: Thank you for commenting on this important aspect.

  In our study, CH_4 mole fractions were reported on the World Meteorological Organization (WMO) X2004A scale. $\delta^{13}CH_4$ values were reported relative to the Vienna Pee Dee Belemnite (VPDB) scale. The $\delta^{13}CH_4$ values of the working standard gases (WS1-5) were established against reference gases (Ref1-3) which were traceable to the INSTAAR laboratory scale using the Picarro G2201-i analyzer.

  The working standard gases were calibrated based on the isotopologue-specific approach. We have now revised the nomenclature of our standard gases to more clearly distinguish reference gases from working standard gases, and we updated the manuscript consistently to reflect this change. Specifically, we removed Ref0, which duplicated the assigned values of Ref2, and we now use Ref2 directly in its place. All related labels and references have been updated throughout the text, figures, tables, and processing description to avoid ambiguity and any impression of circular reasoning. The added text in the revised manuscript of Section 2.2.3 reads as follows:

  "For the assigned reference-gas values (Ref1–3), CH_4 mole fractions were reported on the World Meteorological Organization (WMO) X2004A scale, and $\delta^{13}CH_4$ values were reported relative to the Vienna Pee Dee Belemnite (VPDB) scale. For the remaining working standard gases (WS1–WS5), the values were assigned in our laboratory by calibration against the assigned reference-gas values, thereby tying our

working scale to the Institute of Arctic and Alpine Research (INSTAAR, University of Colorado Boulder) laboratory scale. The working standard gases were calibrated based on the isotopologue-specific approach. Section 2.3 then describes two alternative strategies for applying these assigned reference-gas values to calibrate and correct the laboratory and field measurements."

The revised version of Table 1 is shown below.

Table 1 Certified values of CH_4 molar fraction and $\delta^{13}CH_4$ for dry and humidified reference gases.

Reference Standard	CH_4 (ppb)	$\delta^{13}CH_4$ (‰)	^{13}r	R_{sum}	$^{12}CH_4$ (ppb)	$^{13}CH_4$ (ppb)
					True values	
WS1	2004.32±0.57	-46.80±0.035	0.010656929	1.011223724	1982.07±0.564	21.12±0.006
WS2	3592.80±0.3	-47.01±0.035	0.010654656	1.011221449	3552.93±0.297	37.86±0.030
WS3	5017.03±1.5	-47.16±0.021	0.010652979	1.011219770	4961.36±1.48	52.85±0.020
WS4	1985.35±0.11	-48.00±0.700	0.010645125	1.011274904	1963.34±1.1	20.90±0.010
WS5	1983.94±0.07	-49.06±0.700	0.010632171	1.011261942	1961.97±1.1	20.86±0.010
Ref1	1831.6±0.3	-47.85±0.671	0.010645205	1.011211993	1811.29±0.3	19.28±0.014
Ref2	1979.9±0.3	-48.14±0.534	0.01064203	1.011208816	1957.95±0.3	20.84±0.012
Ref3	2219.2±0.3	-45.98±0.497	0.01066609	1.011232889	2194.55±0.3	23.41±0.013

Note: CH_4 and $\delta^{13}CH_4$ are certified values for each cylinder. ^{13}r denotes the ratio between $^{13}CH_4$ and $^{12}CH_4$, calculated from Eq. (5). R_{sum} is the total isotopologue normalising factor for methane, defined in Eq (6). This correction accounts for the absorption effects introduced by the presence of hydrogen isotopologues such as CH_3D in atmospheric methane, in addition to carbon isotopologues. $^{12}CH_4$ and $^{13}CH_4$ were derived quantities calculated from assigned total CH_4 and $\delta^{13}CH_4$; their uncertainties were obtained by uncertainty propagation. In this cylinder set, WS1–WS5 are working standards calibrated against Ref1–Ref3. Ref1–Ref3 have manufacturer-certified CH_4 and $\delta^{13}CH_4$ values and are used as the reference standards for the calibration of WS1–WS5.

- The manuscript does not sufficiently address whether the derived humidity and calibration correction functions are stable over time. Were humidity tests performed only once or repeated multiple times and averaged (as, for example, required in the ICOS ATC MLab initial tests)? Over what time period were the laboratory experiments conducted? A discussion of temporal stability and the need for periodic repetition of these tests is essential for assessing the long-term applicability of the proposed calibration approach.

  AR: Thank you for this important comment. In this study, the empirical water-vapor correction functions were derived from a single laboratory humidity-response experiment (approximately 6 hours). We did not perform repeated humidity tests over time, so the temporal stability of these fitted water-vapor correction functions could not be assessed in this study. We have now clarified this limitation in the revised manuscript. The added text in the revised manuscript of Section 2.2.1 reads as follow:

  "The main humidity-response experiment was conducted once in this study. As no repeated humidity tests were performed over time, we cannot assess the temporal stability and reproducibility of the fitted correction functions. For long-term applications, periodic validation and, if necessary, re-derivation of the correction functions is recommended."

- Figure 7 shows that the $\delta^{13}CH_4$ values have a cross sensitivity to the CH_4 concentration and the H_2O content of the sample, but does not clearly distinguish between the effects of humidity and CH_4. A clear CH_4 dependency in $\delta^{13}CH_4$ was already reported in other studies (e.g. Miles et al., 2018 https://doi.org/10.5194/amt-

11-1273-2018 and Rella et al., 2015). Additional laboratory tests under dry and wet conditions (e.g. a dilution series with reference gases) are essential to isolate the effects of $\delta^{13}CH_4$ dependency on CH_4 and to enable a robust scientific interpretation of the results.

AR: Thank you for pointing this out. We revised the Jurong processing because the previous calibration handling could introduce time-dependent artifacts that co-vary with CH_4 and H_2O, thereby confounding the interpretation of Fig. 7. Specifically, treating the isotopologue-specific calibration coefficients (a and b) as piecewise constants (interval means) can create step-like biases within and across calibration periods, which may appear as an apparent CH_4 dependence when CH_4 varies systematically in time. In addition, referencing the delta-based one-point correction to the mean of multiple standards can propagate variability among standards into the corrected $\delta^{13}CH_4$ time series. Reducing these processing-induced effects is therefore a prerequisite for disentangling humidity-driven behavior from concentration-related dependency in the field dataset.

For the isotopologue-specific strategy, the multi-concentration calibration coefficients (a and b for $^{12}CH_4$ and $^{13}CH_4$) now linearly interpolated in time between calibration periods replacing the previous interval-mean approximation. For the delta-based strategy, the one-point correction is referenced to the most stable standard gas, rather than using the average of three standards. The following text has been added to Section 2.3.1 and Section 2.3.3, respectively:

"For field processing, these coefficients were linearly interpolated in time between calibration periods, rather than applying an interval-mean approximation, to obtain time-resolved calibration parameters for each measurement."

"In both field measurements, the two reference gases used for the delta-based calibration had very similar $\delta^{13}CH_4$ values. Therefore, a_δ was effectively set to 1 and the delta-based calibration was implemented as a one-point correction."

Using the reprocessed Jurong data, we updated Fig. 7. The revised text in the revised manuscript of Section 3.3 reads as follow:

"…… exhibiting significant correlations with H_2O ($R^2 = 0.91$, $p < 0.001$) and with CH_4 ($R^2 = 0.28$, $p < 0.001$). The overall $\Delta\delta^{13}CH_4$ distribution is centered at −1.06‰ ($\sigma = 1.92$‰; $n = 3126$) (Fig. 7c). This bias varies with CH_4 concentration, and the fitted mean $\Delta\delta^{13}CH_4$ shifts from −1.55‰ (2000–2500 ppb) to +2.20‰ (5000–5500 ppb) (Fig. 7c)."

"To further interpret Fig. 7, we added the Jurong correlation analysis in Fig. A1. The inter-method difference ($\Delta\delta^{13}CH_4$) shows only a weak relationship with CH_4 mole fraction ($R^2 \approx 0.30$) and is not statistically significant ($p > 0.05$; Fig. A1a), which may reflect the intrinsic concentration dependence for the delta-based approach has been reported previously (Rella et al., 2015; Miles et al., 2018).

In contrast, $\Delta\delta^{13}CH_4$ correlates strongly with H_2O ($R^2 \approx 0.91$, $p < 0.001$; Fig. A1b). This is most likely driven by the significant correlation between the calibrated $\delta^{13}CH_4$ using the delta-based approach and H_2O (Fig. A1d), as the correlation between the calibrated $\delta^{13}CH_4$ using the isotopologue-specific approach and H_2O is insignificant (Fig. A1c)."

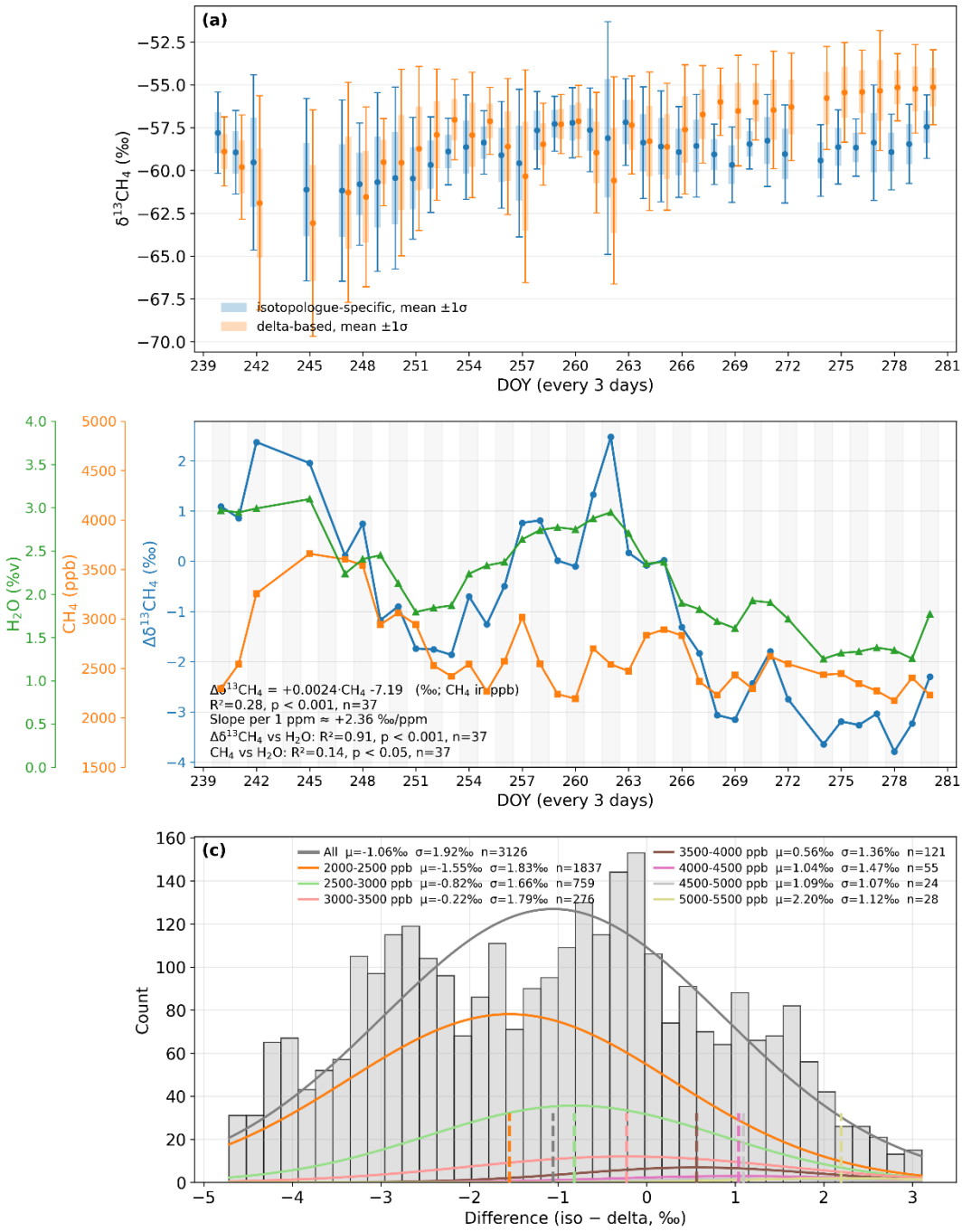

Fig. 7. Field comparison of $\delta^{13}CH_4$ calibration strategies for humid air observations at the rice paddy site (DOY 240–280, 2018). (a) Daily mean $\delta^{13}CH_4$ from isotopologue-specific (blue) and delta-based (orange) calibrations (mean ± 1σ). (b) Time series of the inter-method difference in $\delta^{13}CH_4$ (iso − delta, blue line) together with total CH_4 mole fraction (orange line) and H_2O concentration (green line, 0.93% to 3.5%). (c) Histogram of $\delta^{13}CH_4$ differences (iso − delta) grouped by CH_4 mole-fraction ranges, each fitted with a Gaussian function. The fitted mean (μ) and standard deviation (σ) are reported for each subset. All results are based on 5-min averages.

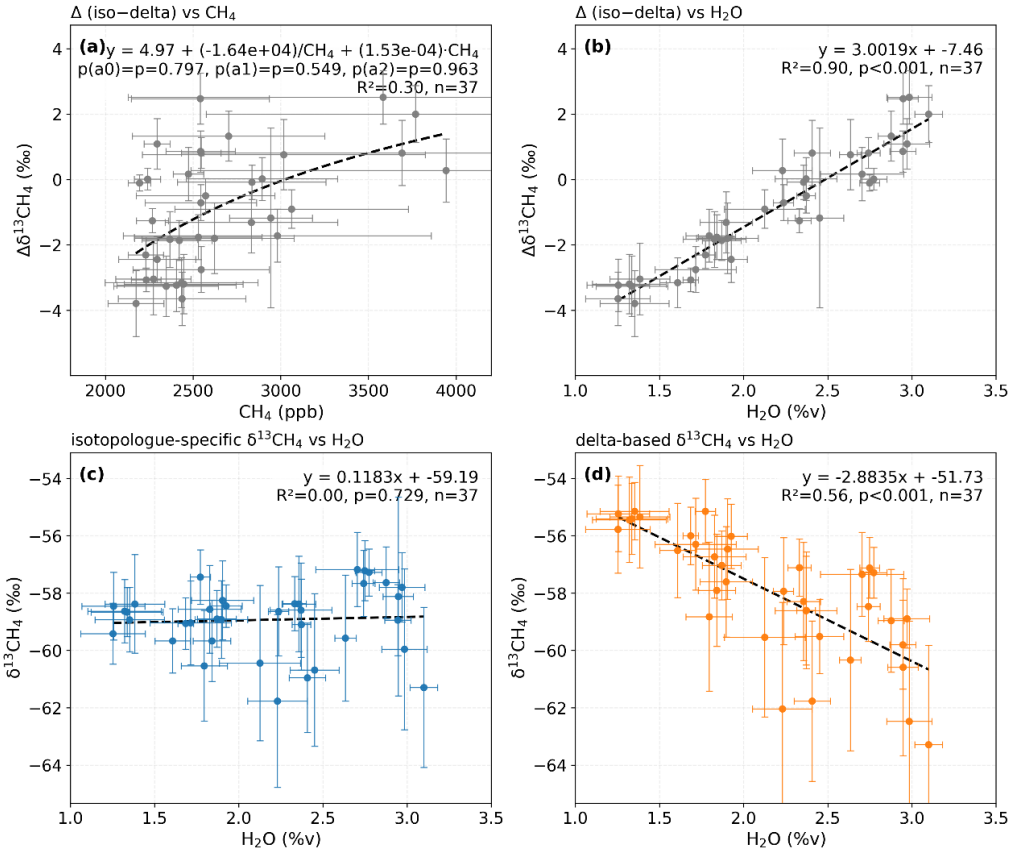

Fig.A1 Correlations between $\delta^{13}CH_4$ calibration results and CH_4 and H_2O for humid air observations at the Jurong site. Relationship between (a) $\Delta\delta^{13}CH_4$ and CH_4; (b) $\Delta\delta^{13}CH_4$ and H_2O; (c) isotopologue-specific $\delta^{13}CH_4$and H_2O; (d) delta-based $\delta^{13}CH_4$ and H_2O. Dashed lines indicate fitted regressions. The fitted equations, R^2, p, and n are reported in each panel.

Specific Comments

- L. 107: It is not completely clear whether the same instrument was used for both lab and field measurements, or if different instruments of the same model were employed. Instrument-specific behavior (e.g., detector response, CH_4 dependency and water vapor sensitivity) can affect instrument performance (see Miles et al., 2018). Moreover, even if the same instrument was used for both lab and field measurements, its characteristics may change over time, which could also influence the results. The authors should clarify how many instruments were used and how any instrument-to-instrument variability was assessed.

  AR: We clarify that the same Picarro G2201-i analyzer was used for all laboratory experiments and field measurements both in the SORPRE and Jurong sites. We also acknowledge that instrument characteristics can evolve over multi-year periods, even for the same analyzer (e.g., Miles et al., 2018). As no repeated humidity tests were performed over time, we cannot assess the temporal stability and reproducibility of the fitted correction functions.

  We added the following sentence in Section 2.1 to clarify this:

  "All laboratory experiments and field measurements reported here were performed using the same G2201-i analyzer, although the field measurements were performed during different campaigns in different years (2018, 2022, and 2025)."

- L. 118: The description of the internal water correction applied by the instrument is unclear. Please specify what type of correction is implemented by the manufacturer (e.g., linear, quadratic, or higher-order empirical function) and whether this correction is instrument-specific or identical for all analyzers of the same model.

  AR: According to the G2201-i user manual, the analyzer applies an internal water-vapor–related correction that is empirically parameterized using the measured H_2O absorption signal (peak75). The manual indicates that this correction is non-linear and includes a quadratic term, rather than a single linear function. The manual does not specify whether the internal correction is identical across analyzers; therefore, we treat it as part of the analyzer's internal implementation and do not use the manufacturer-corrected outputs in our calibration workflow. We have clarified in the revised manuscript that this manufacturer-implemented internal correction is distinct from the externally derived empirical water-vapor correction functions developed in this study. The added text reads as follows:

  "The analyzer applies an internal, empirically parameterized water-vapor interference correction based on the measured H_2O absorption signal (peak75), which is non-linear and includes a quadratic term, as described in the G2201-i user manual."

  "In this study, all water-vapor correction functions were derived from the raw (uncorrected) outputs. All subsequent analyses were performed using raw $^{12}CH_4$, $^{13}CH_4$, and $\delta^{13}CH_4$ values corrected offline with the empirical water-vapor correction functions developed here. Manufacturer-corrected outputs were not used in the calibration workflow to avoid potential double correction."

- L. 165: A 5-minute averaging time is relatively short for high-precision isotopic measurements. Please justify this choice with respect to measurement precision and stability, e.g. provide Allan variance analysis and to typical atmospheric variability at the two sites. Previous studies (e.g. Hoheisel et al., 2019) suggest that longer

averaging times are often necessary to achieve optimal $\delta^{13}CH_4$ precision before drift becomes dominant.

AR: We therefore added an Allan deviation analysis (Fig. A2) to quantify measurement precision and stability. The Allan deviation of $\delta^{13}CH_4$ (raw) decreases monotonically over the accessible τ range, with σ (300 s) = 0.0095‰ and σ (600 s) = 0.0048‰, and no drift-dominated upturn is observed within this range. Although 600 s yields lower random noise, we retain 5-min averaging as a practical compromise that already provides very low short-term noise while improving operational efficiency and reducing the consumption of both reference and working-standard gases.

In addition, the 5-min precision was evaluated empirically from repeated reference-gas and target-gas injections processed with the same workflow as the ambient measurements. For the humidified target gas after water-vapor correction, the residual distributions show Gaussian-fit standard deviations of ~0.24–0.26‰ (Fig. 6d–e), and the inter-method difference distribution is also narrow (σ ~0.14‰; Fig. 6f), supporting repeatability and stability under our measurement conditions. We also clarified the valve-switching and averaging protocol in Section 2.2.3; the 5-min window was chosen to avoid stabilization immediately after switching and to represent the stable plateau of each injection. While previous studies (e.g., Hoheisel et al., 2019) note that longer averaging can be beneficial when drift becomes dominant, our Allan analysis indicates that within the operational range relevant to our datasets, increasing τ continues to reduce random noise, and 5-min averaging is sufficient for the purposes of this study.

The added text in Section 2.2.3 reads as follows:

"The Allan deviation of $\delta^{13}CH_4$ (raw) decreases monotonically over the accessible τ range, with σ (300 s) = 0.0095‰ and σ (600 s) = 0.0048‰ (Fig. A2), and no drift-dominated upturn is observed within this range. Although 600 s yields lower random noise, we retain 5-min averaging as a practical compromise that already provides very low short-term noise while improving operational efficiency and reducing the consumption of both reference and working-standard gases."

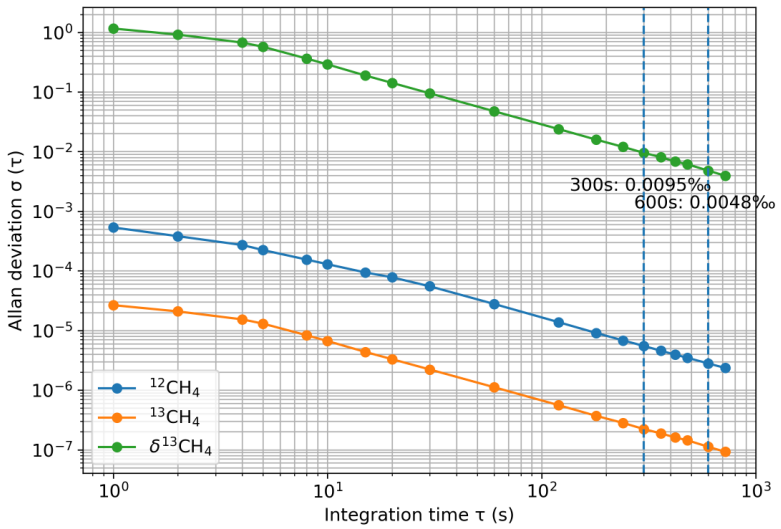

Fig. A2. Allan deviation versus integration time for $^{12}CH_4$, $^{13}CH_4$, and $\delta^{13}CH_4$ from

the Picarro G2201-i, calculated from a stable target-gas segment using 1s data. Vertical lines mark τ = 300 s and 600 s.

- L. 178: "The calibration and correction methods applied to these datasets are described in Section 2.3." Section 2.3 describes two different calibration strategies. Which one was used to calibrate the reference gases ($\delta^{13}CH_4$) against the working standards linked to INSTAAR?

  AR: Thank you for this comment. To clarify the traceability chain, the assigned reference gases (Ref1–3) provide the anchor values (CH_4 on the WMO X2004A scale and $\delta^{13}CH_4$ relative to VPDB). The remaining working standard gases (WS1–WS5) were assigned in our laboratory by calibration against Ref1–3, thereby tying our working scale to the INSTAAR laboratory scale.

  Importantly, the assignment of WS1–WS5 against Ref1–3 was performed using the isotopologue-specific approach. Section 2.3 then describes two alternative strategies (isotopologue-specific vs. delta-based) for applying these assigned working-standard gases to calibrate the laboratory and field datasets, rather than for establishing the assigned values of the working standards themselves.

  We have revised the text around L.178 to make this distinction explicit. Revised text reads:

  "For the assigned reference-gas values (Ref1–3), CH_4 mole fractions were reported on the WMO X2004A scale and $\delta^{13}CH_4$ values were reported relative to the VPDB scale. The remaining working standard gases (WS1–WS5) were assigned in our laboratory by calibration against Ref1–3, thereby tying our working scale to the INSTAAR laboratory scale. The working standard gases were calibrated using the isotopologue-specific approach. Section 2.3 then describes two alternative strategies for applying these assigned values to calibrate and correct the laboratory and field measurements."

- Table 1: It is unclear how the reported true vales for $^{12}CH_4$ and $^{13}CH_4$ were obtained. Please clarify whether these values were measured directly, derived from total CH_4 and $\delta^{13}CH_4$, or taken from assigned reference-gas specifications. Since they are required for isotopologue-specific calibration, their origin should be clearly stated. Furthermore, uncertainties are missing in Table 1.

  AR: This point has been addressed in the revised Table 1 (provided in our response to the major comment). We now state explicitly that the "true" $^{12}CH_4$ and $^{13}CH_4$ values were not measured directly, but were derived from the assigned total CH_4 mole fraction and assigned $\delta^{13}CH_4$ values of the reference gases using standard isotopic relationships. This has been clarified in the Table 1 caption, and the corresponding equations have been added in the Methods.

  We also added uncertainty information to Table 1. For reference gases with manufacturer certificates, the reported uncertainties follow the certificate values. For reference gases without documented certificate uncertainties, we report 1σ repeatability estimated from repeated measurements processed with the same workflow as the ambient data.

- Under section 2.3.1, Eq. 1 and Eq. 2: It should be made explicit whether the instrument's own humidity/water correction (e.g., internal manufacturer algorithms)

was applied to the raw data, or whether all corrections are entirely based on the empirical functions developed here. If both corrected and uncorrected raw data exist, indicate which one was used to derive the humidity correction functions and how this choice affects the calibration.

AR: We have clarified this point explicitly in Section 2.3.1 around Eqs. (1) and (2). The G2201-i data logs provide raw (uncorrected) outputs and, for some reported variables, instrument-corrected ("dry") outputs. In this study, all correction and calibration steps were based on the raw (uncorrected) outputs of $^{12}CH_4$, $^{13}CH_4$, and $\delta^{13}CH_4$. All water-vapor corrections were performed offline using the empirical functions developed in this study, rather than using the manufacturer-corrected outputs. Specifically, the water-vapor correction functions were derived from the raw outputs and then applied consistently to obtain water-corrected $^{12}CH_4$, $^{13}CH_4$, and $\delta^{13}CH_4$ values prior to calibration, ensuring a uniform humidity-treatment workflow and avoiding potential double correction. The added paragraph in the revised version reads as follows:

"In this study, all water-vapor correction functions were derived from the raw (uncorrected) outputs. All subsequent analyses were performed using raw $^{12}CH_4$, $^{13}CH_4$, and $\delta^{13}CH_4$ values corrected offline with the empirical water-vapor correction functions developed here. Manufacturer-corrected outputs were not used in the calibration workflow to avoid potential double correction."

- L. 221-222: "An assumed δD value of -100‰ for atmospheric CH_4 was adopted from Quay et al. (1999)." It is unclear whether this value was applied generally for atmospheric CH_4 or whether the actual δD values of the reference gases used for calibration were known and applied.

  AR: The δD-CH_4 values of the reference cylinders were not available, so we did not apply reference-gas-specific δD values. Instead, we used a single representative atmospheric δD value (−100‰, from Quay et al., 1999) as a common assumption for all conversions involving R_{sum} (following Griffith, 2018). This assumption does not affect the calculation of $\delta^{13}CH_4$, which is derived from the $^{13}CH_4/^{12}CH_4$ ratio. It only affects the conversion from isotopologue mole fractions to total CH_4 via R_{sum}.

  To evaluate the impact of this assumption, we performed a sensitivity analysis by varying the assumed δD–CH_4 over a representative atmospheric range (−50‰ to −150‰). The resulting change in R_{sum} is negligible (≈0.003%), corresponding to <0.1 ppb in derived total CH_4 at typical ambient levels (2000–3000 ppb), which is well below our measurement uncertainty.

  The added text in the revised manuscript reads:

  "A sensitivity estimate was conducted to assess the impact of uncertainty in the assumed atmospheric δD–CH_4 on R_{sum} and total CH_4 mole fraction. Taking δD–CH_4 = −100‰ as a representative value, a conservative ±50% variation (−50‰ to −150‰) changes ^2r by ±5.6%, but propagates to R_{sum} in Eq. (6) at only ±0.0031%. This corresponds to <0.1 ppb uncertainty in total CH_4 for typical atmospheric mole fractions of 2000–3000 ppb, which is negligible relative to measurement uncertainty. Note that R_{sum} is used only to convert $^{12}CH_4$ to total CH_4 and does not enter the calculation of $\delta^{13}CH_4$, which is determined from the ratio $^{13}CH_4/^{12}CH_4$."

- L. 365: The delta-based calibration after water correction is almost identical to the

calibration before water correction. However, the isotopologue-specific calibration shows a substantial change. Could you please explain why the delta-based approach is largely insensitive to the applied water correction here? Could this be due to cancellation effects, the amount of water vapor in the target gas or the specific form of the correction function? Could you also provide the water concentration in the humidified target gas?

AR: The delta-based calibration shows a much smaller response to the water-vapor correction in this case because $\delta^{13}CH_4$ is calibrated as a ratio-based quantity. A substantial fraction of the H_2O interference affects the fitted $^{12}CH_4$ and $^{13}CH_4$ signals in a similar way (common-mode), so part of the effect cancels in the $^{13}CH_4/^{12}CH_4$ ratio and in the subsequent delta-based calibration. In contrast, the isotopologue-specific approach uses absolute $^{12}CH_4$ and $^{13}CH_4$ values that are corrected separately for H_2O interference before reconstructing $\delta^{13}CH_4$. Small differences in the humidity dependence of the two isotopologues, together with the form of the isotopologue correction functions, therefore produce a more pronounced change.

The water vapor level during the humid-air observations at the rice paddy site ranged from 0.93% to 3.5%, as shown in Fig. 7b and reported in Table 2. We have added explicit references to Table 2 and Fig. 7b in Section 3.3 to make the humidified target-gas water-vapor range easier to locate.

- L. 415: The statement should be emphasized, as Fig. 7b shows a clearly stronger correlation between CH_4 mole fraction and $\delta^{13}CH_4$ than between H_2O and $\delta^{13}CH_4$. This could be underlined more explicitly in the text and, if possible, supported by correlation coefficients.

  AR: Thank you for this suggestion. We first reprocessed the Jurong dataset and updated Fig. 7 using the revised calibration workflow (see updated Fig. 7). Based on the updated results, we strengthened the statement around L.415 and now support it with correlation statistics (new Fig. A1). The updated analysis shows that the inter-method difference $\Delta\delta^{13}CH_4$ has only a weak and not statistically significant relationship with CH_4 ($R^2 \approx 0.30$, $p > 0.05$; Fig. A1a), whereas it is strongly correlated with H_2O ($R^2 \approx 0.91$, $p < 0.001$; Fig. A1b). In addition, CH_4–H_2O covariance over the season is relatively weak ($R^2 \approx 0.14$, $p < 0.05$; Fig. 7b). Consistently, isotopologue-specific $\delta^{13}CH_4$ shows no significant dependence on H_2O, while the delta-based $\delta^{13}CH_4$ remains significantly correlated with H_2O (Fig. A1c–d). These updated results indicate that the dominant field behavior at Jurong is a humidity-related residual primarily associated with the delta-based calibration; any concentration-related effect is weak in the field data but is consistent with the intrinsic CH_4 dependence reported for delta-based CRDS calibration (Rella et al., 2015; Miles et al., 2018).

  L. 413-420: The manuscript shows a significant difference between calibration strategies at the humid site, including a clear correlation of $\Delta\delta^{13}CH_4$ with $1/CH_4$. The authors correctly note that part of the observed $\Delta\delta^{13}CH_4$-H_2O relationship may arise from covariance between humidity and CH_4 during high-emission episodes at the Jurong site, rather than from a direct spectroscopic effect of water vapor alone. However, it would be useful to further clarify whether the remaining concentration dependence primarily reflects intrinsic CH_4 non-linearity of the instrument (as reported in previous studies), residual spectral interferences (e.g. pressure or temperature effects), or limitations of the empirical correction approach. A clearer separation between true humidity-driven effects and concentration-related

dependency would strengthen the interpretation of the field results and their implications for source attribution.

AR: Thank you for this comment. We have reprocessed the Jurong dataset and revised Fig.7. We now support the interpretation with the updated Fig. 7 and the added correlation analysis in Fig. A1. In the updated results, the inter-method difference ($\Delta\delta^{13}CH_4$) shows only a weak relationship with CH_4 mole fraction ($R^2 \approx 0.30$) and is not statistically significant ($p > 0.05$), whereas it correlates strongly with H_2O ($R^2 \approx 0.91$, $p < 0.001$). In addition, CH_4–H_2O covariance over the season is comparatively weak ($R^2 \approx 0.14$), indicating that H_2O is not simply a proxy for CH_4 and that the apparent $\Delta\delta^{13}CH_4$–H_2O relationship cannot be explained solely by covariance.

Consistent with this, the isotopologue-specific calibrated $\delta^{13}CH_4$ shows no significant dependence on H_2O, indicating that the isotopologue-level humidity correction is effective, while the delta-based calibrated $\delta^{13}CH_4$ remains significantly correlated with H_2O, suggesting residual sensitivity of the delta-based workflow under humid field conditions. Regarding the remaining concentration-related behavior, the field dataset alone does not robustly constrain its magnitude; however, it is consistent with previously reported intrinsic concentration dependence and non-linearity in delta-based CRDS calibrations (Rella et al., 2015; Miles et al., 2018). We therefore revised the text to avoid over-interpreting the field correlations as a purely spectroscopic water-vapor effect, and to state more explicitly that the residual bias at Jurong most likely reflects limitations of the delta-based calibration under humid conditions, potentially compounded by intrinsic concentration dependence, while additional controlled laboratory dilution tests would be required to fully isolate instrument non-linearity from residual spectral/thermodynamic interferences and empirical-correction limitations.

- L. 480: Please clarify whether an offset of 0.29 ‰ is considered “minor” in the context of $\delta^{13}CH_4$ measurement goals and typical instrumental uncertainty. A brief justification would help readers assess the practical significance of this bias.

  AR: The WMO/GAW compatibility goals for $\delta^{13}C$-CH_4 are very stringent for well-mixed background air (±0.02‰), while an extended goal of ±0.2‰ (k = 2) is commonly used for large-signal, source-influenced conditions. We agree that the term “minor” was ambiguous. We now quantify the practical significance of the 0.29‰ offset relative to established network goals and to our target-gas performance.

  In our humidified target-gas tests, the residual bias of ~0.29‰ is therefore small but not negligible: it is close to the extended ±0.2‰ benchmark (exceeding it by ~0.09‰) and is comparable to the measurement repeatability ($\sigma \approx 0.26$‰). We have revised the text accordingly and avoid describing this offset as “minor.”

  “…a bias of 0.29‰, which is small but non-negligible relative to the extended compatibility goal (±0.2‰, k = 2) and comparable to our target-gas repeatability.”

- L. 470-500 (Discussion): The authors should discuss limitations of their calibration approach regarding different instruments. Are the empirical correction functions transferable, or would they require recalibration for each instrument and site?

  AR: The empirical water-vapor correction functions and calibration parameters reported in this study were derived for the same Picarro G2201-i analyzer and therefore should not be assumed to be directly transferable across instruments or

deployments. Instrument-to-instrument differences (e.g., optical alignment, detector response, and aging) can lead to different humidity sensitivities and concentration dependences, so the empirical coefficients generally require re-derivation (recalibration) for each analyzer and should be validated over the humidity and CH_4 ranges relevant to each field campaign.

At the same time, the function and workflow are transferable that the same correction framework can be applied to other instruments provided that instrument-specific coefficients are determined using controlled humidity tests and routine reference and target-gas checks. We added a brief discussion in Section 4 to state this limitation and to outline recommended validation practices for new instruments and deployments. We added a brief discussion to clarify this limitation. The added text reads:

"The empirical water-vapor correction functions and calibration parameters reported here were derived for a single Picarro G2201-i analyzer and therefore should not be assumed to be directly transferable across instruments or deployments. Instrument-to-instrument differences (e.g., optical alignment, detector response, and aging) may lead to different humidity sensitivities and concentration dependences, so the empirical coefficients generally need to be re-derived and validated for each analyzer and field campaign over the relevant H_2O and CH_4 ranges. While the function and workflow are transferable, applying this framework to other instruments requires instrument-specific characterization using controlled humidity tests and routine reference-/target-gas checks."

Technical Corrections/ Suggestions

- In the abstract and early text, avoid introducing symbols before definitions (e.g., $\Delta\delta^{13}CH_4$ should be defined at first mention).

  AR: We have carefully checked the manuscript to ensure that all abbreviations and symbols are defined at their first occurrence (including $\Delta\delta^{13}CH_4$), and we revised the text accordingly to avoid introducing any symbol before its definition.

  "…… significant inter-method difference ($\Delta\delta^{13}CH_4$) was observed, with which exhibiting a strong correlation with $1/CH_4$……"

- The symbol for methane isotopologues appears incorrectly spaced/formatted in some places. Ensure that isotopologue labels such as $^{12}CH_4$ and $^{13}CH_4$ are formatted consistently throughout (i.e. L. 279 and L. 473). Same with $\delta^{13}C\text{-}CH_4$ and $\delta^{13}CH_4$, please decide for one notation style.

  AR: We have standardized the notation throughout the manuscript. Specifically, methane isotopologues are now consistently formatted as $^{12}CH_4$ and $^{13}CH_4$ (with 12/13 as superscripts and 4 as a subscript), and we use a single isotopic-notation style, $\delta^{13}CH_4$, consistently across the text, figures, and tables. At first mention, we clarify that $\delta^{13}CH_4$ denotes the carbon isotopic composition of methane.

- In 2.1, the manuscript should explicitly state whether all measurements were made with one instrument (Picarro G2201-i) or if multiple analyzers (same model but different instruments) were used (field vs. lab). A reviewer comment already highlights this as missing.

  AR: We have clarified in Section 2.1 that all laboratory experiments and field measurements were performed using the same Picarro G2201-i analyzer.

"All laboratory experiments and field measurements reported here were performed using the same G2201-i analyzer, although the field measurements were performed during different campaigns in different years (2018, 2022, and 2025)."

- L. 28: Correct to "SORPES site".

  AR: We have corrected "SORPES Stie/Site" to "SORPES site" and standardized the capitalization of "site" throughout the manuscript (e.g., "the SORPES site" and "the Jurong site") for consistency.

- L. 135: "$^{12}CH_4$" instead of "$^{2}CH_4$".

  AR: We have corrected this typo and now use $^{12}CH_4$ consistently.

- L. 144: Complete location information with height a.s.l. also for SORPES site.

  AR: We have completed the site descriptions by adding elevation (m a.s.l.) information for both sites to ensure consistent reporting. Specifically, we report the SORPES site elevation as ~40 m a.s.l., and we added the Jurong site elevation as ~15 m a.s.l. based on the site description in Dai et al. (2019).

- L. 158: "Ref4-Ref5" instead of "Ref5-Ref6"

  AR: We have corrected this as suggested.

- L. 162: Update Figure reference to Fig. 1b-c, as no Fig. 1d is available.

  AR: We have corrected the figure reference to Fig. 1b–c.

- L. 179: Subscripts in Table 1 are not properly formatted (low/high positioning is incorrect).

  AR: We have corrected the subscript formatting in the Table 1 title to ensure that all subscripts are properly rendered and consistently positioned.

- L. 182: set "CH3D" to "CH_3D"

  AR: We have corrected this as suggested and now use "CH_3D" consistently.

- Eq. 1 and Eq.2: I think subscript should be "true" instead of "ture"

  AR: Thank you for noticing this typo. We corrected the subscript from "ture" to "true" in Eq. (1) and Eq. (2).

- L. 295: Correct "vapor correction" for consistent terminology

  AR: We have revised the wording at L. 295 to ensure consistent terminology and now use "water vapor correction" consistently throughout the manuscript.

Recommendation

Overall, this manuscript makes a meaningful contribution to the further development of calibration strategies for $\delta^{13}CH_4$ measurements under variable ambient conditions. With clarifications to the calibration methodology and expanded uncertainty characterization, as well as clearer justification of functional corrections and discussion of temporal stability and transferability, as well as conducting and reporting additional laboratory tests the manuscript would be suitable for publication in AMT. I recommend publication in AMT after major revisions addressing the points outlined above.

We thank the reviewer for the constructive overall assessment. We have revised the manuscript extensively in response to your comments. Specifically, we clarified the calibration methodology, notation, and scale anchoring; strengthened the justification and interpretation of the correction functions; expanded the uncertainty characterization; and added discussion on temporal stability and transferability. In addition, we reprocessed and updated the field analyses, including a revised Jurong processing workflow and supplementary correlation analyses, to better distinguish humidity-linked effects from concentration-dependent behavior under humid conditions. We hope that with these revisions, the concerns have been addressed.