

Dear Editors,

In light of the reviewer comments, we have reorganized the manuscript to be organized by topic rather than having separate methods and discussion sections. Thus back and forth references have been dramatically reduced. We also simplified the manuscript by eliminating discussion of calibration methods that were not actually used. We made clear recommendations for similar future networks, e.g., the isotopic ratio of the field tanks, and the lab and field tank sampling times. Our responses to the reviews are in-line, in blue below.

Thank you,

Natasha Miles

General Comments

This manuscript discusses a specific question within the scope of AMT, presenting new data on the difficult task of calibrating laser methane isotope instruments at unmanned tower sites.

The methods are outlined in extensive detail, but this also highlights how a little more consideration of the isotopic signatures of influencing methane sources at the start of the project could have saved significant time spent later refining calibration routines. The choice of enriched calibration mixtures over depleted ones suggests an expectation of ^{13}C -enriched sources.

Thank you for your thoughtful comments. We shortened the paper by removing most of the discussion of choosing the optimal calibration study. Instead we described how we actually processed the data, rather than describing all the options we explored. We agree that the choices of field tanks and the field tank testing times were not ideal and have added specific recommendations for future similar networks to the paper.

In places the manuscript is quite difficult to follow as there are frequent references to previous or subsequent sections. There are also a lot of small errors that need tidying up, and some points that need greater clarification. This includes the selection procedure of points for the individual peak Keeling plots. Some of the subsequently 'unused' EXPTs could be removed from the text (but remain in the table), as they are not used in the revised calibration routine.

We have addressed these concerns by reorganizing the bulk of the paper. Instead of strictly separating methods and results, we have now inter-mixed these, with the paper being organized by topic. We believe this does enhance readability of the paper. While not the traditional method of manuscript organization, it is sometimes used (e.g., Rella et al. 2015). We also added an example of a time series used for the individual peak Keeling plots, as suggested. We removed lengthy discussion of calibration procedures not actually used to process the data.

The manuscript also lack a good final summary, possibly due to the conference volume time constraints, but it would be good to see the following questions discussed at the end: 1) In hindsight what could have been done differently? 2) What are the recommendations for anyone else undertaking the set-up of a similar network? 3) What are the limitations and advantages of the CRDS technique at fixed tower sites compared with IRMS at a similar type of site (e.g. Rockmann et al., 2016)?

We have reworked the final section to include recommendations for future similar deployments and have mentioned the potential of high-temporal-resolution methane isotopic ratio data, when combined with modeling, to constrain regional methane budgets. Previously the utility of networks of such data was not mentioned.

The manuscript also highlights the need for suitable isotopic standards for urban / source region standards within this community, as the measured isotopic ratios fall between those commonly measured at source, or the tight constraints around -47 ‰ needed for background sites. Something like -52 and -42 ‰ at 3 ppm mole fraction are toward the limits of the enhancements measured at such sites, and so would be very useful as standards.

We included recommendations for future work with these thoughts in mind.

Detailed Comments

Abstract (see also later comments)

Page 1 Line 24 – Why only calibration with high methane mole fraction air bottles?

We edited that sentence to read, “Prior to deployment, each analyzer was tested using bottles with various isotopic ratios, from biogenic to thermogenic source values, which were diluted to varying degrees in zero air, and an initial calibration was performed.” We were referring to the Isometric Instruments bottles (as high methane mole fraction air bottles). We did only use the high mole fraction mixture in the initial calibration prior to deployment, because the field calibrations supersede the correction of the mole fraction dependence of the isotopic ratio (and because we didn’t measure the low mole fraction mixtures long enough). Those details are in the text, but are too detailed to describe in the abstract. Thank you for pointing out the confusing statement in the abstract.

Page 2 Line 2-5 - This technique might work here because the Marcellus gases are significantly enriched in ^{13}C , but in many regions (eg. Australia, much of the EC) there are no major sources enriched in ^{13}C relative to atmospheric background, so that the sources would be very difficult to distinguish far away from the point of emission.

Good point. We added a sentence to the conclusions, “We note that the Keeling plot approach to determine source isotopic signatures far from the point of emission will be difficult to apply in regions without sources that are significantly depleted or enriched in $^{13}\text{CH}_4$ compared to ambient.”

Page 2 Line 4 – What is the error on the -31.2 ‰ The literature suggests a range of source values.

We added the standard deviation of values that we determined and edited the last phrase to read, “ within the wide range of values consistent with a deep-layer Marcellus natural gas source. “

Main Text

Page 3 Line 18 – t missing from Schweizke

Corrected.

Page 5 Lines 16-22 – yes it should be symmetrical at 35 and -60 ‰ around a background of -47.5 ‰ but the colours on Fig 1 don’t show an even change to enriched and depleted values around the background composition (see also comment on Figure).

Yes, thank you for pointing out this error. We have corrected the calculations for the figure to reflect -35 and -60‰, with a background of -47.5‰ used. We added further description of the end members to the caption, as requested.

Page 6 Line 3 – the noise is known to be less for higher mole fractions – please cite the source of this

We added the reference for this (Rella et al., 2015).

Page 6 Line 24 – slight clarification needed – was 1 sccm for the standard line and 500 sccm for the zero air line

Yes, and we have added that clarification.

Page 7 Lines 1-7 – it would be useful here to include the flow rate through the instrument to give some idea of how much cal gas was being used.

We noted in this section: “With the flow rate of 0.400 sccm for the isotopic standard bottles, the total volume of standard gas used was 88 cc.” Here we focused on the standard bottles rather than the zero air, since the zero air is inexpensive. We also added in Section 4.3 (In-situ field calibration gas system), “The flow rate of the instruments was 35 cc/min, and the 150A tank size was used, corresponding to 4.021×10^6 cc at standard pressure and temperature. Thus there was gas sufficient for 9800 calibration cycles, or 725 days at this calibration frequency.”

Page 7 Lines 19-32 – some of the terms are not clearly explained here (for example Bdefault) so please either clarify or remove and refer the reader directly to Rella et al. (2015).

We added a description of the meaning of the terms B and Bdefault.

As it would be good to be able to refer to the calibration plot at this point to see the isotopic offset compared to known values for each standard at different mole fractions can Fig.7 go here, or be linked to section 4.2 or at least referred to?

With our reorganization of the paper, the figure is included in the mentioned section.

Page 8 Lines 3-10 – The study area section and accompanying figures seem a little out of place here as the sites are not discussed for a long time. Is there a better position for this?

We moved the section describing the study area (including the figures originally numbered 3 and 4) to just prior to the section describing “Methods for determining enhancements”.

Page 8 Line 16 – not been demonstrated for laser instruments, for IRMS see Rockmann et al.(2016).

We added “CRDS” to the sentence.

Page 8 Line 25 – using the word ‘sampling’ for the measurement of standard cylinders is easy to confuse with the actual measurement of ambient air at the towers. Is there an alternative word that could be used for this procedure?

We changed this to ‘testing’ throughout the text.

Page 8 Lines 26-27 – the reasoning behind choosing the mole fractions for tanks with these isotopic standards is not clear here, and is even less clear later when the -24 ‰ cylinder is dropped from the calibration routine.

Agreed. In retrospect that is clear. We added, “We note that it would have been preferable to utilize calibration tanks closer to the observed air samples in terms of isotopic ratio. In particular, the high tank could have been spiked with the -38.3 ‰ bottle, and/or both the high and the low tanks could have been spiked with a mixture of the -38.3 and -54.5 ‰ bottles.”

Page 9 Lines 11-13 – what is a large flow rate and what is the delay time for air entering the inlet to instrument measurement?

We added, “For the CRDS analyzer, air was drawn down the tube at 1 L/min, with 30 cc/min flow into the analyzer and the remainder purged. The residence time in the tube was about 1 min. Separate tubes were used for the CRDS and flask sampling lines because of the differing flow rates (varying between 0.29 and 3.8 liters per minute) required for the flask samples (Turnbull et. al., 2012) and to ensure independence of the CRDS and flask measurements.”

Page 9 Lines 28-30 – should the aim not be to achieve the best possible precision, not just to reach target compatibility? Would the high mole fraction tanks not achieve better precision if analysed for 32 minutes? If the shorter times are chosen just to save cal gas then it should be made clear somewhere.

We added to that paragraph, “The ideal calibration tank testing time is a balance between minimizing calibration gas usage (and consequently maximizing ambient air sampling time) and achieving sufficient precision.”

Page 10 Lines 18-24 – the -24 ‰ standard at any concentration was never going to be useful at the tower sites as it is even beyond the range of sampling a pure source emission. A clear understanding of the maximum mole fractions and isotopic shifts observed at the towers would have been useful before production of the working standards for these sites, even if it meant not using the first 6 months of isotope data.

Page 13 Lines 26-27 – what are the uncertainties on the NOAA tanks?

Uncertainties have been added to the text.

Page 14 Lines 8-10 – replace tubing with tube, and on-board with internal

Done.

Page 14 Section 5.3 – what is the relevance of this section? Unless the results of this test are shown somewhere, and referred to here, then is this section necessary for the points under discussion?

We added clarification that although not central to the primary results of this project, the performance of the analyzers is important if the data are to be used as part of the continental-scale CO₂ network. The results are shown in Section 6.6.

Page 14 Line 16 – 14:00-15:00 isn't late afternoon

Replaced with “afternoon.”

Page 14 Line 18 – do you mean measured? Either ‘samples were collected’ or ‘flasks were filled’

Replaced with “samples were collected.”

Page 15 Lines 9-10 – ‘An error in isotopic ratio as a function of isotopic ratio’ is not very informative. Figure 7 shows that there is an offset in the measured isotopic ratio as a function of the changing known isotopic ratio, which seems to be quite constant for all instruments at higher mole fractions, but instrument-specific at near-background mole fractions.

We added this clarification to the text.

Section 6.2 – I initially thought that the experiments involved changing over to different cylinders, which seems to be the case when changing to the new calibration routine, but most of these EXPTs seem to be just manipulation of the data for different standards, so please can you make clear at the start of this section that these are mostly changes in the calibration calculations and not changes to the cylinders being analysed. Some of the results of these experiments are not used as they do not improve the required precisions. Do these need to be described in the text?

This section was confusing and we have eliminated all but two schemes (using the target as independent and using the low as independent). We eliminated Table 3 from the original document. And we tried to clearly state when hindsight (and insightful reviewer comments) have revealed things that we would do differently.

Page 16 Line 13 – a result is a result and cannot be changed. Do you mean ‘improve the calculated precision’?

Changed to “improve the calculated accuracy.” Also replaced in three other locations in the text.

Page 18 Line 16 – 0.18 ‰ is stated above this as the daily average, not the hourly average as used here.

We clarified that section to indicate more clearly that averaging over the low tank for each day totaled about 1 hour of data (actually 81 min), and thus the standard deviation of these values is a proxy for the noise due to the calibration scheme in the hourly sample air data.

Section 6.6 – the side-by-side testing results are mentioned here, but have a lot of content overlap with the methodology for side-by-side testing. Does this have to be in two places? It isn't particularly relevant to the core isotopic story.

We were careful to keep the methods and results separate. We added clarification in Section 5.3 that although not central to the primary results of this project, the performance of the analyzers is important if the data are to be used as part of the continental-scale CO₂ network.

Page 20 Line 2 – ‘mean flask’ – is this referring to the rapidly-filled or the hour-long filled flasks?

This was specified in the prior section as being the hour-long filled flasks.

Page 20 Line 5 – ‘hourly flask to in-situ differences for the year’. It isn’t clear what this means

This was repetitive with the prior section and makes it sound more complicated that it is. We removed it.

Page 20 Line 19 – ‘The time scale of the individual data points was 10 min’. Do you mean averaging interval?

Yes, corrected.

Page 20 Lines 32-33 – It isn’t clear why ‘enhancements greater than 6ppb CH₄ in magnitude’ are ‘3 times the target compatibility of 0.2 ‰

Typo corrected.

Page 21 Line 15 – mean of -31.2± ? ‰

-31.2 ± 1.9 ‰. Added to text.

Page 21 Lines 21-23 – this section needs sorting out; firstly they are not peak heights but enhancements over background, and it is a 2.5 to 8.7 ‰ positive shift in measured isotopic ratio. What does ‘reduced methane enhancement at other data points within the peak’ mean? If the maximum enhancement is 2008 ppb, why does it mention 1500 ppb maximum earlier?

We made these changes and removed the confusing phrase as it is too obvious. The maximum enhancement during non-afternoon hours at the central tower was 2008 ppb. The 1500 ppb mentioned refers to afternoon hours at the North tower. We added clarifications.

Page 22 Line 21 – remove the s from fractions

Done.

Page 22 Line 26 – replace improving with improved

Done.

Page 22 Lines 35 to 37 – given the availability of the Isometric flasks, my preference for a calibration would be to create a high and low mole fraction cylinder from both of the -54 and -38 ‰ standards, used in combination with a low ambient cylinder and a similar target gas, or at least a high and low at -38 ‰ as that is the direction your sources are taking the ambient mix. More cylinders, but it should improve the correction in the triangle of ¹³C-mole fraction space where the measurements lie.

We have incorporated these ideas into Table 4 of the revised manuscript, for recommendations for future tower networks of CRDS isotopic methane analyzers.

Page 23 Line 30 – an isotopic ratio enhancement of -0.6 ‰ Two problems with this: surely this should be a + and not a -, and how do you enhance a ratio? Normally this would be heavy or light for a change in ratio as ¹³C increases or decreases, or enrichment or depletion if talking about the individual ¹³C, so it does represent an enrichment in ¹³C.

It seems that “difference from background isotopic ratio” is a better term than “enhancement above background” – I can see how the latter is confusing. I switched the terminology throughout the text. If the source in question is -35 ‰ and background is -47 ‰ (as in the example in the text), the measured isotopic ratio would be lower than background, so I think it is negative. I’m assuming it was just the term “enhancement” that was causing the problem.

Page 24 Line 15 – already presented so should be ‘we have presented.

Done.

Page 24 Lines 15-19 – this is a rather abrupt ending – see general comments for suggestions of what to add in summation.

References

All those present seem to be correct in the text, just move the date on P27 Line 7 up a line.

Done.

The following references are mentioned in the text but are not in the list: Conway et al., 2011 Montzka et al., 2011 Turnbull et al., 2012 Vaughn et al., 2004

These have been added to the references list.

Tables

Table 2 – Not clear what this table is representing from the caption? Are these due to interferences with the $^{12}\text{CH}_4$ and $^{13}\text{CH}_4$ spectral lines? What is the maximum CH_4 at these sites and does this increase or decrease the interference on each of these species?

We clarified the caption of Table 2, “Maximum error estimate attributable to cross-interference due to direct absorption on $\delta^{13}\text{CH}_4$. These estimates were based on typical values for this tower-based application and estimated effects on CRDS measurements (Rella et al., 2015), and assumed 2 ppm ambient CH_4 mole fraction. For water vapor and carbon dioxide, the interferences are independent of CH_4 mole fraction for 1 – 15 ppm. For the other species listed, the interferences are inversely proportional to CH_4 mole fraction. “

Table 3 caption Line 10 – should be at not and

Done.

Table 5 – the alternative strategy looks to be the best, but given the observed range of measurements up to 4 ppm, would it not be better to correct with a lower precision at 4 ppm than a calculation of the offset from 10 ppm?

This is a good question. The choice of the CH_4 mole fraction of the high tank is based on the optimal determination of the calibration coefficients c_0 and χ , rather than the expected range of ambient CH_4 mole fractions. The effect of the offset parameter c_0 on the calibrated δ is largest at low mole fractions, whereas the effect of the slope parameter χ is independent of mole fraction. Thus the ratio of the high and low tank mole fractions determines how separable the two effects are. We therefore chose the high tank mole fraction to be as high as possible without introducing other nonlinearities into the system. We have added this explanation to the text.

Table 5 - You have all of the Isometric standards available, so 'easier to obtain tanks with these ratios' is not an advantage of the reduced compatibility option, but all of the options.

We have reworked Table 5 (now Table 4) to make it more clear.

Figures

Is there a way to sort out the superscript and subscripts on the vertical axes of the graphs or is this a software limitation?

The superscripts and subscripts have been fixed.

Figure 1 - It needs to be mentioned in the caption what are the signatures of the isotopic end members used to create this plot. There is an uneven spread of colour change around the -47 ‰ point, which suggests that this has been calculated with -30 and -60 ‰ source increments and not the -35 and -60 ‰ mentioned in the text.

Yes, thank you for pointing out this error. We have corrected the calculations for the figure to reflect -35 and -60‰, with a background of -47.5‰ used. We added further description of the end members to the caption, as requested.

Figure 2 caption – mention the cylinder volumes

Added: “At standard pressure and temperature, the gas volume of the zero air and working standard tanks was 4021 L and that of the Isometric Instruments bottles was 28 L.”

Figure 4 – The lowest category of wind direction is 0-2 m/s, but the caption states that calm winds below 3.6 m/s are not categories by direction, so two things to correct here. 3.6 m/s is not calm. I think that this is a scale conversion problem from km/h, and calm should be either less than 1 m/s (in which case the lower category displayed should be 1-2 m/s), or <0.1 m/s.

3.6 m/s was a typo. Further clarification to the caption has been added. “These afternoon means were based on hourly reported measurements. For the hourly measurements, calm winds ($< 1.6 \text{ m s}^{-1}$) were not categorized by direction and thus were not included in the afternoon mean. For the hourly measurements, calm winds ($< 1.6 \text{ m s}^{-1}$) were reported as zero and were included in the afternoon mean.”

It isn't clear why Figs 3 and 4 are so early in the manuscript. Until the tower location data is discussed there isn't really the interest in seeing these influences.

We moved the section describing the study area (including the figures originally numbered 3 and 4) to just prior to the section describing “Methods for determining enhancements”.

Figure 5 – What is the flow rate of the pump on the exhaust line (ME3)? It looks like it should be balancing out the ambient inlet flow.

Yes, the flow rate is about 1 LPM. We added that information to the figure, and also changed the labels of the tanks to be consistent with the text.

Figure 6 caption – incorrect 13C shown on first line

Corrected.

Figure 7 – The residual errors seem to be quite consistent between instruments for the higher mole fraction cylinders at all sites. The correction procedure for the small number of elevated mole fraction samples is not very obvious from the text.

The calibration procedure is described in Section 3.2.2.

Figures 8 and 9 – what are the error bars for these data points? If they cannot be added to the graphs can they be alluded to in the caption?

We added to the caption of Figure 8, “The Allan deviation for time period used for each calibration cycle was, for the period prior to the improved tank sampling strategy, 0.2 ‰ for the high tank, and 0.5 ‰ for the low and target tanks. Following the implementation of the improved tank sampling strategy, the Allan deviation for each calibration cycle was 0.1 ‰ for the high tank, and 0.3 ‰ for the low and target tanks.”, but it is not obvious how to calculate an estimation of the error for Figure 9.

Figure 11 – Can North, Central and East be labeled at the top of each column?

Done for this figure, and for Fig. 7.

Figure 13 – add measured before isotopic in the first line of the caption. The big delta small delta use on the left column horizontal axes is not explained. The $\times 10^4$ used in the right column needs correctly positioning, although it would be much clearer just using 0.2 to 0.6 in $1/\text{CH}_4$ ppm.

Done for this figure, and Figs. 14 and 15.

Figure 14 – the 15 ‰ spread of isotopic ratios measured at near background mole fractions reduces confidence in the data. The text suggests that there is a gradual improvement in measurement precision as the mole fraction increases. The Keeling plots suggest that the transition is quite sharp, and the isotope CRDS that I have seen in field operation seem to have a sharp improvement at 7-8 ppm CH_4 , so can this be clarified in the caption or elsewhere in the text.

Yes the range of values during non-afternoon hours at the North tower is 15 ‰, but the standard deviation is much less, only 0.76 ‰. The size of the markers in the figures makes it look more variable than it is. I reduced the marker size, which helped a little, but I don't think we can eliminate this without making the figures very large. Thus I added the median and standard deviation of the isotopic ratios at each tower to the Figs. 13/14 CFI, (Fig 14/15 in the updated document) in order to clarify the variability. The noise in the isotopic ratio measurement does increase near background mole fractions, but it's about 0.4 ‰ for 10 min averages at 2 ppm CH₄. We added this to the captions for Figs. 13 and 14 and in the text. We added a figure showing the exponential decrease of the standard deviation of the methane isotopic ratio at 2, 3, 6, and 7 ppm.

Figure 15 – the Keeling plots are good but indicate very little variation in measured isotopic ratio for a given mole fraction even at near background mole fractions. Are all of the points from these peaks used? It would be good to see the actual time graph of the isotopes for 1 of these graphs to see the points that have been selected for use in the Keeling plot. The source intercept calculation is without errors. Can these be calculated?

We added a figure to the Methods Section indicating an example of the CH₄ time series for which the Keeling plot approach was applied. The time during which the tower was in the plume of the source was obvious, and only these points were included in the calculation, thus explaining the lack of variability. Added to the text, “Propagating a potential error (attributable of analyzer uncertainty) of 0.2 ‰ at the heavy end of the Keeling plots and –0.2 ‰ at the light end, and vice versa, the potential range of the mean is from –32.0 to –30.4 ‰. “

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2017-364, 2017.

Anonymous Referee #2

Received and published: 7 December 2017

Review of the manuscript: "Calibration and Field Testing of Cavity Ring-Down Laser Spectrometers Measuring CH₄, CO₂, and $\delta^{13}\text{CH}_4$ Deployed on Towers in the Marcellus Shale Region", submitted to Atmos. Meas. Tech., by Natasha Miles et al.

The paper is describing the measurements of atmospheric mole fractions of $\delta^{13}\text{CH}_4$, CH₄ and CO₂ at four sites in Pennsylvania. More precisely the manuscript describes the optimization of the technical setup based on lab and field tests.

The manuscript needs to be reorganized to reduce the back and forth between test descriptions and their results, which makes the reading quite difficult. There are too many redundancies, and unclear statement. When doing that I also suggest to shorten the manuscript. Some conclusions appear obvious, like for example the statement that field calibrations significantly improved the measurements compatibility. Also the so-called optimal calibration strategy refers to the design which was decided a priori and slightly modified during the campaign, but there was no plan to really evaluate alternative design. The conclusion should be written in a more concise way, focusing on the recommendations gained from the experiment.

Thank you for your thoughtful comments. We have addressed these concerns by reorganizing the bulk of the paper. Instead of strictly separating methods and results, we have now inter-mixed these, with the paper being organized by topic. I believe this does enhance readability of the paper. We shortened the paper by removing most of the discussion of choosing the optimal calibration study. As the other reviewer indicated, the isotopic ratio of the low tank (-23.9 per mil) would be better chosen to be closer to the measured values. We also eliminated switching from the target tank being independent to the low tank being independent in the text. Instead we just described how we actually processed the data, rather than describing all the options we explored. We have reworked the conclusions, focusing on the recommendations gained from our experiment and the potential of high-temporal-resolution isotopic methane data to constrain regional methane budgets.

Introduction: the introduction need to be reorder in order. For example the first paragraph of page 4 describing the interest of tower versus aircraft, appears between two paragraphs discussing more technical points about CRDS measurements

We reordered the introduction and added context concerning the utility of high-temporal-resolution isotopic methane data.

Page 4 / Line 13: “three field calibration tanks. . .”: I would rather say two calibration tanks plus one target tank used as quality control and not used in the calibration.

We changed this throughout the text to refer to these tanks as field tanks, since as you mention, one of them is independent of the calibration.

Allan variances tests; calibrations tests (Page 6 / Line 31): there are many back and forth between description of the set up and the results, which confuse the paper.

We have reorganized the paper to eliminate this concern. Instead of strictly separating methods and results, we have now inter-mixed these, with the paper being organized by topic. We believe this does enhance readability of the paper. While not the traditional method of manuscript organization, it is sometimes used (e.g., Rella et al. 2015).

Page 9: In-situ field calibration: is the Nafion required for the setup ? Have you compared possible biases due to the use of the Nafion versus the water vapor correction ? I am not fully convinced by the strategy of humidifying the dry calibration tanks.

Rella et al. (2015) noted that the effect of water vapor on the isotopic ratio of methane is large (up to 1 ‰) and nonlinear. Thus no water vapor correction is available. They recommend drying to less than 0.1% mole fraction. We have added clarification on our reasoning for drying to the text. We also added a reference to Andrews et al. (2014) who document the technique of drying the sample and humidifying the calibration gases.

Page 9: 4 min flushing: how do you estimate those 4 minutes as sufficient for the flushing ?

Added to the text: After this time, the CO₂ and CH₄ mole fractions have stabilized.

Page 12: background site: why don't you select the background site as a function of wind direction rather than picking up one site for the full period ?

We added to this discussion. It now reads, “ The predominant wind direction for the Marcellus region is from the west (Fig. 4). For westerly winds, the South tower is a reasonable choice for a background tower. The South tower measured the lowest overall mean afternoon methane mole fraction (1960.2 ppb CH₄). The mean afternoon methane mole fractions of the other towers, averaged only when data for the South tower exist, were 8.7, 7.0, and 2.9 ppb higher, at the North, Central, and East towers, respectively. For future analyzes, a wind direction dependent background tower (South or North) could be used, but the North tower did have the largest mean enhancement in CH₄ mole fraction compared to the South tower. “

Page 14: Allan results: For CH₄ and CO₂ it should be noted that the results seem to be not as good as the performances obtained with G1301/G2401 analyzers. Do you know the reason which could explain a difference of the performances between those analyzers ?

We have added further clarification to the Allan deviation results section and the side-by-side testing section. The performance of the G2132-i analyzers in terms of CO₂ precision is worse than that of the G2301/G2401 analyzers primarily because a weaker spectral line is used. Whereas the spectral line for CH₄ is the same between the two types of analyzers, for CO₂, the absorbance of the spectral line used in the G2132-i analyzers is a factor of 11 times less, meaning the precision is dramatically reduced.

Page 16: Calibration scheme: the presentation of the different tests should probably be shortened. Is there a difference between Expt E and H designs ?

We have eliminated most of the EXPTs in order to simplify. We eliminated Table 3 and shortened Table 4 from the original document. EXPTs E and H are no longer described.

I would appreciate an evaluation of the optimum frequency of the field calibration sequences (intermediate between 0 and once per day). From the variabilities shown on Fig.8 and 9 it looks like a reduction of the calibration frequency to once every few days would not affect by much the measurements.

We added to the text, ‘Considering shorter term changes, the day to day changes in the calibration were less than 0.5 ‰ for December 2016. Less frequent calibrations, e.g., twice per week, could be considered, but the reduction in field tank use is not

large considering the low flow rates of the instruments and steady changes up to 2 ‰ in the raw data over the time scale of days were observed in Rella et al. (2015). ‘

Fig.8: the legend is misleading since the so-called target tank is used as a calibration tank. To make it clear you should add comments in the legend of each figure (e.g. Target tank (used as CAL))

We added to the caption, “The target tank was used in the isotopic ratio calibration, whereas the low tank was independent.”

Page 17: Fig. 9B and 9C should rather be 9A and 9B

Corrected.

Page 20 Line 27: suppress ‘For the daily afternoon averages,’. Not clear what you mean by a ‘reduction’ of 0.6-0.7pmil.

We have clarified these confusing statements. The text now reads, “The standard deviation of the daily afternoon averages (rather than 10-min averages) was 0.6 – 0.7 ‰. Thus the observed width of the distribution appears to be persistent throughout the afternoon and not merely measurement noise. “

Page 20 Line 32: Why do you compare CH₄ enhancements (6ppb) with ¹³CH₄ target compatibility (0.2pmil) ?

Typo corrected.

Page 21: lines 22/23: Unclear statement about the dilution of local source.

We have removed this statement.

Page 21: lines 22/23: The discussion about the source signature need to be clarified, or preferably merged in the discussion section.

Hopefully having the paper organized by topic, rather than having the method for each topic separated from the results, makes this discussion more clear. We also added a figure of an example time series of a peak for which we applied this method.

Page 23: lines21/22: unclear statement.

A misplaced parenthesis made this statement confusing. Corrected.

Conclusion: in your last sentence I would like to see also a comment or discussion that the strategy of using continuous measurements at four tower is maybe not the optimal one for the quantification of such sources.

We added, ‘ For determination of the source signature for a specific known location, the tower-based approach is not ideal. Instead the strength of the tower-based approach lies in covering larger areas and many potential source locations, and for longer periods of time than is feasible by other approaches, e.g., short-term mobile techniques. ‘ We also added to the last section discussion about the utility of high-temporal-resolution methane isotopic ratio data for constraining regional methane budgets.

Please also note the supplement to this comment: <https://www.atmos-meas-tech-discuss.net/amt-2017-364/amt-2017-364-RC2-supplement.pdf>

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2017-364, 2017.

Calibration and Field Testing of Cavity Ring-Down Laser Spectrometers Measuring CH₄, CO₂, and δ¹³CH₄ Deployed on Towers in the Marcellus Shale Region

Natasha L. Miles¹, Douglas K. Martins^{1,2}, Scott J. Richardson¹, Christopher W. Rella³, Caleb Arata^{3,4}, Thomas Lauvaux¹, Kenneth J. Davis¹, Zachary R. Barkley¹, Kathryn McKain⁵, Colm Sweeney⁵

¹Department of Meteorology and Atmospheric Science, The Pennsylvania State University, University Park, Pennsylvania, 16802, USA

²FLIR Systems, Inc, West Lafayette, Indiana, 47906, USA (current affiliation)

³Picarro, Inc., Santa Clara, California, 95054, USA

⁴University of California, Berkeley, California, 94720, USA

⁵National Oceanic and Atmospheric Administration / University of Colorado, Boulder, Colorado, 80305, USA

Correspondence to: N. L. Miles (nmiles@psu.edu)

Abstract. Four in-situ cavity ring-down spectrometers (G2132-i, Picarro, Inc.) measuring methane dry mole fraction (CH₄), carbon dioxide dry mole fraction (CO₂) and the isotopic ratio of methane (δ¹³CH₄) were deployed at four towers in the Marcellus Shale natural gas extraction region of Pennsylvania. In this paper, we describe laboratory and field calibration of the analyzers for tower-based applications, and characterize their performance in the field for the period January – December 2016. Prior to deployment, each analyzer was tested using bottles with various isotopic ratios, from biogenic to thermogenic source values, which were diluted to varying degrees in zero air, and an initial calibration was performed. Furthermore, at each tower location, three field tanks were employed, from ambient to high mole fractions, with various isotopic ratios. Two of these tanks were used to calibrate the analyzers on a daily basis. A method to correct for cross interference from ethane is also described. Using an independent field tank for evaluation, the standard deviation of 4-hour means of the isotopic ratio of methane difference from the known value was found to be 0.26 ‰ δ¹³CH₄. Following improvements in the field tank testing scheme, the standard deviation of 4-hour means was 0.11 ‰, well within the target compatibility of 0.2 ‰. Round robin style testing using tanks with near ambient isotopic ratios indicated mean errors of −0.14 to 0.03 ‰ for each of the analyzers. Flask to in-situ comparisons showed mean differences over the year of 0.02 and 0.08 ‰, for the East and South towers, respectively.

Regional sources in this region were difficult to differentiate from strong perturbations in the background. During the afternoon hours, the median differences of the isotopic ratio measured at three of the towers, compared to the background tower, were −0.15 to 0.12 ‰ with standard deviations of the 10-min isotopic ratio differences of 0.8 ‰. In terms of source attribution, analyzer compatibility of 0.2 ‰ δ¹³CH₄ affords the ability to distinguish a 50 ppb CH₄ peak from a biogenic source (at −60 ‰, for example) from one originating from a thermogenic source (−35 ‰), with the exact value dependent upon the source isotopic ratios. Using a Keeling plot approach for the non-afternoon data

Deleted: The calibration of the continuous isotopic methane analyzers used in this study required both a linear calibration and a mole fraction correction, and a correction for cross-interference from ethane.

Deleted: calibrated using high methane mole fraction air

Deleted: .

Deleted: calibration

Deleted: By testing multiple calibration schemes, we determined an optimized field calibration method.

Deleted: calibration

Deleted: sampling

Deleted: enhancements

Deleted: enhancements

1 at a tower in the center of the study region, we determined the source isotopic signature to be $-31.2 \pm 1.9\text{‰}$, within
2 the wide range of values consistent with a deep-layer Marcellus natural gas source.
3
4

Deleted: ‰,

1 Introduction

Quantification of regional greenhouse gas emissions resulting from natural gas extraction activities is critical for determining the climate effects of natural gas usage compared to coal or oil. Studies have shown that the emission rates as a percentage of production vary significantly from reservoir to reservoir. An aircraft-based mass balance study in the Uintah basin in Utah (Karion et al., 2013; Rella et al., 2015) found a methane emission rate of 6.2–11.7 % of production, exceeding the 3.2 % threshold for natural gas climate benefits compared to coal determined by Alvarez et al. (2012). In the Denver-Julesburg basin in Colorado, Pétron et al. (2014) found an emissions rate of 4 % of production, again using an aircraft mass balance approach. The Barnett Shale, one of the largest production basins in the United States with 8 % of total U.S. natural gas production, was found to exhibit a lower emission rate of 1.3–1.9 % (Karion et al., 2015). Using a model optimization approach for aircraft data, Barkley et al. (2017) found the weighted mean emission rate from unconventional natural gas production and gathering facilities in the Marcellus region in northeastern Pennsylvania, a region with mostly dry natural gas, to be only 0.36 % of total gas production.

~~Aircraft-based studies cover large areas, but the temporal coverage is limited. Tower-based networks offer a complementary approach, making continuous measurements over long periods of time. At the Boulder Atmospheric Observatory (BAO) tall tower, daily flask measurements are found to contain enhanced levels of methane and other alkanes, compared to the other tall towers in the National Oceanic and Atmospheric Administration (NOAA) network (Pétron et al., 2012). Tower measurements allow for continuous measurements in the well mixed boundary layer which are influenced by both nearby sources and the integrated effect of the upstream emissions. While towers provide near continuous coverage of regional emissions, specific emissions sources with specific isotopic signatures are often diluted by mixing, making the differences from background very small.~~

Differentiating CH₄ emissions from natural gas activities from other sources (e.g., wetlands, cattle, landfills) is key to documenting the greenhouse gas impact of natural gas production and to evaluate the effectiveness of emissions reduction activities. The isotopic ratio of methane ($\delta^{13}\text{CH}_4$) is particularly useful in this regard (Coleman et al., 1995). In general, heavy isotope ratios are characteristic of thermogenic CH₄ sources (i.e., fossil-fuel based) and light isotope ratios are characteristic of biogenic sources (Dlugokencky et al., 2011). ~~Schwietzke~~ et al. (2016) compiled a comprehensive database of isotopic methane source signatures, indicating signatures of –44.0 ‰ for globally averaged fossil-fuel sources of methane, –62.2 ‰ for globally averaged microbial sources such as wetlands, ruminants, and landfills, and –22.2 ‰ for globally averaged biomass burning sources. Atmospheric measurements of $\delta^{13}\text{CH}_4$ have been used to partition emissions of CH₄ into source categories (e.g., Mikaloff Fletcher et al., 2004a,b; Kai et al., 2011). It is important to note, however, that for fossil-fuel sources of methane, isotopic ratios of methane vary significantly from reservoir to reservoir (e.g., Townsend-Small et al., 2015; Rella et al., 2015), and with depth in a single reservoir (Molofsky et al., 2011; Baldassare et al., 2014).

The isotopic ratio of methane has traditionally been measured in the laboratory with continuous flow gas chromatography/ isotope ratio mass spectrometry, with repeatability of $\pm 0.05\%$ (Fisher et al., 2006). ~~Röckmann et al.~~

Moved (insertion) [1]

Deleted: Schwietzke

Deleted: Cavity ring-down spectroscopy (CRDS) analyzers measuring continuous isotopic ratio of methane in addition to methane and carbon dioxide (Rella et al., 2015), have recently become commercially available, are more robust in the field, and have the potential to aid with source attribution of the methane signals.

(2016) recently compared continuous in-situ measurements of methane isotopic ratio using a dual isotope mass spectrometric system (IRMS) and a quantum cascade laser absorption spectroscopy (QCLAS)-based technique at the Cabauw tower site in the Netherlands. They showed that high-temporal-resolution methane isotopic ratio data can be used in conjunction with a global and a mesoscale model to evaluate CH₄ emission inventories. Röckmann et al. (2016) also used a moving Keeling plot approach to identify source isotopic ratios.

Cavity ring-down spectroscopy (CRDS) is another technique for measurement of continuous in-situ isotopic ratio of methane (Rella et al., 2015). CRDS is a laser-based technique in which the infrared absorption loss caused by a gas in the sample cell is measured to quantify the mole fraction of the gas. The analyzers utilize three highly reflective mirrors such that the flow cell has an effective optical path length of 15–20 km, allowing highly precise measurements. The temperature and pressure of the sample cell is tightly controlled, improving the stability of the measurements (Crosson 2008). Rella et al. (2015) documented the operation of CRDS (Picarro, Inc., model G2132-i) analyzers, including cross-interference from other gases, and general calibration approach.

Furthermore, Rella et al. (2015) described the use of two tanks to correct for analyzer drift of the isotopic ratio measured by the G2132-i analyzers. In this approach, the variables of interest, i.e., the total methane mole fraction and the isotopic ratio, are directly calibrated. The drift terms in the calibration equations have differing dependence on mole fraction, requiring the use of at least two tanks for calibration. For this study, three field tanks were deployed at each tower location, two for the field calibration and one as an independent test.

In this paper, we describe a network of four tower-based atmospheric observation locations, measuring CH₄ and CO₂ dry mole fractions and $\delta^{13}\text{CH}_4$ using CRDS (Picarro, Inc., model G2132-i) analyzers in the Marcellus shale region in north-central Pennsylvania at towers referred to here as the North, South, East and Central towers. We focus on the specific application of tower-based measurements of isotopic methane using CRDS analyzers. Instead of describing the methods and results separately, we combine these for each topic. First, we describe laboratory calibration of the G2132-i analyzers, field calibration approach, and calibration results. We determine the compatibility achieved for the isotopic measurements in the current field deployment, using an independent field tank, round-robin style testing and comparisons to flasks as our primary metrics. We also evaluate the performance of the G2132-i analyzers in terms of CH₄ and CO₂ mole fractions measurements by comparing to a G2301 analyzer. We then describe the tower locations and compare differences in CH₄ mole fraction and isotopic ratio observed at the towers and use a Keeling plot approach to determine source isotopic signatures. Finally, we describe recommendations for future isotopic methane CRDS tower-based networks.

2 Compatibility goals

Because this is the first network of multiple isotopic ratio of methane continuous analyzers to date, the needed compatibility has not yet been defined. Thus, our compatibility goals for CO₂ and CH₄ mole fractions follow the

Deleted: describe

Moved up [1]: Aircraft-based studies cover large areas, but the temporal coverage is limited. Tower-based networks offer a complementary approach, making continuous measurements over long periods of time. At the Boulder Atmospheric Observatory (BAO) tall tower, daily flask measurements are found to contain enhanced levels of methane and other alkanes, compared to the other tall towers in the National Oceanic and Atmospheric Administration (NOAA) network (Pétron et al., 2012). Tower measurements allow for continuous measurements in the well mixed boundary layer which are influenced by both nearby sources and the integrated effect of the upstream emissions. While towers provide near continuous coverage of regional emissions, specific emissions sources with specific isotopic signatures are often diluted by mixing, making the

Deleted: enhancements above background very small. -

Deleted: calibration

Moved (insertion) [2]

Deleted: ¹³CH₄ and ¹²CH₄ mole fractions), instead of the total methane and the isotopic ratio. This approach

Moved down [3]: has been applied to Fourier Transform infrared and isotope ratio infrared spectrometers measuring $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ of CO₂ in air (Griffith et al., 2012; Wen et al., 2013; Flores et al.,

Deleted: -

... [1]

Deleted: calibrate the isotopologues of methane (i.e.,

Formatted: Font color: Text 1

Deleted: 2017). -

... [2]

Deleted: . We

Deleted: optimized

Deleted: strategy, calibration results

Deleted: comparison to flasks. Our first objective is determining how best to calibrate the isotopic measurements in a field setting, given the field calibration tanks that were deployed for the network.

Deleted: address several questions related to the effects of applying the calibration using different combinations of the tanks. The second objective is to

Deleted: .

Moved up [2]: -

... [3]

Deleted: .

WMO compatibility recommendation for global studies: 0.1 ppm for CO₂ (in the Northern Hemisphere) and 2 ppb for CH₄ (GAW Report No. 229, 2016). Here we use the term compatibility, as advised in the GAW Report No. 229 (2016), to describe the difference between two measurements, rather than the absolute accuracy of those measurements.

For δ¹³CH₄, we set our target compatibility at 0.2 ‰, thought to be a reasonable goal based on laboratory testing prior to deployment and the results shown in Rella et al. (2015). This goal corresponds to the WMO extended compatibility goal for the isotopic ratio of methane, which ~~was deemed~~ sufficient for regionally focused studies with large local fluxes. The measured signal at the towers is a mixture of the source and the background (Pataki et al., 2003), and the ability to distinguish between a biogenic and thermogenic source depends on the difference of the source isotopic signature from background and the peak strength in terms of methane mole fraction. Equating the slope of a source and the background with the slope of a mixture and the background on a Keeling plot (Keeling, 1961), the measured isotopic ratio ~~difference (Δδ)~~ is given by

$$\Delta\delta = (\delta_{\text{src}} - \delta_{\text{back}}) \frac{\Delta\text{CH}_4}{\text{CH}_{4,\text{meas}}} \quad (1)$$

where δ_{src} and δ_{back} are the isotopic ratios of the source and the background, CH_{4,meas} is the measured methane mole fraction, and ΔCH₄ is the difference between the measured mole fraction and the background. This equation is represented graphically in Fig. 1. If there are two possible sources in a region, a biogenic source at -60 ‰ and a thermogenic source at -35 ‰, for example, the difference in isotopic ratio ~~difference~~ is at least three times the compatibility goal of 0.2 ‰ (and thus distinguishable) for a peak strength of 50 ppb CH₄ or greater, assuming a measured CH₄ mole fraction of 2000 ppb and a background isotopic ratio of -47.5 ‰. In this case, the biogenic source would measure 0.3 ‰ above the background, as opposed to the thermogenic source measuring 0.3 ‰ below the background. As shown in Fig. 1, sources closer to the background in isotopic ratio require a larger peak in CH₄ and those further from the background can be attributed with a smaller peak in CH₄.

3 Allan standard deviation testing

Allan standard deviation testing (Allan, 1966) is a useful tool for testing the noise and drift response of instrumentation. The Allan standard deviation for each averaging interval is proportional to the range of values for each averaging interval. This range typically decreases for increasing averaging interval, as the noise is ~~reduced~~ through averaging. As the averaging interval increases, however, analyzer drift may contribute, placing an upper bound on the optimal averaging interval. Thus, the Allan deviation results are critical for defining the minimum averaging time required for a given target compatibility.

To calculate the Allan standard deviation of the G2132-i analyzers used in this study, one tank containing an ambient mole fraction of CH₄ (1.9 ppm), and CO₂ (~400 ppm) mole fraction and one tank containing high mole fraction of

Deleted: , somewhat arbitrarily,

Formatted: Not Superscript/ Subscript

Deleted: is

Deleted: enhancement

Deleted:

Deleted: enhancement

Deleted: 3 Methods: Laboratory testing

... [4]

Deleted: averaged out. Thus, the results are critical for defining the minimum averaging time required for a given target compatibility.

1 CH₄ (9.7 ppm) and an ambient mole fraction of CO₂ (~400 ppm) were ~~tested~~ with an analyzer for 24 hours. For
 2 simplicity, we call these the “high” and “low” tanks, respectively, and they are described further in Section ~~5.1~~. We
 3 tested both as the noise is known to be less for higher mole fractions, and at least one tank with higher CH₄ mole
 4 fraction is necessary for the isotopic ratio calibration (~~Rella et al., 2015~~).

Deleted: sampled

Deleted: 4.2

Deleted: .

6 ~~The resulting Allan standard deviations for $\delta^{13}\text{CH}_4$, CH₄ and CO₂ are shown in Fig. 2. For the high tank, the Allan~~
 7 ~~deviation for $\delta^{13}\text{CH}_4$ (Fig. 2A) was < 0.2 ‰ (our target compatibility) for an averaging interval of 2 min (the averaging~~
 8 ~~interval used each field calibration cycle of the high tank). To reduce the noise to < 0.1 ‰, an averaging interval of 4~~
 9 ~~min is sufficient (in addition to the time required for the transition between gases). For the low tank, in order for the~~
 10 ~~Allan standard deviation to be < 0.2 ‰, 32 min were required and 64 min for 0.1 ‰ noise. Note that for much of the~~
 11 ~~deployment, the near ambient mole fraction target tank was not sampled sufficiently within each day for the desired~~
 12 ~~compatibility goals.~~

Moved (insertion) [4]

Moved (insertion) [5]

Moved (insertion) [6]

Moved (insertion) [7]

14 ~~For CH₄ (Fig. 2B), both the high and low tank Allan deviation were < 1 ppb for even a 1-min averaging interval. The~~
 15 ~~CO₂ levels in the high and low tanks were similar (~400 ppm), and an averaging interval of 6 min corresponded to~~
 16 ~~Allan standard deviations of 0.3 ppm, and 64 min were necessary for 0.1 ppm (Fig. 2C). The performance of the~~
 17 ~~G2132-i analyzers in terms of CO₂ precision is worse than that of the G2301/G2401 analyzers primarily because a~~
 18 ~~weaker spectral line is used (Rella et al., 2015).~~

Moved (insertion) [8]

Deleted: 3.2 Analyzer laboratory

19 **4 Laboratory calibration**

20 **4.1 Experimental set-up**

Deleted: 3.2

21 Prior to field deployment, each analyzer was calibrated for CH₄ and CO₂ mole fraction. Four NOAA-calibrated
 22 tertiary standards (traceable to the WMO X2004 scale for CH₄ and the WMO X2007 scale for CO₂) were used for the
 23 linear mole fraction calibration, as described in Richardson et al. (2017). These NOAA tertiary standards ranged
 24 between 1790 and 2350 ppb CH₄, and between 360 and 450 ppm CO₂.

25 To calibrate the $\delta^{13}\text{CH}_4$ measurement prior to deployment, four different target mixing ratios, each at four different
 26 known isotopic ratios were ~~tested~~ by the four analyzers using the experimental setup in Fig. 3. Commercially-available
 27 isotopic standard bottles (Isometric Instruments, Inc., product numbers L-iso1, B-iso1, T-iso1 and H-iso1) were
 28 diluted with zero air to produce mixtures with varying CH₄ mixing ratios and $\delta^{13}\text{CH}_4$. The gravimetrically-determined
 29 zero air (Scott Marrin, Inc.) was natural ultra-pure air, containing no methane or other alkanes but ambient levels of
 30 CO₂. The isotopic calibration standard bottles each contained approximately 2500 ppm of CH₄ at -23.9, -38.3, -54.5,
 31 and -66.5 ‰ $\delta^{13}\text{CH}_4$, with uncertainty of ± 0.2 ‰ reported by the supplier. ~~These isotopic ratios were tied to the~~
 32 ~~Vienna Pee Dee Belemnite (VPDB) scale.~~ Mass flow-controllers (MC-1SCCM and MC-500SCCM, Alicat Scientific,
 33 Inc.) and a 6-port rotary valve (EUTA-2SD6MWE, Valco Instruments Co., Inc.) were used to direct the standard bottle

Deleted: sampled

Deleted: 2

air for each isotopic calibration standard bottle into a mixing volume (~4 m of 1/8 in, 0.32 cm OD stainless steel tubing; TSS285-120F, VICI Precision Sampling, Inc.) at 0.400 sccm and mixed with zero CH₄ air at 137, 161, 303, and 555 sccm to create target CH₄ mole fractions of 7.3, 6.2, 3.3, and 1.8 ppm, respectively. Thus 16 CH₄ mole fraction/isotopic ratio pairs were produced. The accuracy of the mass flow controllers can be a significant source of error in making mixtures. Here the nominal range of the mass flow controllers was 1 sccm for the standard bottle line and 500 sccm for the zero air line, and the accuracy was ±0.2 % of full scale. To avoid isotopic fractionation at the head of the low-flow mass flow controller, the flow of the zero air was varied rather than the isotope standard. It is possible that fractionation did occur due to the tees used to direct gas into the individual analyzers. For this reason, it would have been preferable to set up the analyzers to each sample directly from the mixing volume.

The first mixture of each isotopic standard was tested for 60 minutes to flush out the span gas line and to avoid isotopic fractionation at the head of the span mass flow controller. Subsequent dilutions using the same isotopic standard were tested for 20 minutes each and each dilution was repeated twice. With the flow rate of 0.400 sccm for the isotopic standard bottles, the total volume of standard gas used was 88 cc. Observations were collected at ~0.5 Hz and the final 5 minutes of data for each dilution were averaged to compare against the target value. The standard deviation of the raw data collected during these tests (Fig. 4) decrease exponentially with increasing mole fraction.

Averaged methane isotopic ratios prior to calibration are shown in Fig. 5. There is an offset in the measured isotopic ratio as a function of the changing known isotopic ratio. For higher mole fractions, this offset is fairly constant, but for near ambient mole fractions it is analyzer-specific. We note that the precision of these results could be improved by averaging over longer periods. We now describe the calibration technique to remove these offsets.

4.2 Application of calibration equations

The first step in the calibration process for the analyzers is to remove the nearly linear error that is a function of isotopic ratio. We applied methods leading from the theoretical framework developed by Rella et al. (2015) to calibrate the isotopic ratio data. Applying a linear fit to highest mole fraction values (7.3 ppm) measured in the laboratory for known δ¹³CH₄ values (−23.9, −38.3, −54.5, −66.5 ‰) for each analyzer, we determined the linear calibration coefficients p_1 and p_0 .

$$[\delta^{13}CH_4]_{intermediate} = p_1[\delta^{13}CH_4]_{measured} + p_0. \quad (2)$$

For this step, we used only the highest mole fraction values because δ¹³CH₄ is more precise for higher mole fractions (Fig. 4). We note that these laboratory tests were completed prior to the Allan standard deviation testing and that the averaging times were not sufficient to achieve the desired compatibility at ambient mole fractions. Ambient mole fractions could be used for this step if measured for sufficient durations.

Deleted:

Deleted: , respectively

Deleted: a common

Moved down [9]: A 3-way solenoid valve (091-0094-900, Parker Hannifin Corp.)

Deleted: was used just downstream of the mixing volume to stop flow from the zero air tank and Isometric Instrument bottles and allow flow from the working standards. In this way, the working standards were sequentially calibrated. The results of these calibrations are described in Section 4.2 and these working standards are referred to as the “high” and “low” field calibration tanks. -

Deleted: sampled

Deleted: sampled

Deleted: We note that these laboratory tests were completed prior to the Allan standard deviation testing and that the averaging times were not sufficient to achieve the desired compatibility for all but the highest mole fractions. Only the highest mole fractions are used for the linear calibration of the analyzers. The resulting averaged data are shown in Fig. 3. An error in isotopic ratio as a function of isotopic ratio is apparent, as is the mole fraction dependence of the isotopic ratio response

Formatted: Subscript

Deleted: 3.2

Deleted: , as is apparent from the Allan standard deviation tests (Section 6.1).

To correct for the CH₄ mole fraction dependence of the measured $\delta^{13}\text{CH}_4$, the two time-dependent drift parameters described in Rella et al. (2015) c_0 and χ must be determined. Here c_0 varies because of spectral variations in the optical loss of the empty cavity and χ varies because of errors in the temperature or pressure of the gas, or changes in the wavelength calibration. These parameters are defined in Eq. (15) of Rella et al. (2015). A coefficient describing the changes in the crosstalk between the two methane isotopologues was ignored, following Rella et al. (2015). For the laboratory calibration, we determined c_0 and χ using measurements at -23.9‰ for a high mole fraction (7.3 ppm) and a low mole fraction (1.8 ppm). We then applied Eq. (12) of Rella et al. (2015)

$$[\delta^{13}\text{CH}_4]_{\text{calibrated}} = [\delta^{13}\text{CH}_4]_{\text{intermediate}} + \frac{c_0}{c_{12}} + \chi([\delta^{13}\text{CH}_4]_{\text{intermediate}} - B), \quad (3)$$

to correct for the CH₄ mole fraction dependence of $\delta^{13}\text{CH}_4$. Here c_{12} is the measured $[\text{CH}_4]$ and

$$B = p_1 B_{\text{default}} + p_0, \quad (4)$$

with B_{default} being -1053.59‰ . B_{default} is the intercept of the fit of the isotopic ratio to the ratio of the absorption peak heights for the standard calibration and B is the updated value, specific to the analyzer. We followed Rella et al. (2015) and ignored the contribution of an additional offset term that depends on neither mole fraction nor isotopic ratio. Note that the slope of the linear calibration was the only component of the calibration that was not adjusted in the field using field tanks (Section 5.4).

5 Methods: Field deployment

5.1 In-situ field tanks

At each tower site, three field tanks were utilized, as listed in Table 1. One tank at each tower site was calibrated by the National Oceanic and Atmospheric Administration (NOAA) for CH₄ and CO₂ mole fractions and by the Institute of Arctic and Alpine Research (INSTAAR) for $\delta^{13}\text{CH}_4$. This tank was tested quasi-daily (every 21 hours) and used to adjust the intercept for the CH₄ and CO₂ mole fraction calibrations (Richardson et al., 2017). The constituents of this tank were at typical ambient levels (as listed in Table 1), and for the purposes of this paper, we call it the “target” although it was not independent.

Two additional tanks were tested at each of the tower sites (Table 1). These tanks were filled using ultra-pure air and spiked (using Isometric Instruments, Inc bottles) by Scott Marrin, LLC, (one at $1.9\text{--}2.1\text{ ppm CH}_4$ and $-23.9\text{‰ } \delta^{13}\text{CH}_4$) and one at $9.7\text{--}10.5\text{ ppm CH}_4$ at $-38.3\text{‰ } \delta^{13}\text{CH}_4$). Recall that these are called the “low” and “high” tanks, for simplicity. These tanks contained ambient levels of CO₂ (368 – 407 ppm). The choice of the CH₄ mole fraction of the high tank is based on the optimal determination of the calibration coefficients c_0 and χ , rather than the expected

Deleted: We

Deleted: calibration

Deleted: 6.2

Deleted: 4

Moved down [10]: .1 Study area ... [5]

Moved down [11]:). The South and North towers were located on the southern and northern edges of the unconventional gas well region, respectively, and were intended to measure background values depending on the wind direction.

Moved down [12]: Measurements began in May 2015, but a complete set of field tanks necessary for calibration of $\delta^{13}\text{CH}_4$ was not deployed until January 2016.

Deleted: 4

Moved down [13]: In addition to the continuous G2132-i analyzers, the East and South towers were also equipped with NOAA flask sampling systems (Turnbull et al.

Deleted: 4

Deleted: Winds in the study area were most often from the west (Fig. 5).

Deleted: This study thus describes results from the period January – December 2016.

Deleted: 2012), measuring a suite of > 55 gases (including greenhouse gases, hydrocarbons, and halocarbons) and $\delta^{13}\text{CH}_4$ [6]

Deleted: calibration

Deleted: calibration

Deleted: We anticipated using one

Deleted: ,

Deleted: , for the CH₄ and CO₂ mole fraction calibration and as an independent check of the $\delta^{13}\text{CH}_4$ calibration, and using two additional tanks for the $\delta^{13}\text{CH}_4$ calibration. Since no long-term field calibration of isotopic measurements have yet been demonstrated, we experimented with different ways to use these three tanks and present two optimized scheme. This analysis also includes suggestions for improving this approach in future field deployments. We first describe the preparation of the field calibration tanks, then the method used to test their application to field measurements. .

Deleted: ”.

Deleted: .

Deleted: sampled

Deleted: They

1 range of ambient CH₄ mole fractions. The effect of c_0 on the calibrated isotopic ratio is largest at low mole fractions,
 2 whereas the effect of χ is independent of mole fraction. Thus the ratio of the high and low tank mole fractions
 3 determines how separable the two effects are. We therefore chose the high tank mole fraction to be as high as possible
 4 without introducing other nonlinearities into the system.

5 The high and low tanks for each tower were calibrated for $\delta^{13}\text{CH}_4$ in the laboratory prior to deployment. First we
 6 applied a linear calibration for $\delta^{13}\text{CH}_4$ using measurements from each of four Isometric Instruments bottles (–23.9, –
 7 38.3, –54.5, –66.5 ‰), diluted with zero air to 10.3 – 10.4 ppm CH₄. A 3-way solenoid valve (091-0094-900, Parker
 8 Hannifin Corp.) was used just downstream of the mixing volume in the laboratory calibration system to stop flow
 9 from the zero air tank and Isometric Instrument bottles and allow flow from the working standards. Then a mole
 10 fraction correction was applied using the –23.9‰ bottle diluted to 10.4 ppm CH₄ and the –38.3 ‰ bottle diluted to 1.9
 11 ppm. These calibration results are shown in Table 1. The values assigned to the tanks differed slightly (with the
 12 differences ranging in magnitude from 0.01 to 0.38 ‰) from the bottles used for spiking. Possible reasons for these
 13 slight differences include noise in the measurement, fractionation upon tank-filling, bottle assignment error with the
 14 0.2 ‰ uncertainty reported by the supplier (Isometric Instruments, Inc.) and insufficient testing times for the tanks at
 15 ambient mole fractions (5 min). We note that it would have been preferable to utilize calibration tanks closer to the
 16 observed air samples in terms of isotopic ratio. In particular, the low tank could have been spiked with the –38.3 ‰
 17 bottle, or a mixture of the –38.3 and –54.5 ‰ bottles.

18 5.2 In-situ field calibration gas sampling system

19 The flow diagram of the field calibration system is shown in Fig. 6. Polyethylene/aluminum composite tubing (¼ in,
 20 0.64 cm OD, Synflex 1300, Eaton Corp.) was used to sample from the top of each tower for the CRDS analyzer and
 21 a separate sample line made from ⅜ in (0.95 cm) OD Synflex 1300 tubing was used for the flask sampling packages.
 22 The top end of each tube was equipped with a rain shield to prevent liquid water from entering the sampling line. For
 23 the CRDS analyzer, air was drawn down the tube at 1 L/min, with 30 cc/min flow into the analyzer and the remainder
 24 purged. The residence time in the tube was about 1 min. Separate tubes were used for the CRDS and flask sampling
 25 lines because of the differing flow rates required for the flask samples (varying between 0.29 and 3.8 liters per minute)
 26 (Turnbull et. al., 2012) and to ensure independence of the CRDS and flask measurements.

27
 28 For the continuous in-situ measurement system, switching between sample and calibration gases was accomplished
 29 using a 6-port rotary valve (EUTA-2SD6MWE, Valco Instruments Co, Inc.). Stainless steel tubing (1/8 in, 0.32 cm
 30 OD, TSS285-120F, VICI Precision Sampling, Inc.) and single-stage regulators (Y11-C444B590, Airgas, Inc.) were
 31 used for testing the field tanks. Rella et al. (2015) noted that the effect of water vapor on the isotopic ratio of methane
 32 measurement is up to 1 ‰ and nonlinear, and recommended drying to less than 0.1% H₂O mole fraction. Thus we
 33 used a Nafion dryer (MD-070-96S-2, PermaPure) in the reflux configuration, with an additional pump (ME1,
 34 Vacuubrand, Inc.) on the outlet of the Nafion dryer (Fig. 6). The sample air was dried to ~0.06 % H₂O, and the

Deleted: by first applying

Formatted: Normal (Web)

Moved (insertion) [9]

Deleted: and

Deleted: .).

Deleted: .

... [7]

Deleted: lines

Deleted: large

Deleted: sampling

Deleted: calibration

Deleted: A

Deleted: using

Deleted: 6), was used to dry the

Deleted: (~

Deleted:)

Deleted: humidify

1 calibration gases were humidified to 0.02 % H₂O, in a manner similar to Andrews et al. (2014). The CH₄ mole fraction
2 was corrected for water vapor following Rella et al. (2015 supp), and the CO₂ mole fraction following Chen et al.
3 (2010).

4
5 A cycle including 90 min of ambient sampling, 6 min testing the high mole fraction field tank, and 10 min testing the
6 low mole fraction field tank was repeated 12 times, then the target tank was tested for 10 min (occurring every ~21
7 hours, to test for diurnal effects). Thus, there were 13.5 calibration cycles for the high and low tanks each day, on
8 average. The first 4 min of data were discarded each time after switching gases to ensure sufficient flushing of the
9 sample cell. After this time, the CO₂ and CH₄ mole fractions stabilized. The ideal calibration tank testing time is a
10 balance between minimizing calibration gas usage (and consequently maximizing ambient air sampling time) and
11 achieving sufficient precision. Note that the Allan standard deviation results indicate that testing for 4 min for the
12 high tank and for 32 min for the low and target tanks is required to achieve our target compatibility of 0.2 ‰ δ¹³CH₄.
13 Thus, this averaging time was achieved in two calibration cycles for the high tank (excluding flushing time), but in
14 5.3 calibration cycles for the low and target tanks (completed in about 10 hours in the case of the sampling scheme
15 utilized for most of the deployment). An improved sampling strategy was implemented on 3 December 2016 and is
16 discussed in Section 5.4.

17
18 The flow rate of the instruments was 35 cc/min, and the 150A tank size was used, corresponding to 4.021 x 10⁶ cc at
19 standard pressure and temperature. Thus there was sufficient gas to test each tank for about one hour per day, for about
20 five years, as a general guideline.

21 5.3 Cross-interference from other species

22 5.3.1 Overview

23 The effects of cross-interference from other species must be considered for spectroscopic measurements. Rella et al.
24 (2015) give proportional relationships for cross-interference from various species for the G2132-i analyzers. Listed
25 in Table 2 are species with potential to affect the isotopic methane calibration, and their estimated effects for tower-
26 based applications. We based these estimates on typical maximum values determined by flask (level at which 99 %
27 of flask measurements at the South and East towers were below; for carbon monoxide, propane, butane, ethylene, and
28 ethane), by in-situ measurements at the towers in this deployment (for water vapor and carbon dioxide), and by typical
29 values (Warneck and Williams, 2012; for ammonia and hydrogen sulfide). There are no known ambient estimates for
30 methyl mercaptan (Barnes, 2015), so the odor threshold (Devos et al., 1990) was used as a maximum value.

31
32 For the Picarro G-2132i analyzers, ethane contributed the largest interference and a correction to the isotopic ratio was
33 applied (Section 4.4.2). Because of water vapor effects, the sample was dried and the calibration gases humidified.
34 The effects of other species were neglected.

Deleted: (

Deleted: sampling

Deleted: calibration

Deleted: sampling

Deleted: calibration

Deleted: sampled

Deleted: (Section 6.1)

Deleted: sampling

Deleted: one

Deleted: cycle

Deleted: ,

Deleted: 6.2

Deleted: 4.4 Method for determination of optimized application

Formatted: Font color: Auto, English (US)

Deleted: field calibration gases .
4.

Deleted: applied to the in-situ data. The optimized strategy for application of the field calibration gases in terms of which tanks to use

Deleted: how often to apply the calibration was determined via testing multiple strategies on the data from October 2016 at the South tower. Our primary metric for determination of performance of these calibration strategies was the bias of the target

Deleted: (difference from the known value, averaged over the month). We also considered the

Deleted: deviation of the difference from the known value for the calibration cycles over the month. Through these "experiments", we address the following questions: .
4.

Deleted: for our target compatibility level? .
4.

Deleted: ? -
4.

Deleted: calibration strategy, we further analyzed the results by testing the strategy, with variations, on all four towers, and evaluated the results again using the bias of the target tank as our primary metric. Although the low tank was planned to be used in the isotopic ratio calibration and the target as an independent test, we compare this strategy to that of instead using the target tank in the calibration and keeping the low tank as the independent test, given the isotopic ratios and sampling times utilized in the field for the majority of the deployment. In this case, the low tank was a metric for determining the compatibility of the isotopic ratio calibration.

Deleted: 4.

Deleted: 4.

Deleted: 5

5.3.2 Ethane correction

Ethane (C_2H_6) is co-emitted with methane during natural gas extraction and its cross-interference with the isotopic ratio of methane is significant. The magnitude of the effect of ethane on the isotopic methane is proportional to its mole fraction and inversely proportional to the methane mole fraction. The two Scott-Marrin field tanks at each site were scrubbed of alkanes (including ethane), but the one NOAA/INSTAAR field tank at each site contained ambient levels of these species. Typical mole fractions of C_2H_6 (1.3 ppb) compared to the Scott-Marrin tanks containing no ethane would lead to a 0.04‰ bias, if uncorrected. Furthermore, flask measurements at the South and East towers indicated ethane up to 8 ppb, which corresponds to a 0.23 ‰ error.

The G2132-i analyzers reported an ethane measurement, but were not designed for high-compatibility C_2H_6 measurements at levels near background. In this deployment, 99 % of the flask measurements, which were taken in the afternoon, were less than 8.0 ppb C_2H_6 . In comparison, the drives near natural gas sources conducted by Rella et al. (2015) indicated C_2H_6 mole fractions up to 13 ppm (note unit change). The ethane signal is subject to strong cross-interference from water vapor, methane and carbon dioxide. Rella et al. (2015; Eq. (S20)) report coefficients for these corrections. These coefficients indicate corrections larger in magnitude than the ethane mole fractions measured in this deployment. We have thus not attempted to analyze the ethane results themselves. The ethane output was however used to correct the isotopic methane data. To do so, we first developed a linear calibration using the Scott-Marrin high field tank containing zero ethane and the NOAA/INSTAAR target tank which we assumed contained a background level of 1.5 ppb ethane (Peischl et al., 2016). This calibration is clearly a rough estimate. Note that we determined the linear relationship between the reported ethane of each analyzer and its calibrated value initially, and assumed that this relationship does not change throughout the deployment. Newer models of the $\delta^{13}CH_4$ analyzer (G2210-i, Picarro Inc.) measure C_2H_6 at ppb levels, simplifying this correction process.

We then corrected the isotopic methane for the effects of ethane cross-interference. For example, 1.3 ppb of ethane in an air sample of 2 ppm CH_4 would, if uncorrected, shift the $\delta^{13}CH_4$ measurement higher by $[+58.56 \text{ ‰ ppm } CH_4(\text{ppm } C_2H_6)^{-1} \times [0.0013 \text{ ppm } C_2H_6]/[2 \text{ ppm } CH_4]=+0.04 \text{ ‰}]$. Note that the calibration coefficient for ethane has been updated from that indicated in Rella et al. (2015). The correction to compensate for this error was applied to all data, using the estimated ethane and measured methane values.

5.3.3 Water vapor and carbon dioxide

Water vapor can have a significant effect on the measurements of isotopic methane (up to $\pm 1 \text{ ‰}$ for up to 2.5 % H_2O) (Rella et al., 2015). Thus, the sample air was dried and the calibration gases slightly humidified such that this effect is minimized (estimated to be $< 0.02 \text{ ‰}$). For the range of ambient CO_2 observed in this study ($\sim 375 - 475 \text{ ppm}$), the difference from the calibration gases was $\sim 100 \text{ ppm}$, and the effect was estimated to be $< 0.03 \text{ ‰}$ (Table 2). The isotopic ratio of methane was thus not corrected for CO_2 effects.

Deleted: 4.

Deleted: calibration

Deleted: calibration

Formatted: Subscript

Formatted: Subscript

Deleted: calibration

Deleted: 4.

Deleted: ($< 0.06 \text{ ‰}$)

Deleted: (0.02 ‰)

5.3.4 Oxygen, argon, and carbon monoxide

The ambient variability in oxygen, argon, and carbon monoxide is expected to have a negligible effect on the isotopic ratio measurements (Rella et al., 2015) and no corrections for these constituents were applied to the isotopic methane data.

5.3.5 Other species

Ammonia, hydrogen sulfide, methyl mercaptan, propane, butane, ethylene are components of natural gas, but their cross-interference effects were small for our tower-based application for which the sources are relatively far from the measurement location. The effects of these species may be significant for other applications, such as automobile-based measurements. Like for ethane, the magnitude of the effect of these gases on the isotopic methane is proportional to the mole fraction of the contaminant species and inversely proportional to the methane mole fraction. In Table 2, maximum mole fractions from the flasks if available, or typical mole fractions from the literature, were used to estimate the effect of these species for our application. The cross-interference from these species was insignificant for our application, < 0.01 %.

5.4 Field calibration

The linear calibration was determined in the laboratory as described in Section 4.2. We then used the daily average of the high and target field tanks to adjust the mole fraction correction (terms c_0 and γ) for the field data. The low tank was used as an independent test. For October 2016, the mean errors for the low tank at the South tower are 0.2 ± 0.7 ‰, for example (Table 3, SCHEME B). Here the standard deviation was calculated using all of the calibration cycles during the month. The errors near the isotopic ratio of the target tank are likely less in magnitude. Instead using the low tank in the calibration and keeping the target tank independent yielded similar magnitudes of errors (Table 3, SCHEME A), but minimized bias near the low tank (about -23.9 ‰) rather than near the target tank (about -47.2 ‰). Therefore, despite increased testing of the low tank throughout the majority of the deployment, we chose to use the target tank in the calibration to minimize errors near ambient isotopic ratios. ▲

On 3 December 2016, an improved tank testing strategy was implemented, in which the target tank testing time was increased from 6 min/day to 54 min/day (excluding transition times), achieved by sampling for 20 min every 420-min cycle (3.4 times/day, on average). The calibration times were completed using multiple cycles in order to avoid not sampling the atmosphere for long periods and to measure possible changes in analyzer response throughout each day. The low tank was tested using an identical strategy (20 min every 420-min cycle), with the total amount of testing time per day changing from 81 min to 54 min. The high tank was tested on average 1.7 times per day (every 840 min) for 10 min. Excluding the transition times, the high tank testing time was thus reduced from 26 min/day to about 10

Deleted: 4.

Deleted: 4.

Deleted: .6 Methods

Formatted: Normal, Space Before: 0 pt, After: 0 pt, Don't keep with next

Formatted: Font color: Auto, English (US)

Deleted: determining enhancements

Formatted: Font color: Auto, English (US)

min/day. Following the implementation of the improved strategy, the mean error of the independent low tank at the sites was similar but the standard deviation was reduced from 0.5 to 1.3 ‰ to 0.3 to 0.9 ‰ (Table 3).

As an example of the effects of calibration, the tank results (differences from known values) using only the laboratory calibration for isotopic ratio, and following the SCHEME B are shown in Fig. 7, for the period September – December 2016. For the results using only the laboratory calibration, analyzer drift is apparent for all three tanks. Without a field calibration, the isotopic ratio was biased by up to 2 ‰. The target tank measurement was used in the calibration; hence the apparent drift following final calibration was necessarily zero. The noise apparent in Fig. 7B prior to 3 December 2016 when the calibration scheme was improved is at least partially due to insufficient sampling times of the target tank.

The relative effects of the calibration terms are illustrated in Fig. 8. The terms c_0 (Fig. 8A) and χ (Fig. 8B) in Eq. (3) are time-dependent drift terms. These terms vary because of spectral variations in the optical loss of the empty cavity (c_0), and because of errors in the temperature or pressure of the gas, or changes in the wavelength calibration (χ). Recall that the parameters c_0 and χ were calculated following Eq. (15) in Rella et al. (2015). The calculation of the parameter c_0 used measurements from the high and target tank. The calculation of the parameter χ used measurements of the high tank and was not independent from p_0 . The largest calibration effect was from the c_0 term, which increased the calibrated isotopic ratios by -0.5 to 4 ‰ during September to December 2016. The χ term increased the final calibrated isotopic ratios by a smaller amount, -0.6 to 0.2 ‰. Thus over this period, there were large changes in the calibration effect of these terms, although no software or hardware changes were applied. Considering shorter term changes, the day to day changes in the calibration were less than 0.5 ‰ for December 2016. Less frequent calibrations, e.g., twice per week, could be considered, but the reduction in field tank use is not large considering the low flow rates of the instruments and steady changes up to 2 ‰ in the raw data over the time scale of days were observed in Rella et al. (2015).

6 Evaluation of the compatibility of in-situ tower measurements

6.1 Independent low tank

The low tank was treated as an ambient sample, independent of the calibration. To evaluate the noise in the calibrated ambient samples that results from noise in the calibration, we calculated the standard deviation over the period September 1 – December 2 of the individual low-tank calibration cycles (6 min each), of the calibration cycles averaged over 1 day (81 min total), and of the calibration cycles averaged over 3 days (4.1 hours total). These results are a proxy for the noise in the calibrated ambient samples over those testing periods.

The low tank differences from known values, averaged over differing intervals, are shown in Fig. 9. The standard deviation of individual low-tank calibration cycles (6 min each) over the period September 1 – December 2 is 0.62

Moved (insertion) [14]

Formatted: Font color: Auto, English (US)

Formatted: Normal, Space Before: 0 pt, After: 0 pt, Don't keep with next

Moved (insertion) [15]

Moved (insertion) [16]

Moved (insertion) [17]

Moved (insertion) [18]

Moved (insertion) [19]

1 %. During this period, the calibration used 6 min/day measurements of the target tank. The standard deviation of the
2 low tank calibration cycles was similar to expectations based on the Allan standard deviation (Fig. 2). The low tank
3 was tested a total of 81 min (1.35 hour) per day. Thus calculating the standard deviation of the low tank values
4 averaged over each day is a measure of the noise due to the calibration scheme for hourly averages of sample data.
5 The standard deviation of daily averages for the low tank (81 min total) was 0.40 %. Based on this result, differences
6 in the hourly average between towers of less than 0.40 % were likely not significant. For 3-day means (a total of 4.1
7 hours), the standard deviation over the three-month period was 0.26 %. For the period after the calibration tank
8 sampling scheme was improved (primarily by sampling the target tank for 54 min/day instead of 6 min/day), December
9 3 – December 31, the standard deviation of the individual cycles reduced substantially, to 0.25 %, and that of the 81-
10 min (4.1 hour) mean of the cycles was 0.18 % (0.11 %). Therefore, according to this metric, after the improved
11 calibration scheme was implemented, differences in the hourly average between towers of greater than 0.18 % were
12 significant.

Moved (insertion) [20]

Moved (insertion) [21]

Moved (insertion) [22]

13 **6.2 Round-robin testing**

14 Post-deployment round-robin style tests were completed in the laboratory in March 2017 for three of the analyzers, to
15 assess the compatibility achievable via our calibration method. The analyzer deployed at the South tower was not
16 included in these tests, as it was still in the field. Two NOAA/INSTAAR tanks (JB03428: -46.82‰ $\delta^{13}\text{CH}_4$, 1895.3
17 ppb CH_4 and 381.63 ppm CO_2 ; and JB03412: -45.29‰ $\delta^{13}\text{CH}_4$, 2385.2 ppb CH_4 and 432.71 ppm CO_2) were tested
18 and treated as unknowns. The uncertainty for these NOAA tertiary standards was 0.1 ppm CO_2 , including scale
19 transfer (Hall 2017; Zhao and Tans 2006), and 1 ppb CH_4 (GAW Report No. 185, 2009). The reproducibility based
20 on the calibration results was 0.06 ppm CO_2 and 0.4 ppb CH_4 . The isotopic ratio was tied to the VPDB scale but was
21 not an official calibration (Michel and Vaughn, personal communication, 2015). The precision of the determined
22 values assigned to the tanks was 0.04‰ ([https://instaar.colorado.edu/research/labs-groups/stable-isotope-](https://instaar.colorado.edu/research/labs-groups/stable-isotope-laboratory/services-detail/)
23 [laboratory/services-detail/](https://instaar.colorado.edu/research/labs-groups/stable-isotope-laboratory/services-detail/)). High, low, and target tanks were tested, with the calibration applied as in the field for
24 ambient samples (as described in Section 5.4). The high mole fraction tank was tested for 20 min and the all ambient
25 mole fraction tanks were tested for 70 min, with 8 min ignored after each gas transition. Four to six tests were
26 completed for each analyzer. We used these tests as a means of evaluating the compatibility of the analyzers, in terms
27 of both mole fractions and the isotopic ratio.

Moved (insertion) [23]

Moved (insertion) [24]

28 ■.....
29 The results for the round-robin style laboratory testing are shown in Fig. 10. The mean of the errors (measured –
30 NOAA known value) for each analyzer/tank pair was -0.08 to 0.04 ppm CO_2 within the 0.1 ppm WMO compatibility
31 recommendation for global studies of CO_2 (GAW Report No. 229, 2016). The standard error, indicating an estimate
32 of how far the sample mean is likely to be from the true mean, for the means of the CO_2 tests were $0.03 - 0.10$ ppm.
33 The mean difference was -0.03 to 0.02 ppm CO_2 for the analyzers, averaged over the two round-robin tanks (analogous
34 to averaging over the entire range of CO_2 during the flask comparison, for example). For CH_4 , the means of the errors
35 were $0.03 - 0.07$ ppb CH_4 , for the NOAA/INSTAAR tank measuring 2385.2 ppb, and -0.83 to -0.70 ppb CH_4 for the

Moved (insertion) [25]

Formatted: Font color: Auto, English (US)

Formatted: List Paragraph, Space Before: 0 pt, After: 0 pt, Don't keep with next

Moved (insertion) [26]

1 NOAA/INSTAAR tank measuring 1895.3 ppb CH₄. Therefore, there was a slight error in the slope of the linear
2 calibration, possibly attributable to tank assignment errors. However, the error was well within the WMO
3 recommendations for global studies of 2 ppb CH₄ (GAW Report No. 229, 2016), and the range of NOAA/INSTAAR
4 tanks encompassed the majority of the CH₄ mole fraction observed during the study. We also note that the standard
5 error for the means of the CH₄ tests were 0.07 – 0.12 ppb. Averaging over the two round-robin tanks, the mean
6 difference was –0.40 to –0.32 ppm CH₄ for the analyzers. For $\delta^{13}\text{CH}_4$, the mean errors for each analyzer/tank pair
7 were –0.33 to 0.24 ‰ for these tanks within the range of ambient isotopic ratio and the standard errors were 0.05 –
8 0.10 ‰. The mean errors were –0.14 to 0.03 ‰ for each analyzer.

9 6.3 Side-by-side testing

10 The precision and drift characteristics are not optimized for CO₂ for the G2132-i analyzers, compared to the G2301
11 and G2401 analyzers, which measure mole fractions and not isotopic ratios. Whereas the spectral line for CH₄ is the
12 same between the two types of analyzers (Rella et al., 2014), for CO₂, the absorbance of the spectral line used in the
13 G2132-i analyzers is a factor of 11 times less, meaning the precision is dramatically reduced. Although not central to
14 the primary results of this project, the performance of the analyzers in terms of CO₂ is important if the data are to be
15 used as part of the continental-scale CO₂ network. To test the performance of the G2132-i analyzers for consideration
16 of the data for this use, G2301 and G2132-i (Picarro, Inc.) analyzers were run side-by-side for one month (June 2016)
17 at the South tower. The sampling system for the G2132-i was as described in Section 5.2. A separate ¼” (0.64 cm)
18 tube was used for the G2301 analyzer and an intercept calibration using the target tank was applied daily. The sample
19 air for the G2301 analyzer was not dried and the internal water vapor correction was used.

20
21 This testing resulted in mean differences of 0.06±0.41 ppm CO₂ and 0.9±1.5 ppb CH₄ with the G2132-i analyzer
22 measuring slightly lower for both species. Here the standard deviation was based on the 10-min average calibrated
23 values for the month for all times of the day. The standard error of the differences was 0.01 ppm CO₂ and 0.02 ppb
24 CH₄. These results indicate that the performance of the G2132-i is similar for CO₂ and CH₄ mole fractions, at least in
25 terms of the long-term mean. In terms of utilizing the mole fraction data in atmospheric inversions, the multi-day
26 mean afternoon differences are most appropriate. The five-day mean afternoon difference for the month was
27 0.05±0.08 ppm CO₂ and –0.7±0.1 ppb CH₄. The G2132-i analyzers are thus appropriate for use in the atmospheric
28 inversions and in the global network where 0.1 ppm CO₂ and 2.0 ppb CH₄ have been identified as criteria. For these
29 results, recall that the target tank was tested for a total of 30 min in five days. To optimize results on a daily time
30 scale, sampling the target tank for 60 min per day would be preferable for improving CO₂ results. We also note that
31 round robin testing of these instruments requires 60 min sampling per tank.

Moved (insertion) [27]

Moved (insertion) [28]

Moved (insertion) [29]

Moved (insertion) [30]

6.4 Flask to in-situ comparison

In addition to the continuous G2132-i analyzers, the East and South towers were also equipped with NOAA flask sampling systems (Turnbull et al. 2012). These flask measurements were used for independent validation and error estimation of the continuous CO₂, CH₄ and $\delta^{13}\text{CH}_4$ in-situ measurements. In addition, the flasks were measured for a suite of species including N₂O, SF₆, CO, H₂ (Dlugokencky et al., 2017), halo- and hydro-carbons (Montzka et al. 1993) and stable isotopes of CH₄ (Vaughn et al., 2004). The flasks were filled over a 1-hour time period in the afternoon (1400–1500 LST), thereby yielding a more representative measurement compared to most flask sampling systems, which collect nearly instantaneous samples (e.g., ~10 sec). Samples were collected only when winds were blowing steadily out of the west or north (~45–225°) to ensure that the samples were sensitive to and representative of the broader Marcellus shale gas production region that is the focus of this study. For the in-situ data, ten-minute segments were reported. These were averaged over the hour for comparison with the flask measurements. For CH₄, data points with high temporal variability (standard deviation of the 10-min means within the hour > 20 ppb) were excluded, on the basis that the ambient variability was large, making comparisons difficult.

For January – December 2016, the mean flask to in-situ CH₄ difference at the East tower was -1.2 ± 2.2 ppb CH₄, and at the South tower was -0.9 ± 1.4 ppb CH₄ (Fig. 11A). Here the standard deviation reported is that of the hourly flask to in-situ differences. Thus, at the South tower, for example, on 67% of the sampled afternoons indicated differences for CH₄ within 1.4 ppb of the mean of -0.9 ppb. The standard error was 0.24 ppb at the East tower and 0.14 ppb at the South tower. Thus, there is high confidence that the difference between the in-situ and flask measurements at both towers is more compatible than the WMO recommendation. As for the side-by-side testing, the G2132-i analyzers were slightly lower than the “known”, in this case, the flask results. The difference, was however, less than the target compatibility, and the flasks could in theory be biased.

Although CO₂ is not the focus of this paper, the differences were -0.21 ± 0.31 ppm for the East tower and 0.21 ± 0.35 ppm for the South tower (Fig. 11B). The standard error was 0.03 ppm at the East tower and 0.04 ppm at the South tower. The magnitude of CO₂ differences was somewhat larger in the growing season. The mean flask to in-situ differences were thus larger than the WMO recommendation of 0.1 ppm, but at the extended compatibility goal of 0.2 ppm CO₂ (GAW Report No. 229, 2016).

For the isotopic ratio of methane, the mean flask to in-situ differences were $0.08 \pm 0.54\text{‰}$ and $0.02 \pm 0.38\text{‰}$ at the East and South towers, respectively (Fig. 11C). The standard error of the differences was 0.06‰ and 0.04‰ at the East and South towers, respectively. The range of $\delta^{13}\text{CH}_4$ throughout the project (including day and night) was relatively small: one standard deviation (67%) of the data points are between 46.7 – 48.2 ‰, a range of 1.5 ‰. Errors for isotopic ratios outside the calibration range (further from the high and target calibration tanks) would likely be larger. For example, the mean error of the independent low tanks (averaging over all calibration cycles during a one month period) at the towers (Table 3) were 0.2 – 0.7 ‰.

Moved (insertion) [13]

Moved (insertion) [31]

Moved (insertion) [32]

Moved (insertion) [33]

Moved (insertion) [34]

Moved (insertion) [35]

Moved (insertion) [36]

Moved (insertion) [37]

Moved (insertion) [38]

7 Network comparisons

7.1 Study area

Four CRDS isotopic CH_4 analyzers (G2132-i, Picarro, Inc.) were deployed on commercial towers 46–61 m AGL in northeast Pennsylvania (Fig. 12). The South and North towers were located on the southern and northern edges of the unconventional gas well region, respectively, and were intended to measure background values depending on the wind direction. Measurements began in May 2015, but a complete set of field tanks necessary for calibration of $\delta^{13}\text{CH}_4$ was not deployed until January 2016. The Central tower measured only mole fractions for the period June – December 2016. For inter-tower comparisons, we focused on the period January – May 2016 when all sites measured both CH_4 and $\delta^{13}\text{CH}_4$.

7.2 Inter-network differences in CH_4 and $\delta^{13}\text{CH}_4$

A background value is required to calculate differences in CH_4 and $\delta^{13}\text{CH}_4$. For this simple analysis, we chose a single tower to represent the background for the entire period. The predominant wind direction for the Marcellus region is from the west (Fig. 13). For westerly winds, the South tower is a reasonable choice for a background tower. The South tower measured the lowest overall mean afternoon methane mole fraction (1960.2 ppb CH_4). The mean afternoon methane mole fractions of the other towers, averaged only when data for the South tower exist, were 8.7, 7.0, and 2.9 ppb higher, at the North, Central, and East towers, respectively. For future analysis, a wind direction-dependent background tower (South or North) could be considered, but the North tower did have the largest mean enhancement in CH_4 mole fraction compared to the South tower. As noted by Barkley et al. (2017), the area encompassing southwestern Pennsylvania and northeastern West Virginia contains large sources of CH_4 , with emissions from conventional gas, unconventional gas, and coal mines all having significant contributions to the total. These large sources complicated the interpretation of the signals, as does changing wind direction. For this overview analysis, we calculated differences above the South background tower to determine overall signal strength to compare with our target compatibility. We first examine the afternoon (defined here as 1700 – 2059 UTC), when the atmospheric is well mixed, allowing simpler interpretation of the measurements and more tractable modeling. We then consider non-afternoon hours, when the atmosphere is less mixed and signals are typically larger.

In the first set of plots, we focus on the majority of the afternoon data points by truncating the scale for the probability distribution functions of methane mole fraction and isotopic ratio (Fig. 14A, B, D, E, G, and H). The averaging interval of the individual data points was 10 min and the data were afternoon only (1700–2059 UTC, 1200–1559 LST) for the time period January – May 2016. The median differences for both isotopic ratio (–0.15 to 0.12 ‰) and methane mole fraction (less than 1 ppb) were less in magnitude than the compatibility of the analyzers. This result is generally consistent with the results of Barkley et al. (2017), who found the emission rate of methane due to natural gas extraction activities to be very low, 0.36 % of total production. The standard deviation of 10-min segments of isotopic

Moved (insertion) [19]	[... [18]
Moved (insertion) [11]	[... [9]
Deleted: enhancements	
Deleted: 5...3). For westerly winds, the South tower	[... [10]
Moved up [14]: .	[... [22]
Formatted	[... [23]
Deleted: described in Section 3.1, two tanks were sam	[... [24]
Moved up [4]: The resulting Allan standard deviation	[... [25]
Moved up [5]: . For the high tank, the Allan deviat	[... [26]
Deleted: Keeling plots (Keeling 1961; Röckmann et a	[... [11]
Moved up [23]: The analyzer deployed at the South	[... [12]
Moved up [24]: Four to six tests were completed for	[... [14]
Moved up [27]: analyzers were run side-by-side for	[... [16]
Moved up [18]:) in Eq. (3) are time-dependent drift	[... [34]
Deleted: sampled for 70 min, with 8 min ignored after	[... [13]
Deleted: There is variability in the calibration effect	[... [35]
Deleted: 5.3 Side-by-side testing .	[... [15]
Deleted: 4.3. A separate 1/4" (0.64 cm) tubing was us	[... [17]
Deleted: 7	
Moved up [31]:), halo- and hydro-carbons (Montzka	[... [18]
Moved up [32]: afternoon (1400–1500 LST), thereb	[... [20]
Deleted: 7A	
Moved up [6]:) was < 0.2 ‰ (our target compatibil	[... [27]
Deleted: 7B), both the high and low tank Allan devia	[... [28]
Moved up [8]: The CO_2 levels in the high and low ta	[... [29]
Deleted: The flasks were filled over a 1-hour time pe	[... [19]
Deleted: 7C). .	[... [30]
Moved up [15]: for the period September – Decemb	[... [31]
Deleted: The target tank was sampled only once per	[... [32]
Moved up [16]: .	[... [33]
Deleted: 9B) and χ (Fig. 9C	
Deleted: Samples were measured only when winds w	[... [21]
Deleted: 9	
Moved up [17]: . The terms c_0 (Fig.	
Moved up [19]: The low tank differences from know	[... [36]
Deleted: 10. As described in Section 5.1, the standar	[... [37]
Moved up [20]: During this period, the calibration us	[... [38]
Deleted: 7). Averaging over the cycles in one day (a	[... [39]
Moved up [21]: Based on this result, differences in	[... [40]
Deleted: For 3-day means (a total of 4.1 hr), the stand	[... [41]
Moved up [22]: Therefore, according to this metric, a	[... [42]
Deleted: 6.4 Round-robin testing	
Moved up [25]: .	[... [43]
Formatted	[... [44]
Deleted: of two NOAA/INSTAAR tanks are shown i	[... [45]
Moved up [26]: 229, 2016). The standard error, ind	[... [46]
Moved up [28]: resulted in mean differences of 0.06	[... [48]
Deleted: .	[... [47]
Deleted: more appropriate. The five-day mean after	[... [49]
Moved up [29]: The G2132-i analyzers are thus	[... [50]
Moved up [30]: To optimize results on a daily time s	[... [51]
Deleted: sampled for a total of 30 min in five days.	
Deleted: 6.6 Flask to in-situ comparison .	
Moved up [33]: For January – December 2016, the m	[... [52]
	[... [54]
	[... [53]
	[... [55]
	[... [56]
	[... [57]
	[... [58]
	[... [59]

ratio differences was 0.8 ‰ at each of the towers. We note that the Allan standard deviation for 10-min averaging times for ambient levels of methane was 0.4 ‰ $\delta^{13}\text{CH}_4$. The standard deviation of the daily afternoon averages (rather than 10-min averages) was 0.6 – 0.7 ‰. Thus the observed width of the distribution appears to be persistent throughout the afternoon and not merely measurement noise. For isotopic ratio, 43 – 54 ‰, depending on the tower, of the 10-min segments were greater than 0.6 ‰ in magnitude (3 times the target compatibility) (Fig. J4A, D, and G) and are thus detectable by the analyzers. The standard deviations of the methane mole fraction differences were 60.7, 30.0 and 33.8 ppb for the North, Central, and East towers, respectively (Fig. J4 B, E, and H). 57 – 66 % of the data points indicated differences greater than 6 ppb CH_4 in magnitude (3 times the target compatibility) for the North, Central, and East towers, respectively (Fig. J4 B, E, and H) and are thus detectable. The majority of afternoon data points indicated relatively few local sources of contamination.

There are however a few outliers during the time period with large values above the background tower during the afternoon hours (up to 1500 ppb enhancement at the North tower). The isotopic as a function of inverse methane mole fraction at each non-background tower are shown in Fig 13C, F, and I. While the range of measured isotopic ratios is large, the majority of the 10-min means lie close to the ambient values: the standard deviation of the 10-min means of the measured isotopic ratios during the afternoon were 0.6 – 0.8 ‰.

During non-afternoon hours (0000–1659 and 2100–2359 UTC), the median isotopic ratio difference from the South tower were still indistinguishable from zero (Fig. J5A, D, and G). The median methane mole fraction enhancement was slightly higher than during the afternoons, at 3.5, 6.8, and 9.8 ppb for the North, Central, and East towers, respectively (Fig. J5B, E, and H). There were however more outliers, particularly at the Central tower (Fig. J5C, F, and I). Applying a best fit line to all of the data shown in Fig 14F gave a poor correlation coefficient ($r^2=0.22$) because there were many data points with no local sources.

7.3 Keeling plots

Keeling plots (Keeling 1961; Röckmann et al., 2016) are used to infer the isotopic ratio of the methane source as the intercept of the best fit line of the isotopic ratio as a function of the inverse methane mole fraction. We used this approach to estimate the source isotopic ratio of the eight largest peaks observed during non-afternoon hours at the Central tower. The time series of CH_4 encompassing the peak observed on DOY 55 is shown in Fig. 16, as an example. The time during which the tower was in the plume was clear (lasting about 1.5 hours) and only those points were included in the calculation of the linear fit.

The Keeling plots for each of the eight largest peaks in the non-afternoon methane data are shown in Fig. J7. The intercepts of the best fit lines for the peaks indicate that the sources contributing to the peaks have a mean isotopic ratio of -31.2 ± 1.9 ‰. The correlation coefficients were high ($r^2=0.92 - 1.0$) except for one peak, which was excluded from the statistics. Propagating a potential error (attributable of analyzer uncertainty) of 0.2 ‰ at the heavy end of

Deleted: enhancements

Deleted: For the daily afternoon averages, the

Deleted: was reduced only slightly, to

Deleted: majority of the

Deleted: real.

Deleted: %

Deleted: 13A

Deleted: enhancements

Deleted: 13

Deleted: enhancements

Deleted: of 0.2 ‰)

Deleted: 13

Deleted: enhancements above

Deleted: 14A

Deleted: 14B

Deleted: 14C

Formatted: Normal, Indent: Left: 0"

Deleted: The

Deleted: (i.e., the Keeling plot)

Deleted: 15

Deleted: gave

Deleted: ‰, with values ranging from -33.2 to -28.0 ‰.

Deleted:),

1 the Keeling plots and -0.2% at the light end, and vice versa, the potential range of the mean is from -32.0 to -30.4
 2 $\%$.
 3
 4 Compared to mobile measurements near the ground, for example, the footprints of towers are large, which is ideal for
 5 determining regional emissions. But the emissions sources with specific isotopic signatures are diluted by mixing,
 6 making the enhancements above background small, particularly for this region/time period with small leakage rates.
 7 For these eight non-afternoon peaks at the Central tower, the enhancements over background were $334.1 - 2007.8$ ppb
 8 CH_4 and the differences of isotopic ratio were -2.5 to -8.7% .

9 8 Discussion

10 In this paper, we present the methods used to calibrate a network of four CRDS methane isotopic ratio analyzers
 11 (Picarro G-2132i). Evaluation of the calibration results using an independent tank, round-robin style testing and flask
 12 comparisons showed that the analyzers are compatible within 0.2% . The calibration required consideration of 1) the
 13 isotopic ratio linear calibration, 2) the mole fraction dependence of the isotopic ratio calibration (using high and
 14 ambient mole fraction tanks), 3) the correction due to ethane cross interference (using one tank without ethane and
 15 one tank with ambient ethane), and 4) drift in the CO_2 and CH_4 mole fractions (using at least one tank near ambient
 16 isotopic ratio and mole fraction). The isotopic ratios and CH_4 mole fractions of these tanks as used in the present
 17 deployment are graphically represented in Fig. 18A. Prior to implementation of the improved field tank testing
 18 strategy, the high and low tanks were tested for 26 and 52 min/day (excluding transition time between gases), as listed
 19 in Table 4. The testing times throughout each day for the high and low standards are sufficient for Allan deviation $<$
 20 0.1% , but neither of those tanks were at ambient ranges of $\delta^{13}\text{CH}_4$. If the calibrations and analyzer response were
 21 both linear, we would expect negligible errors in the target tank if kept independent, but we found a bias between $-$
 22 0.3 and -0.8% , which is very large compared to the ambient differences observed (one standard deviation of the
 23 tower measurements at all times of day were between -48.2 and -46.7%). Thus, we instead chose to minimize mean
 24 error at ambient values (target tank) rather than at the isotopic ratios of the low tanks (-23.9%). This procedure
 25 added noise to the ambient data because the daily sampling time for the target tank was only 6 min/day. On 3
 26 December 2016, we implemented an improved tank testing strategy, primarily by increasing the testing time for the
 27 target tank to 54 min/day.

29 Our recommendation for future similar studies is to choose both target and low tanks closer to the expected range of
 30 isotopic ratios, in addition to being near ambient CH_4 mole fractions. For example, suggested values for the low and
 31 target tanks are 2.1 ppm CH_4 at -46.5% and 1.9 ppm CH_4 at -47.5% (Table 4 and Fig. 18B). The testing time
 32 required is dependent upon the compatibility goals. After implementing our improved tank testing time strategy, we
 33 tested each target and low tank for about an hour per day, to achieve Allan deviations of 0.2% . Source attribution
 34 using mobile measurements, rather than tower measurements, for example, is less demanding in terms of compatibility
 35 needed, due to the relatively large ambient signals typically encountered. The estimated testing time required to

- Deleted: .
- Deleted: The footprints of towers are large compared
- Deleted: Thus, specific
- Deleted: very
- Deleted: .
- Deleted: ,
- Deleted: peak heights
- Deleted: -
- Deleted: The reduced methane enhancement at other data points within the peak was indicative of increased dilution of the local source by the background signal.
- Deleted: 7
- Deleted: and Summary
- Deleted: We found that field calibrations (including both a linear calibration and a correction for mole fraction dependence of $\delta^{13}\text{CH}_4$) significantly improved the compatibility of the measurements (as seen by comparing EXPTs H and F with EXPT D in Table 3). There was, however, no significant drift within a single day (as seen by comparing variations of EXPT D in Table 4). Using (... [60])
- Deleted: in the deployment prior
- Deleted: 3 December 2016 (Table 5). Instead of using (... [61])
- Deleted: of methane. (... [62])
- Formatted: Font color: Text 1
- Deleted: . These results were within the WMO (... [63])
- Formatted: Font color: Text 1
- Deleted: with reduced sampling times of about 10 min (... [64])
- Deleted: to in-situ comparison for $\delta^{13}\text{CH}_4$ (Fig. 12) shown (... [65])
- Deleted: require
- Deleted: (preferably using high mole fractions tanks),
- Deleted: low
- Deleted: 16A
- Deleted: improving
- Deleted: calibration
- Deleted: sampling
- Deleted: sampled
- Deleted: 5
- Deleted: sampling
- Deleted: about 0.1% error
- Deleted: EXPT D in Table 4 indicate
- Deleted: enhancements
- Deleted: Note that the high mole fraction tanks are identical (... [66])
- Formatted: Normal, Indent: Left: 0"
- Deleted: The requirements to achieve compatibility
- Deleted: $\delta^{13}\text{CH}_4$ for tower measurements as described (... [67])
- Deleted: If compatibility of 0.5% is sufficient for a given (... [68])

1 achieve Allan deviations less than 0.4 ‰, for example, can be achieved in 8 min. In general, it is desirable to distribute
2 the tank testing time throughout the time period, in our case, one day. In this case, persistent changes in analyzer
3 response over the day, if any, would be averaged over rather than an extreme value used in the calibration. This
4 procedure also avoids not sampling the ambient air for extended periods. We did not find any evidence of variability
5 in the calibrations on scales less than one day, compared to the precision possible given our tank testing times, but this
6 possibility could be further explored by testing the field tanks for longer periods of time.

7
8 The high tanks used in this network contained methane with about -38.3 ‰ CH₄. This specific isotopic ratio is
9 available commercially, and depending on the compatibility goals of the project, may not require laboratory calibration
10 of the tank. For our case, however, it may have been beneficial to utilize isotopic ratios closer to the observed range,
11 perhaps -44 ‰ (Table 4). Another possibility is to add an additional high tank (Fig. 18B) in the range of -54.5 ‰ to
12 -52 ‰ (with -52 ‰ more closely bracketed the observed isotopic ratios in the present study). In this case, laboratory
13 linear calibration of the analyzers is not necessarily required. Both the slope and intercept of the linear calibration can
14 be adjusted in field, rather than just the intercept, which may improve the calculated accuracy and precision. However,
15 the laboratory calibration in the present study utilized four different isotopic ratios, rather than two, and it is unknown
16 which is more important – improving linear calibration frequency or avoiding over-constraining the calibration.

17
18 In this paper, we calibrated the total CH₄ and the isotopic ratio of methane. An alternative calibration approach is to
19 separately calibrate the individual isotopologues (in this case, ¹³CH₄ and ¹²CH₄ dry mole fractions), as has been applied
20 to Fourier Transform infrared and isotope ratio infrared spectrometers measuring δ¹³C and δ¹⁸O of CO₂ in air (Griffith
21 et al., 2012; Wen et al., 2013; Flores et al., 2017). This approach has the advantage of simple calibration equations,
22 but has the disadvantage that the quantities of interest (e.g., total mole fraction and isotopic ratio) are calculated rather
23 than directly calibration. Like the approach applied in this paper, it also requires at least two standard tanks, and could
24 utilize an independent tank for testing. Rella et al. (2015) list further practical reasons to calibrate δ¹³CH₄ including
25 the lack of primary standards for ¹³CH₄. However, a comparison of performance using each of these techniques on
26 the same dataset would be beneficial. ▲

27
28 The signals observed in the study region were generally small, but the isotopic ratio differences were larger than would
29 be expected based on the methane mole fraction enhancements from local sources. For afternoon hours at the Central
30 tower, for example, 43 % of the differences in δ¹³CH₄ were detectable above background with magnitudes > 0.6 ‰,
31 3 times the analyzer compatibility. For a thermogenic source with isotopic ratio of -35 ‰, a background isotopic
32 ratio of -47 ‰, and assuming a measured CH₄ mole fraction of 2000 ppb, a measured isotopic ratio difference of -
33 0.6 ‰ corresponds to a 100 ppb peak in CH₄ above background, following Eq. (1). Enhancements in CH₄ of 100 ppb
34 were rarely encountered, however (Fig. 14B, E, and H). Using Eq. (1) to predict differences of isotopic ratio based on
35 the observed methane mole fraction enhancements corresponded to only 3 % of the isotopic ratio differences expected
36 to be > 0.6 ‰ in magnitude. Thus during the afternoon hours, most of the deviations from background were not likely
37 directly from local sources. These larger than expected differences in isotopic ratio are not primarily attributable to

Deleted: this case, the noise is increased by reducing sampling time requirements. Another possible simplification

Deleted: utilize the isotopic ratio of

Deleted: Isometric Inc. bottle used to spike

Deleted: high and low tank,

Deleted: calibrating them

Deleted: laboratory. The

Deleted: the tanks is shown to be <0.4 ‰ magnitude different from the bottle value (Table 1), and varies by tank. If

Deleted: one analyzer

Deleted: used to sample a plume, this bias

Deleted: unimportant.

Moved (insertion) [3]

Formatted: Font color: Text 1

Deleted:) (Flores et al.,

Formatted: Font color: Auto

Deleted: .)

Formatted: Font color: Auto

Deleted: enhancements

Deleted: enhancements

Deleted: enhancement

Deleted: 13B

Deleted: enhancements

Deleted: enhancements

analyzer noise. The Allan Deviation (Fig. 2) is 0.4 ‰ for 10-min means at ambient mole fractions of 2 ppm CH₄. We also note that we focused on the period January – May 2016 in this work. Larger differences were observed in the latter half of 2016.

Deleted: enhancements

During the morning hours, however, several peaks resulting from local sources were observed. The mean source isotopic signal indicated by Keeling plot analysis of the eight largest peaks at the Central tower was -31.2 ± 1.9 ‰, fairly heavy even for oil/natural gas sources. In general, the isotopic signature for natural gas sources varies from region to region, and even within one region. The mean isotopic ratio of methane in gas wells in the northeastern Pennsylvania section of the Marcellus region has been shown to vary by depth, from -43.42 ‰ with a standard deviation of 6.84 ‰ for depths of 0 to 305 m, to -32.46 ‰ with a standard deviation of 3.84 ‰ for depths greater than 1524 m (Baldassare et al., 2014). Similarly, Molofsky et al. (2011) found that the isotopic signatures of gases from the deeper layers of the Marcellus Shale in Susquehanna County, Pennsylvania, to be heavier than the shallower Middle and Upper Devonian deposits, with values for the deep layers ranging from -30 to -21 ‰. Thus, the source signature determined here is consistent with a natural gas source originating from deep wells in the Marcellus region. The peaks occurred during the morning hours, when the boundary layer is typically stable, making modeling more difficult, and the winds prior to the peaks were not from a consistent direction. Determination of the location of the specific emitter(s) contributing to these peaks is thus beyond the scope of this paper. Based on the lack of consistent wind direction, it seems likely that more than one location (with potentially different source signatures) contributed to these peaks. We note that the Keeling plot approach to determine source isotopic signatures far from the point of emission will be difficult to apply in regions without sources that are significantly depleted or enriched in ¹³CH₄ compared to ambient.

Deleted: was

Deleted: Determining

For determination of the source signature for a specific known location, the tower-based approach is not ideal. Instead the strength of the tower-based approach lies in covering larger areas and many potential source locations, and for longer periods of time than is feasible by other approaches. The instrumental performance demonstrated here could be used to disaggregate methane sources in areas of stronger enhancements and differing source isotopic signatures. Networks of high-temporal-resolution methane isotopic ratio data have the potential to constrain regional methane budgets when used within a modeling framework.

Deleted: In this paper, we present details of the laboratory and field calibrations of cavity ring-down spectroscopy (Picarro, Inc, G2132-i) analyzers measuring CH₄, δ¹³CH₄, and CO₂. Numerous calibration schemes were tested, and the effects on measurement compatibility of each scheme were quantified. We have demonstrated that the network of four G2132-i analyzers, deployed on towers in the Marcellus Shale region in Pennsylvania, was able to detect the small differences in CH₄ and δ¹³CH₄.

Formatted: Font color: Auto

Data Availability

Miles, N.L., D.K. Martins, S.J. Richardson, T. Lauvaux, K.J. Davis, B.J. Haupt, and C. Rella, 2017. In-situ tower atmospheric methane mole fraction and isotopic ratio of methane data, Marcellus Shale Gas Region, Pennsylvania, USA, 2015-2016. Data set. Available on-line [http://datacommons.psu.edu] from The Pennsylvania State University Data Commons, University Park, Pennsylvania, USA. <http://dx.doi.org/10.18113/D3SG6N>.

Competing Interests

1 TL, SJR, NLM, and KJD are co-owners of a related company, Carbon Now Cast, LLC.

2

3

4 *Acknowledgments.* The authors thank B. Vaughn and S. Englund Michel (Institute of Arctic and Alpine Research,
5 University of Colorado) for providing analysis of methane isotopic ratios of the flask data and for advice regarding
6 gas handling techniques involving isotopic ratios. The authors also acknowledge R.P. Barkley (Tunkhannock Area
7 Middle School) for his contributions to maintaining instrumentation at the tower sites. This work was funded by the
8 Department of Energy National Energy Technology Laboratory (DE-FOA-0000894).

9

10

References

Allan, D. W.: Statistics of atomic frequency standards, P. IEEE, 54, 221–230, 1966.

Alvarez, R.A., Pacalab, S.W., Winebrake, J.J., Chamedies, W.L., and Hamburg, S.P.: Greater focus needed on methane leakage from natural gas infrastructure, P. Natl. Acad. Sci. USA, 109, 6435–6440, doi:10.1073/pnas.1202407109, 2012.

Andrews A.E, Kofler, J.D., Trudeau, M.E., Williams, J.C., Neff D.H., et al.: CO₂, CO, and CH₄ measurements from tall towers in the NOAA Earth System Research Laboratory's Global Greenhouse Gas Reference Network: Instrumentation, uncertainty analysis, and recommendations for future high-accuracy greenhouse gas monitoring efforts. Atmos. Meas. Tech. 7, 647–687. www.atmos-meas-tech.net/7/647/2014/ doi:10.5194/amt-7-647-2014, 2014.

Baldassare, F.J., McCaffrey, M.A., and Harper, J.A.: A geochemical context for stray gas investigations in the northern Appalachian Basin: Implications of analyses of natural gases from Neogene-through Devonian-age strata, Amer. Assoc. Petr. Geol. Bull., 98(2):341–372, doi:10.1306/061113121782014, 2014.

Barnes, I.: Tropospheric Chemistry and Composition: Sulfur Chemistry, Organic in Encyclopedia and Atmospheric Sciences, edited by North, G. R., Pyle, J. A. and Zhang, F., Elsevier, Vol. 6, p. 257, 2015.

Barkley, Z. R., Lauvaux, T., Davis, K. J., Deng, A., Cao, Y., Sweeney, C., Martins, D., Miles, N. L., Richardson, S. J., Murphy, T., Cervone, G., Karion, A., Schwietzke, S., Smith, M., Kort, E. A., and Maasakkers, J. D.: Quantifying methane emissions from natural gas production in northeastern Pennsylvania, Atmos. Chem. Phys., 13941 – 13966, doi.org/10.5194/acp-17-13941-2017, 2017.

Chen, H., Winderlich, J., Gerbig, C., Hoefer, A., Rella, C. W., Crosson, E. R., Van Pelt, A. D., Steinbach, J., Kolle, O., Beck, V., Daube, B. C., Gottlieb, E. W., Chow, V. Y., Santoni, G. W., and Wofsy, S. C.: High-accuracy continuous airborne measurements of greenhouse gases (CO₂ and CH₄) using the cavity ring-down spectroscopy (CRDS) technique, Atmos. Meas. Tech., 3, 375–386, 2010.

Coleman, D.D., Liu, C.-L., Hackley, K.C., and Pelphrey, S.R.: Isotopic identification of landfill methane, Environ. Geosci., 2(2), 95–103, 1995.

Plugokencky, E.J., Lang, P.M., Mund, J.W., Crotwell, A.M., Crotwell, M.J., and Thoning, K.W.: Atmospheric carbon dioxide dry air mole fractions from the NOAA/ESRL Carbon Cycle Global Cooperative Network, 1968-2016, version: 2017-07-28, ftp://ftp.cmdl.noaa.gov/data/trace_gases/co2/flask/surface/, 2017.

Deleted:

Deleted:

Deleted:

Deleted:

Deleted:

Deleted: .

Deleted: . Discuss.,

Deleted: :

Deleted: -200, in review

Crosson, E. R.: A cavity ring-down analyzer for measuring atmospheric levels of methane, carbon dioxide, and water vapor, *Appl. Phys. B-Lasers O.*, 92, 403–408, doi:10.1007/s00340-008-3135-y, 2008.

Devos, M., Patte, F., Rouault, J., Lafort, P., and Van Gemert, L.J.: Standardized Human Olfactory Thresholds. Oxford: IRL Press. p. 101, 1990.

Dlugokencky, E. J., Nisbet, E. G., Fisher R., and Lowry, D.: Global atmospheric methane: budget, changes and dangers, *Phil. Trans. R. Soc. A*, 369, 2058–2072, doi:10.1098/rsta.2010.0341, 2011.

Fisher R., Lowry, D., Wilkin, O., Sriskantharajah, S., and Nisbet, E. G.: High-precision, automated stable isotope analysis of atmospheric methane and carbon dioxide using continuous-flow isotope-ratio mass spectrometry, *Rapid Commun. Mass Spectrom.*, 20, 200–208, doi:10.1002/rcm.2300, 2006.

Flores, E., Viallon, J., Moussay, P., Griffith, D.W.T., and Wielgosz, R.I.: Calibration strategies for FT-IR and other isotope ratio infrared spectrometer instruments for accurate $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ measurements of CO_2 in air, *Anal. Chem.*, 89, 3648–3655, doi:10.1021/acs.analchem.6b05063, 2017.

GAW Report No. 269, 18th WMO/IAEA Meeting on Carbon Dioxide, Other Greenhouse Gases and Related Tracers Measurement Techniques (GGMT-2015), La Jolla, CA, USA, 2016.

[GAW Report No. 185, Guidelines for the measurement of methane and nitrous oxide and their quality assurance, World Meteorological Organization, Geneva, Switzerland, 2009.](#)

Griffith, D.W.T., Deutscher, N.M., Caldow, C., Kettlewell, G., Riggenbach, M., and Hammer, S.: A Fourier transform infrared trace gas and isotope analyser for atmospheric applications, *Atmos. Meas. Tech.*, 5, 2481–2498, doi:10.5194/amt-5-2481-2012, 2012.

[Hall, B.: Calculation of expanded uncertainties for CO2, NOAA ESRL, available at https://www.esrl.noaa.gov/gmd/cc1/cc1/uncertainties_co2.html.](#)

Kai, F. M., Tyler, S. C., Randerson, J. T., and Blake, D. R.: Reduced methane growth rate explained by decreased Northern Hemisphere microbial sources, *Nature*, 476, 194–197, 2011.

Karion, A., Sweeney, C., Pétron, G., Frost, G., Hardesty, R. M., Kofler, J., Miller, B. R., Newberger, T., Wolter, S., Banta, R., Brewer, A., Dlugokencky, E., Lang, P., Montzka, S. A., Schnell, R., Tans, P., Trainer, M., Zamora, R., and Conley, S.: Methane emissions estimate from airborne measurements over a western United States natural gas field, *Geophys. Res. Lett.*, 40, 4393–4397, doi:10.1002/frl.50811, 2013.

Deleted: https://

Deleted: .org/

Karion, A., Sweeney, C., Kort, E. A., Shepson, P. B., Brewer, A., Cambaliza, M., Conley, S. A., Davis, K., Deng, A., Hardesty, M., Herndon, S. C., Lauvaux, T., Lavoie, T., Lyon, D., Newberger, T., Pétron, G., Rella, C., Smith, M., Wolter, S., Yacovitch, T., and Tans, P.: Aircraft-based estimate of total methane emissions from the Barnett Shale region, *Environ. Sci. Technol.*, 49(13), doi:10.1021/acs.est.5b00217, 2015.

Keeling, C. D.: The Concentration and Isotopic Abundances of Carbon Dioxide in Rural and Marine Air, *Geochim. Cosmochim. Ac.*, 24, 277–298, 1961.

Mikaloff Fletcher, S. E., Tans, P. P., Bruhwiler, L. M., Miller, J. B., and Heimann, M.: CH₄ sources estimated from atmospheric observations of CH₄ and its ¹³C/¹²C isotopic ratios: 1. Inverse modeling of source processes, *Global Biogeochem. Cy.*, 18, GC4004, doi:10.1029/2004GB002223, 2004a.

Mikaloff Fletcher, S. E., Tans, P. P., Bruhwiler, L. M., Miller, J. B., and Heimann, M.: CH₄ sources estimated from atmospheric observations of CH₄ and its ¹³C/¹²C isotopic ratios: 2. Inverse modeling of CH₄ fluxes from geographical regions, *Global Biogeochem. Cy.*, 18, GC4005, doi:10.1029/2004GB002224, 2004b.

Molofsky, L.J., Connor, A.A., Farhat, S.K., Wylie Jr., A.S., and Wagner, T.: Methane in Pennsylvania water wells unrelated to Marcellus shale fracturing, *Oil Gas J.*, 109(19):54-67, 2011.

Montzka, S. A., Myers, R. C., Butler, J. H., Elkins, J. W., and Cummings, S.: Global tropospheric distribution and calibration scale of HCFC-22, *Geophys. Res. Lett.*, 20, 703–706, 1993.

Pataki, D. E., J. R. Ehleringer, L. B. Flanagan, D. Yakir, D. R. Bowling, C. J. Still, N. Buchmann, J. O. Kaplan, and J. A. Berry: The application and interpretation of Keeling plots in terrestrial carbon cycle research, *Global Biogeochem. Cycles*, 17, 1022, doi:10.1029/2001GB001850, 2003.

Peischl, J., Karion, A., Sweeney, C., Kort, E. A., Smith, M. L., Brandt, A. R., Yeskoo, T., Aikin, K. C., Conley, S. A., Gvakharia, A., Trainer, M., Wolter, S., and Ryerson, T. B.: Quantifying atmosphere methane emissions from oil and natural gas production in the Bakken shale region of North Dakota, *J. Atmos. Res. D.*, doi:10.1002/2015JD024631, 2016.

Pétron, G., Frost, G., Miller, B. R., Hirsch, A. I., Montzka, S. A., Karion, A., Trainer, M., Sweeney, C., Andrews, A. E., Miller, L., Kofler, J., Bar-Ilan, A., Dlugokencky, E. J., Patrick, L., Moore Jr., C. T., Ryerson, T. B., Siso, C., Kolodzey, W., Lang, P. M., Conway, T., Novelli, P., Masarie, K., Hall, B., Guenther, D., Kitzis, D., Miller, J., Welsh, D., Wolfe, D., Neff, W., and Tans, P.: Hydrocarbon emissions characterization in the Colorado Front Range: A pilot study, *J. Geophys. Res.*, 117, D04304, doi:10.1029/2011JD016360, 2012.

Deleted:

Deleted:

Deleted: 2011.

Pétron, G., Karion, A., Sweeney, C., Miller, B. R., Montzka, S. A., Frost, G. J., Trainer, M., Tans, P., Andrews, A., Kofler, J., Helmig, D., Guenther, D., Dlugokencky, E., Lang, P., Newberger, T., Wolter, A., Hall, B., Novelli, P., Brewer, A., Conley, S., Hardesty, M., Banta, R., White, A., Noone, D., Wolfe, D., and Schnell, R.: A new look at methane and nonmethane hydrocarbon emissions from oil and natural gas operations in the Colorado Denver-Julesburg Basin, *J. Geophys. Res.-Atmos.*, 119, 6836–6952, doi:10.1002/2013JD021272, 2014.

Rella, C. W., Hoffnagle, J., He, Y., and Tajima, S.: Local- and regional-scale measurements of CH₄, δ¹³CH₄, and C₂H₆ in the Uintah Basin using a mobile stable isotope analyzer, *Atmos. Meas. Tech.*, 8, 4539–4559, doi:10.5194/amt-8-4539-2015, 2015.

Richardson, S.J., Miles, N.L., Davis, K.J., Lauvaux, T., Martins, D., Turnbull, J., and Sweeney, C.: Tower measurement network of in-situ CO₂, CH₄, and CO in support of the Indianapolis FLUX (INFLUX) Experiment, *Elem. Sci. Anth.*, 5(59), doi.org/10.1525/elementa.140, 2017.

Röckmann, T., Eyer, S., van der Veen, C., Popa, M. E., Tuzson, B., Monteil, G., Houweling, S., Harris, E., Brunner, D., Fischer, H., Zazzeri, G., Lowry, D., Nisbet, E., Brand, W., Necki, J., Emmenegger, L., and Mohn, J.: In situ observations of the isotopic composition of methane at the Cabauw tall tower site, *Atmos. Chem. Phys.*, 16, 10469–10487, 2016 www.atmos-chem-phys.net/16/10469/2016/ doi:10.5194/acp-16-10469-2016, 2016.

Schwietzke, S., Sherwood, O. A., Bruhwiler, L. M. P., Miller, J. B., Etiope, G., Dlugokencky, E. J., Michel, S. E., Arling, V. A., Vaughn, B. H., White, J. W. C., Tans, P. P.: Upward revision of global fossil fuel methane emissions based on isotope database, *Nature*, 538, 88–91, doi:10.1038/nature19797, 2016.

Townsend-Small, A., Marrero, J.E., Lyon, D.R., Simpson, I.J., Meinardi, S., and Blake, D.R.: Integrating source apportionment tracers into a bottom-up inventory of methane emissions in the Barnett Shale hydraulic fracturing region, *Environ. Sci. Technol.*, 49, 8175–8182, doi:10.1021/acs.est.5b00057, 2015.

Turnbull, J., Guenther, D., Karion, A., Sweeney, C., Anderson, E., Andrews, A., Kofler, J., Miles, N., Newberger, T., Richardson, S., and Tans, P.: An integrated flask sample collection system for greenhouse gas measurements. *Atmos. Meas. Tech.*, 5, 2321–2327, doi:10.5194/amt-5-2321-2012, 2012.

Vaughn, B., Miller, J., Ferretti, D.F., and White, J.: Stable isotope measurements of atmospheric CO₂ and CH₄, in *Handbook of Stable Isotope Analytical Techniques*, 272–304, doi:10.1016/B978-044451114-0/50016-8, 2014.

Warneck P., and Williams, J.: *The Atmospheric Chemist's Companion: Numerical data for use in the atmospheric sciences*, Springer, doi:10.1007/978-94-007-2275-0, 2012.

Deleted: Elementa, in review, 2017.

Deleted: :/

1 Wen, X.-F., Meng, Y., Zhang, X.-Y., Sun, X.-M., and Lee, X.: Evaluating calibration strategies for isotope ratio
2 infrared spectroscopy for atmospheric $^{13}\text{CO}_2$ / $^{12}\text{CO}_2$ measurement, Atmos. Meas. Tech., 6, 1491–1501,
3 doi:10.5194/amt-6-1491-2013, 2013.▲

4
5 Zhao, C., and P.P. Tans: Estimating uncertainty of the WMO mole fraction scale for carbon dioxide in air, J. Geophys.
6 Res., 111, D08S09, doi:10.1029/2005JD006003, 2006.

Deleted: :/

Formatted: Font:Bold

Tables

Table 1. Field tanks used at the tower locations. The high and target tanks were used for the field calibration of $\delta^{13}\text{CH}_4$. Only the target tank is used for field adjustment of the CH_4 and CO_2 mole fraction calibration. The CH_4 and CO_2 mole fractions for the high and low tanks are less certain than that of the target tanks.

*Determined via laboratory measurements.

**NOAA/INSTAAR calibration (WMO X2004A scale for CH_4 and WMO X2007 for CO_2).

*** Field calibration – values not used.

Tank number	Deployment location	Measured isotopic ratio $\delta^{13}\text{CH}_4$ (‰)	CH_4 mole fraction (ppb)	CO_2 mole fraction (ppb)	Used for field calibration of $\delta^{13}\text{CH}_4$	Independent test of $\delta^{13}\text{CH}_4$ calibration	Used for field adjustment of CH_4 and CO_2 mole fraction calibration (intercept only)	Used for ethane correction
CA06418	North-High	-38.31*	9701*	397.75** *	✓			✓
CA05551	North-Low	-23.67*	1926.8*	402.70** *		✓		
CB10825	North-Target	-47.26**	1867.59* *	399.71**	✓		✓	✓
CA05419	Central-High	-38.48*	10534*	399.66** *	✓			✓
CA06438	Central-Low	-23.80*	2064.6*	397.82** *		✓		
CB10734	Central-Target	-47.25**	1878.53* *	397.09**	✓		✓	✓
CA05330	South-High	-38.68*	10152*	403.10** *	✓			✓
CC11499 9	South-Low	-23.72*	1999.2*	402.58** *		✓		
CB10727	South-Target	-47.24**	1868.33* *	399.68**	✓		✓	✓
CA06410	East-High	-38.52*	10414*	407.45** *	✓			✓

Deleted: calibration

Deleted: high and low tanks were planned to be used for the field calibration of $\delta^{13}\text{CH}_4$, with the target as an independent test. However, as described in Section 6.2 and indicated in the table, the

CA06357	East-Low	-24.02*	2079.7*	368.47** *		✓		
CB10718	East-Target	-47.26**	1867.94* *	399.67**	✓		✓	✓

Table 2. Maximum error estimate attributable to cross-interference due to direct absorption on $\delta^{13}\text{CH}_4$. These estimates were based on typical values for this tower-based application and estimated effects on CRDS measurements (Rella et al., 2015), and assumed 2 ppm ambient CH_4 mole fraction. For water vapor and carbon dioxide, the interferences are independent of CH_4 mole fraction for 1 – 15 ppm. For the other species listed, the interferences are inversely proportional to CH_4 mole fraction. Typical maximum values determined by flask^k (level at which 99 % of (afternoon) flask measurements at the South and East towers are below), by in-situ measurements at Marcellus towers^l, or by typical values^l (Warneck and Williams, 2012). ^aNo known ambient estimates (Barnes, 2015) / odor threshold (Devos et al., 1990).

Gas Species	Typical maximum value or range	Estimated maximum error
Carbon monoxide	Range ^f : 107.5-200.7 ppb	0.01‰
Water vapor, dried sample	Range ^l : 0.02 – 0.06%	0.02‰
Water vapor, ambient moisture	Range: 0 – 2.5%	±1‰ (Rella et al., 2015)
Carbon dioxide	Range ^l : 375 – 475 ppm	0.03‰
Propane	Max ^f 3.6 ppb	0.01‰
Butane (i-Butane + n-Butane)	Max ^f 1788 ppt	0.01‰
Ammonia	Typical ^l 90 ppt	0.01‰
Hydrogen sulfide	Typical ^l 30 ppt	0.01‰
Methyl mercaptan	Odor threshold ^a : 1 ppb	0.01‰
Ethylene	13.0 ^f ppt	0.01‰
Ethane	Max ^f 8.0 ppb (typical background ^l : 1.3 ppb)	0.23‰ (0.04‰ typical)

Deleted: ,

Deleted: assuming

Table 3. Results for the four Marcellus towers using two possible calibration schemes. Tank errors are shown for the using the high and low tank in the calibration (SCHEME A) and using the high and target tank in the calibration (SCHEME B). The third set of results are for SCHEME B, but following the change in field tank testing times on 3 Dec 2016. Results fare from October 2016 for the South, East and North towers, but are from May 2016 for the Central tower, as the analyzer was at the manufacturer for repairs during October 2016. Note that the daily means of the field tanks are used in the calibrations.

	Tower	High tank error (‰) mean ± standard deviation for one month (standard error)	Low tank error (‰) mean ± standard deviation for one month (standard error)	Target tank error (‰) mean ± standard deviation for one month (standard error)
SCHEME A	South	Used in cal	Used in cal	-0.3±0.4 (0.1)
SCHEME A	East	Used in cal	Used in cal	-0.8±0.5 (0.1)
SCHEME A	Central	Used in cal	Used in cal	-0.5±0.3 (0.1)
SCHEME A	North	Used in cal	Used in cal	-0.4±0.7 (0.1)
SCHEME B	South	Used in cal	0.2±0.7 (0.0)	Used in cal
SCHEME B	East	Used in cal	0.7±0.6 (0.0)	Used in cal
SCHEME B	Central	Used in cal	0.4±0.5 (0.0)	Used in cal
SCHEME B	North	Used in cal	0.3±1.3 (0.1)	Used in cal
(following change in field tank testing times on 3 December 2016)				
SCHEME B	South	Used in cal	0.3±0.3 (0.0)	Used in cal
SCHEME B	East	Used in cal	0.6±0.5 (0.0)	Used in cal
SCHEME B	Central	Used in cal	0.4±0.3 (0.0)	Used in cal
SCHEME B	North	Used in cal	-0.4±0.9 (0.0)	Used in cal

Deleted: Table 3. Results for the South tower for	...	[69]
Deleted: EXPT H		[70]
Deleted: EXPT D		
Deleted: EXPT D		
Deleted: EXPT D		
Formatted Table		[71]
Deleted: EXPT D-PRE-HI -T		
Formatted		[72]
Deleted: 0.2±0.7 (0.0)		
Formatted		[73]
Deleted: EXPT D-PRE-HI -T		
Formatted		[74]
Deleted: 0.7±0.6 (0.0)		
Formatted		[75]
Deleted: EXPT D-PRE-HI -T		
Formatted		[76]
Deleted: 0.5±0.5 (0.0)		
Formatted		[77]
Deleted: EXPT D-PRE-HI-T		
Formatted		[78]
Deleted:		
Deleted: 0.3±		
Formatted		[79]
Formatted		[80]
Deleted: EXPT D-HI-T		
Formatted		[81]
Formatted		[82]
Deleted: 0.2±		
Formatted		[83]
Deleted: 7		
Formatted		[84]
Formatted		[85]
Deleted: EXPT D-HI-T		
Formatted		[86]
Formatted		[87]
Deleted: 0.7±		
Formatted		[88]
Deleted: 6		
Formatted		[89]
Formatted		[90]
Deleted: EXPT D-HI-T		
Formatted		[91]
Formatted		[92]
Formatted		[93]
Deleted: 5		
Formatted		[94]
Formatted		[95]
Deleted: EXPT D-HI-T		
Formatted		[96]
Formatted		[97]
Deleted: 0.3±1.3 (
Formatted		[98]
Deleted: 1		
Formatted		[99]
Formatted		[100]

Table 4. Possible field tanks and sampling strategies, including those employed in the present study. The “Improved strategy” column suggests a possible strategy in which three field tanks and one independent tank are employed, and thus laboratory calibration is not required. Estimated tank testing times (excluding transition times) are listed for various compatibility requirements.

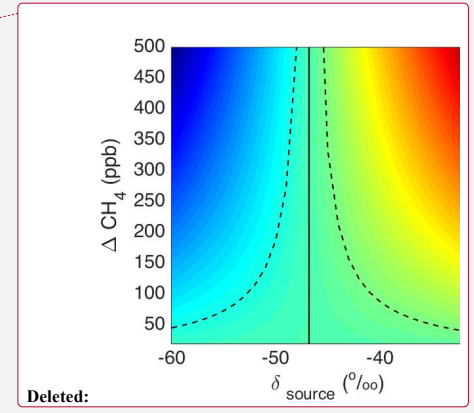
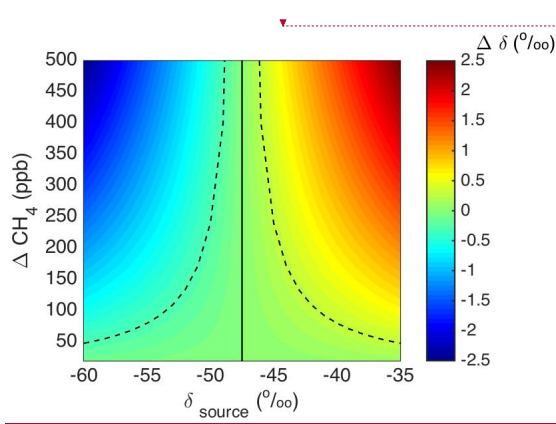
	Present study prior to 3 December 2016	Present study 3 December 2016 and thereafter	Improved strategy	
Laboratory calibration needed?	Yes, for linear calibration	Yes, for linear calibration	No	
High CH ₄ mole fraction tank(s)	HIGH (10 ppm, −38.3‰, 26 min/day)	HIGH (10 ppm, −38.3‰, 10 min/day)	HIGH (10 ppm, −38.3‰ to −44‰, 8 min/day, for 0.1‰ Allan deviations, 1 for 0.2 ‰, 1 for 0.4 ‰)	
	-	-	HIGH (10 ppm, −54.5‰ to −52‰, 8 min/day, for 0.1‰ Allan deviations, 1 for 0.2 ‰, 1 for 0.4 ‰)	
Low CH ₄ mole fraction tanks	LOW (2 ppm, −23.9‰, 81 min/day) independent	LOW (2 ppm, −23.9‰, 54 min/day) independent	LOW (2.1 ppm, −46.5‰ (ambient), 120 min/day, for 0.1‰ Allan deviations, 60 for 0.2 ‰, 8 for 0.4 ‰)	
	TARGET (2 ppm, −47.2‰, 6 min/day)	TARGET (2 ppm, −47.2‰, 54 min/day)	TARGET (1.9 ppm, −47.5‰ (ambient), 120 min/day, for 0.1‰ Allan deviations, 60 for 0.2 ‰, 8 for 0.4 ‰) independent	

Deleted: 5
Deleted: calibration
Deleted: Alternate
Deleted: two high
Deleted: to achieve the linear calibration for isotopic ratio
Deleted: The compatibility may be improved slightly using this strategy, as both the slope and intercept of the linear isotopic ratio calibration
Deleted: updated in the field. Furthermore, an independent tank near the ambient range would be preferable for evaluation of performance. In situations when the
Deleted: are not as stringent (e.g., >0.5 ‰), the strategy detailed in the last column may be appropriate. Sampling times listed exclude transition time between gases.
Formatted Table
Deleted: Alternate
Deleted: Strategy for reduced compatibility requirements (~0.5‰)
Deleted Cells
Deleted: and mole fraction correction
Deleted: No
Deleted: and mole fraction correction
Deleted: .
Deleted: ‰, 4
Deleted: HIGH (10 ppm, - ... [101]
Deleted:)
Deleted: .
Deleted: ‰, 4
Deleted: -
Deleted:)
Deleted: /SECONDARY TARGET
Deleted: near -47
Deleted: LOW (2 ppm, - ... [102]
Deleted: 64
Deleted:)
Deleted: 2
Deleted: near
Deleted: TARGET (2 ppm, - Advantage
Deleted: 64
Deleted:)

1
2
3

Notes		Reduced noise in calibration due to increased target tank sampling time	Does not necessarily require laboratory calibration of analyzers. Range of ideal isotopic ratios for the high tanks is given. Utilizing the isotopic ratios of commercially available bottles for spiking (i.e., 38.3‰ and –54.5‰) may avoid the need for laboratory calibration of these tanks. Using low/target tanks near ambient isotopic ratio range (but not exactly the same isotopic ratio, and preferably not exactly the same mole fraction) is more accurate reflection of compatibility and range of the isotopic ratio of the high tanks better encompasses expected values. For applications with reduced compatibility requirements (e.g., 0.4 ‰), utilizing low/target tanks at commercially available –38.3‰ and –54.5‰ may be sufficient. It is advantageous to distribute field tank testing throughout the day, to avoid not sampling ambient air for long periods and to measure potential changes in analyzer response.
--------------	--	---	--

Deleted: Advantage
Deleted: Does not require laboratory calibration. Low (secondary target) near ambient range is more accurate reflection of compatibility.
Deleted: Does not require laboratory calibration. Reduced sampling time of low and target tanks reduces gas usage. Easier to obtain tanks with these specified isotopic ratios.
Deleted Cells



Deleted:

- Deleted: enhancement above
- Deleted: .
- Formatted: Font color: Text 1, English (US)
- Formatted: Subscript
- Deleted: enhancement above

Figure 1. Isotopic ratio difference from background ($\Delta\delta$) resulting of a mixture of background and source signatures, as a function of source isotopic ratio (δ_{source}) and CH_4 mole fraction enhancement above background (ΔCH_4). Here the source end members are -60‰ and -35‰ . Background CH_4 mole fraction was assumed to be 2000 ppb and background isotopic ratio -47.5‰ (vertical solid line). Dashed lines indicate -0.3‰ and 0.3‰ difference from background.

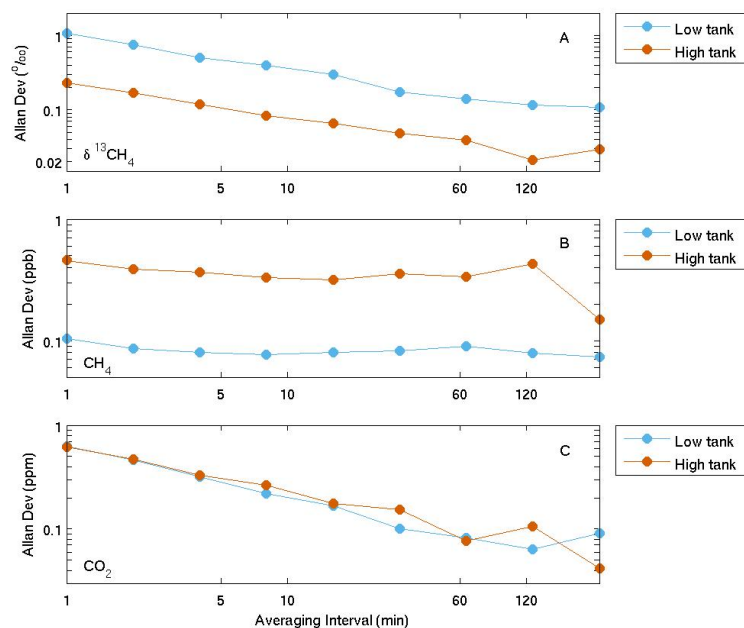


Figure 2. Allan standard deviation for (A) $\delta^{13}\text{CH}_4$, (B) CH_4 , and (C) CO_2 for a high CH_4 mole fraction tank (9.7 ppm CH_4 , ~400 ppm CO_2 , ~38.3 ‰ $\delta^{13}\text{CH}_4$) (orange) and a low (1.9 ppm CH_4 , ~400 ppm CO_2 , ~23.7 ‰ $\delta^{13}\text{CH}_4$) tank (blue). The x-axis is truncated to focus on minimum averaging times required to achieve the desired compatibility goals.

Deleted: Page Break
... [103]

Formatted: Don't keep with next
Moved (insertion) [39]
Formatted: Superscript
Formatted: Subscript

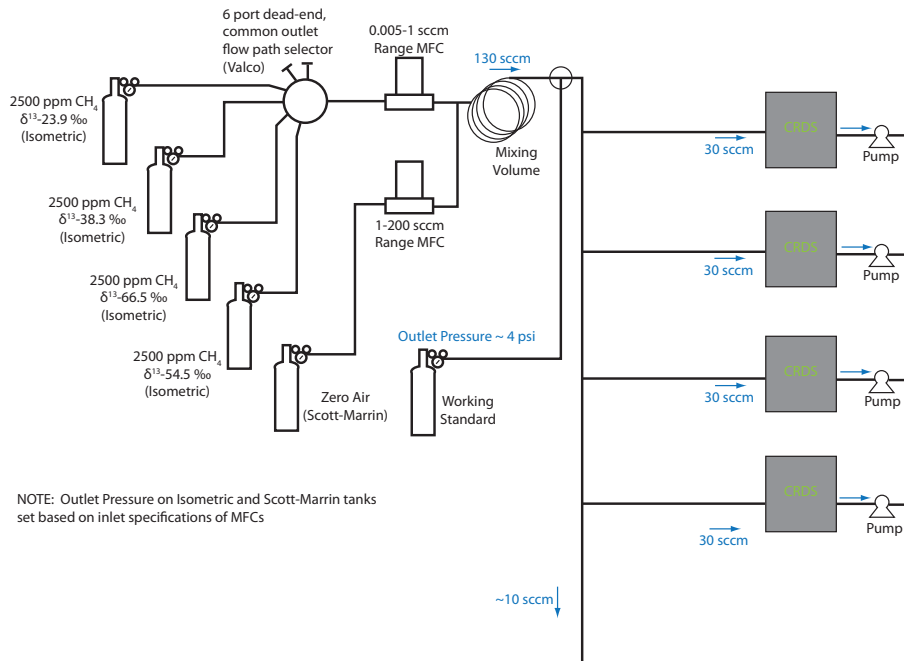


Figure 3. Flow diagram of the experimental setup used for the laboratory calibration of the analyzers and the field tanks (working standards). At standard pressure and temperature, the gas volume of the zero air and working standard tanks was 4021 L and that of the Isometric Instruments bottles was 28 L.

Formatted: Not Superscript/ Subscript

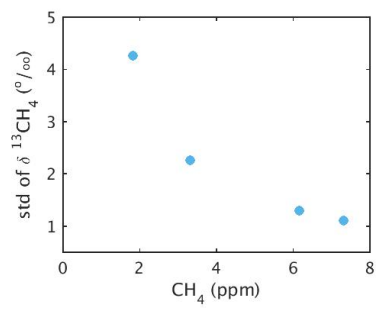
Formatted: Left

Deleted: .

Deleted: Page Break

... [104]

1
2
3
4
5
6
7
8
9



10
11
12

Figure 4. Standard deviation of the CH₄ isotopic ratio during the test results shown in Fig. 5.

Deleted: 3. Isotopic

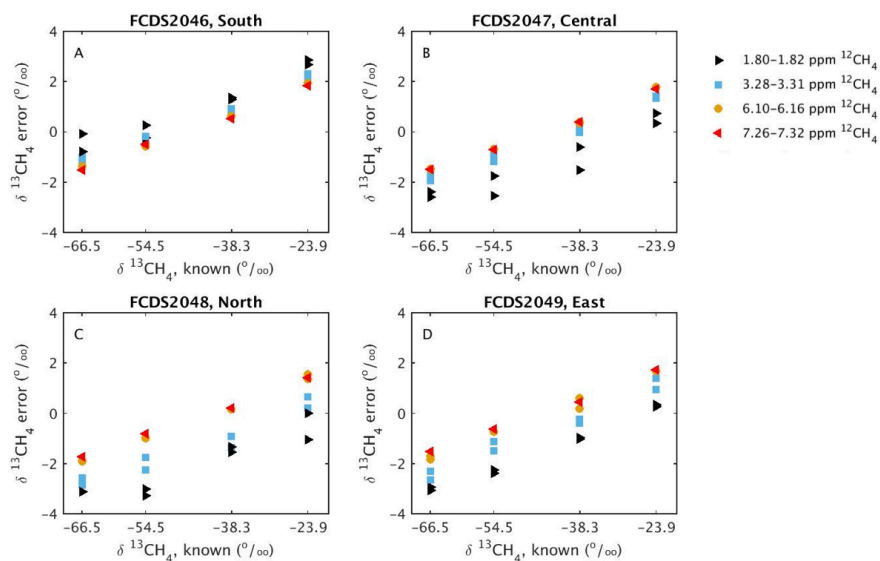


Figure 5. Measured isotopic ratio error as a function of known isotopic ratio for each of the four analyzers (A – D), prior to calibration. The colors indicate the $^{12}\text{CH}_4$ mole fraction, as shown in the legend. The serial numbers (FCDS2046, FCDS2047, FCDS2048, and FCDS2049) of the analyzers are indicated as well. These analyzers were deployed at the South, Central, North and East towers, respectively. Interpolating from the Allan standard deviation results (Fig. 2), the estimated precision is 0.40 ‰ for the 1.80–1.82 ppm CH_4 tests, 0.34 ‰ for 3.28–3.32 ppm CH_4 tests, 0.24 ‰ for 6.10–6.16 ppm CH_4 tests, and 0.20‰ for 7.26–7.32 ppm CH_4 tests. ▼

Deleted:).

Deleted: 7

Formatted: Subscript

Deleted: The precision could be improved by averaging over longer periods.

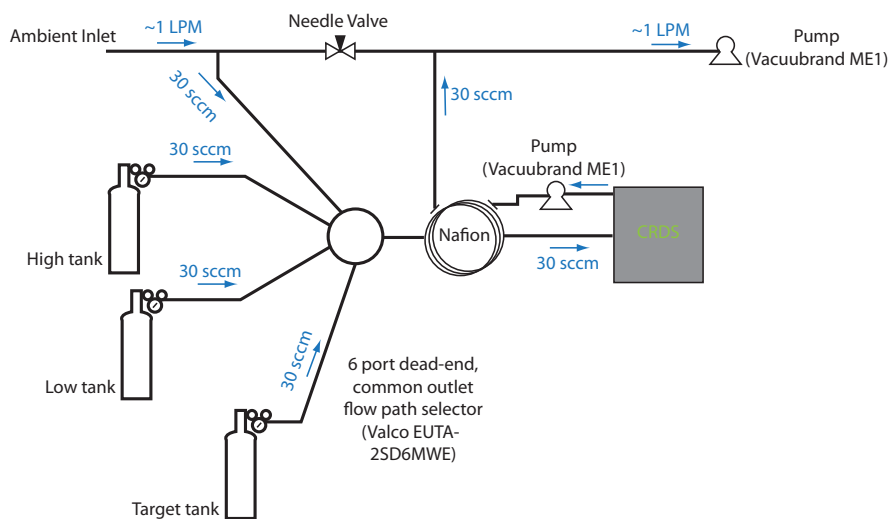
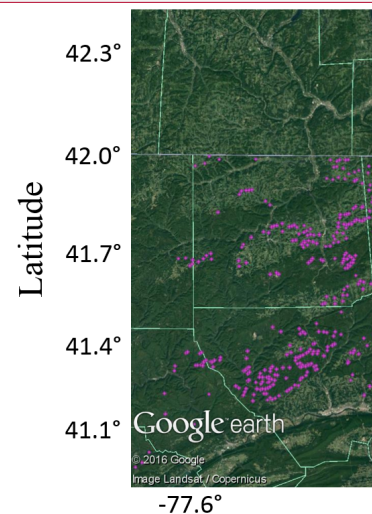


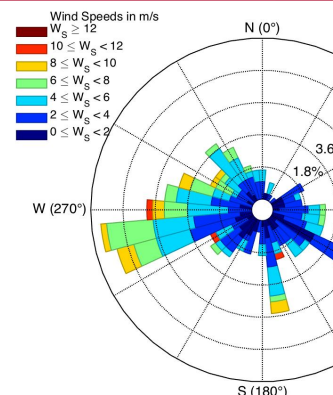
Figure 6. Flow diagram of the field calibration system. At standard pressure and temperature, the gas volume of the field tanks was 4021 L.



Deleted: ... [105]

Deleted: 4.

Moved down [40]: Map of Pennsylvania with permitted unconventional natural gas wells (magenta dots) and network of towers with methane and stable isotope analyzers (Picarro G2132-i). The East and South towers were also equipped with NOAA flask sampling systems. The Binghamton Airport is also indicated. ... [106]



Deleted: ... [107]

Moved down [41]: Wind rose for surface station at Binghamton, NY airport for the period April 2015 – April 2016 (using the mean of the afternoon hours for each day). The magnitude of wedges indicates relative frequency for each wind direction and the wind speeds are indicated by color.

Deleted: Calm winds (< 3 .

Deleted: m s^{-1}) are not categorized by direction. ... [108]

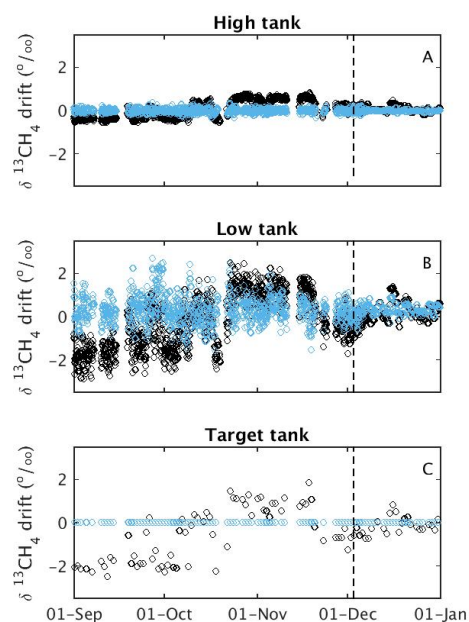
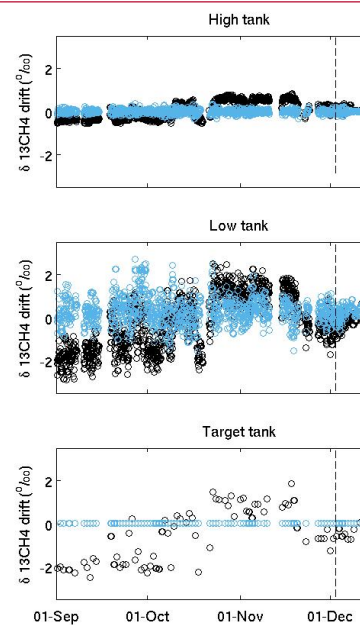


Figure 7. Results following isotopic ratio laboratory calibration only (black) and following calibration (blue) for the South tower for September - December 2016 for the “high” CH₄ mole fraction tank (A), “low” CH₄ mole fraction tank (B), and target tank (C). The target tank was used in the isotopic ratio calibration, whereas the low tank was independent. An improved calibration tank sampling strategy was implemented on 3 December 2016 (indicated by vertical dashed lines). The Allan deviation for time period used for each calibration cycle was, for the period prior to the improved tank sampling strategy, 0.2 ‰ for the high tank, and 0.5 ‰ for the low and target tanks. Following the implementation of the improved tank sampling strategy, the Allan deviation for each calibration cycle was 0.1 ‰ for the high tank, and 0.3 ‰ for the low and target tanks.

Moved up [39]: Allan standard deviation for (A) $\delta^{13}\text{CH}_4$, (B) CH_4 , and (C) CO_2 for a high CH_4 mole fraction tank (9.7 ppm CH_4 , ~400 ppm CO_2 , -38.3 ‰ $\delta^{13}\text{CH}_4$) (orange) and a low (1.9 ppm CH_4 , ~400 ppm CO_2 , -23.7 ‰ $\delta^{13}\text{CH}_4$) tank (blue). The x-axis is truncated to focus on minimum averaging times required to achieve the desired compatibility goals. .

Formatted: Not Superscript/ Subscript



Deleted:

Formatted: Superscript

Deleted: the EXPT-D-HI-T protocol

Formatted: Subscript

Deleted: low

Deleted: independent of

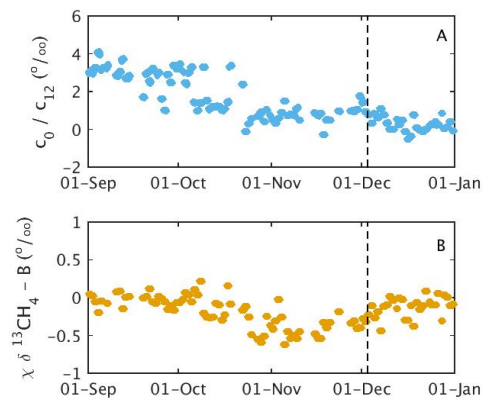
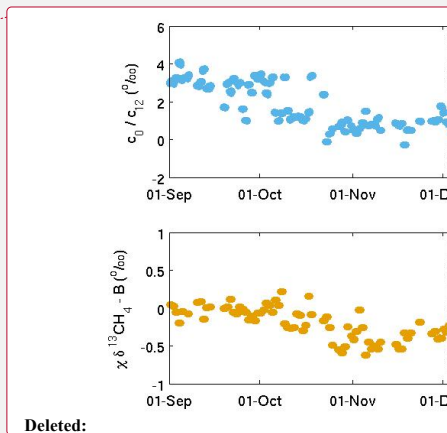


Figure 8. Effect of each of the calibration coefficient terms for the South tower for September - December 2016 for the optimized calibration scheme. The terms c_0 (A) and χ (B) in Eq. (3) are time-dependent drift terms. Note the differing scales. An improved calibration tank sampling strategy was implemented on 3 December 2016 (indicated by vertical dashed lines).



Deleted:

Deleted: 9

Deleted: These parameters vary because of spectral variations in the optical loss of the empty cavity (c_0), and because of errors in the temperature or pressure of the gas, or changes in the wavelength calibration (χ).

Deleted: Page Break

... [111]

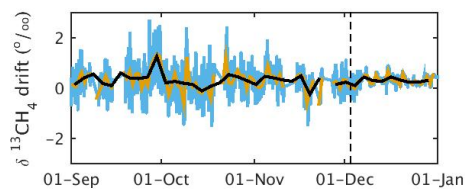


Figure 9. Low tank methane isotopic ratio differences from known value, for the individual calibration cycles (blue), and for 1-day (red) and 3-day (black) means of the calibration cycles, for the South tower for September – December 2016. An improved calibration tank sampling strategy was implemented on 3 December 2016 (indicated by the vertical dashed line). The low tank is independent of the isotopic ratio calibration.

Deleted: 10

Deleted: Page Break
 ... [112]

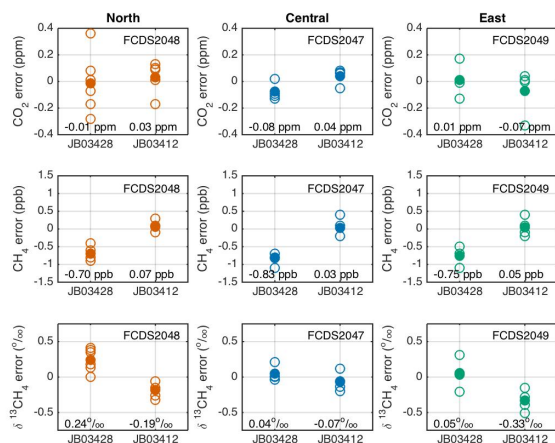


Figure 10. Results from round-robin style testing using two NOAA/INSTAAR tanks (JB03428: $-46.82 \text{ ‰ } \delta^{13}\text{CH}_4$, 1895.3 ppb CH_4 and 381.63 ppm CO_2 ; and JB03412: $-45.29 \text{ ‰ } \delta^{13}\text{CH}_4$, 2385.2 ppb CH_4 and 432.71 ppm CO_2) for CO_2 (top row), CH_4 (middle row), and $\delta^{13}\text{CH}_4$ (bottom row), for the analyzer deployed at the North tower (serial number FCDS2048; left column), at the Central tower (serial number FCDS2047; middle column), and at the East Tower (serial number FCDS2049; right column). These tests were completed in the laboratory, post deployment (March 2017). The analyzer deployed at the South tower (serial number FCDS2046) was not included in these tests. Open circles are individual tests and filled circles are the means of the individual tests for each analyzer/constituent. The mean error for each analyzer/tank/constituent is indicated in the plots.

Deleted: 11

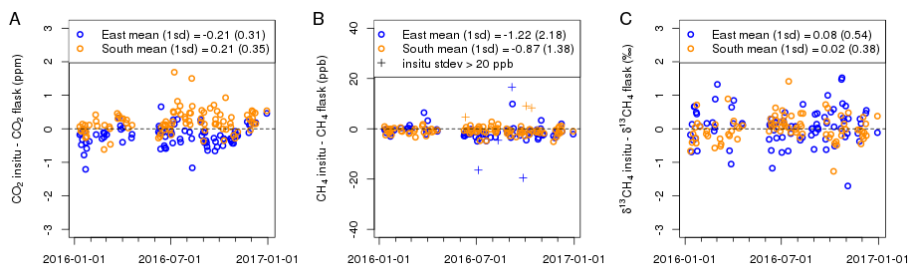


Figure 11. Afternoon in-situ to flask differences for January – December 2016 for the East (blue) and the South towers (orange) for A) CO_2 , B) CH_4 , and C) $\delta^{13}\text{CH}_4$. For CH_4 , data points with high temporal variability (standard deviation of raw ~2sec data within the 10-min segments > 20 ppb) are indicated by '+' symbols and have been excluded. The standard deviation of the in-situ to flask differences are shown in parentheses on each plot. The standard errors, indicating an estimate of how far the sample mean is likely to be from the true mean, is 0.24 ppb CH_4 , 0.03 ppm CO_2 and 0.06 ‰ at the East tower and 0.14 ppb CH_4 , 0.04 ppm CO_2 and 0.04 ‰ at the South tower.

Moved (insertion) [42]

Deleted: Page Break

Moved down [43]: Figure 12.

Moved (insertion) [44]

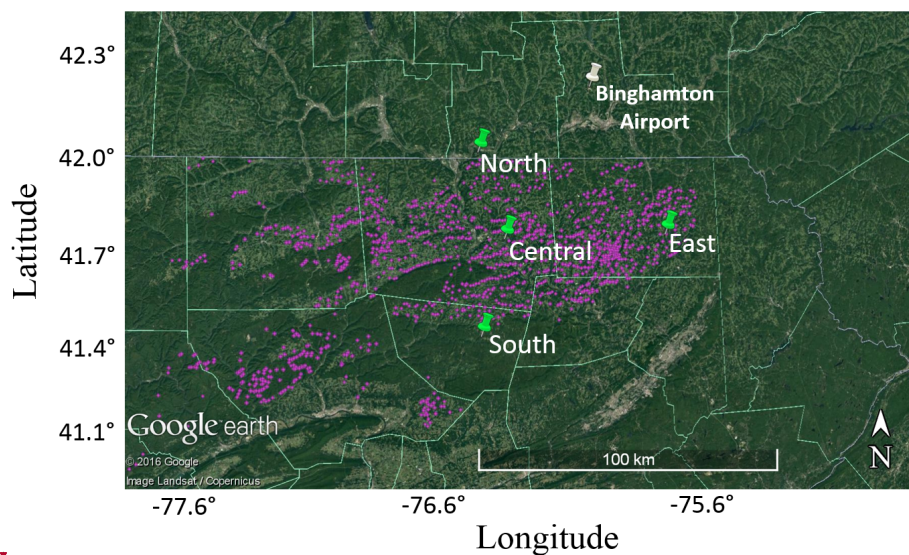


Figure 12. Map of Pennsylvania with permitted unconventional natural gas wells (magenta dots) and network of towers with methane and stable isotope analyzers (Picarro G2132-i). The East and South towers were also equipped with NOAA flask sampling systems. The Binghamton Airport is also indicated.

Moved up [42]: Afternoon in-situ to flask differences for January – December 2016 for the East (blue) and the South towers (orange) for A) CO₂, B) CH₄, and C) $\delta^{13}\text{CH}_4$. For CH₄, data points with high temporal variability (standard deviation of raw ~2sec data within the

Deleted: hour > 20 ppb) are indicated by '+' symbols and have been excluded.

Moved up [44]: The standard deviation of the in-situ to flask differences are shown in parentheses on each plot. The standard errors, indicating an estimate of how far the sample mean is likely to be from the true mean, is 0.24 ppb CH₄, 0.03 ppm CO₂ and 0.06 ‰ at the East tower and 0.14 ppb CH₄, 0.04 ppm CO₂ and 0.04 ‰ at the South tower.

... [114]

Deleted: -

... [115]

Moved (insertion) [43]

Moved (insertion) [40]

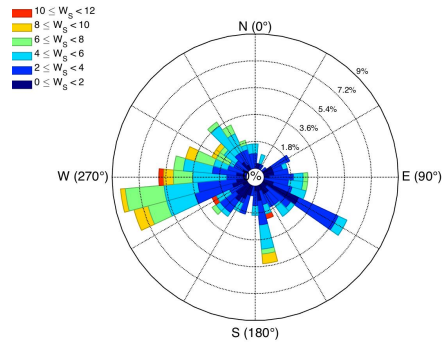


Figure 13. Wind rose for surface station at Binghamton, NY airport for the period April 2015 – April 2016 (using the mean of the afternoon hours for each day). The magnitude of wedges indicates relative frequency for each wind direction and the wind speeds are indicated by color. These afternoon means were based on hourly reported measurements. For the hourly measurements, calm winds ($\leq 1.6 \text{ m s}^{-1}$) were not categorized by direction and thus were not included in the afternoon mean. For the hourly measurements, calm winds ($\leq 1.6 \text{ m s}^{-1}$) were reported as zero and were included in the afternoon mean.

Moved (insertion) [41]

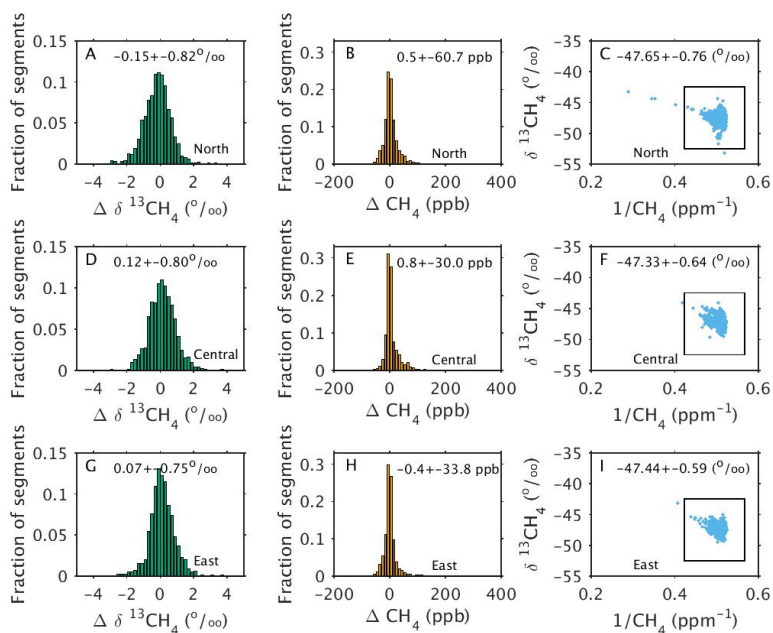


Figure 14. Probability distribution function of measured isotopic ratio differences from the background South tower ($\Delta \delta^{13}\text{CH}_4$) for the A) North, D) Central, and G) East towers for afternoon hours (1700–2059 UTC, 1200–1559 LST). The averaging interval of the individual data points for all plots is 10 min and the time period is January – May 2016. The bin size for A), D) and G) is 0.2 ‰. The median and standard deviation of the differences are indicated on the plots. Probability distribution function of measured methane mole fraction enhancements (ΔCH_4) for the B) North, E) Central, and H) East towers. Note that the scale for B), E), and H) has been truncated to focus on majority of the data points. The bin size is 10 ppb CH_4 . Keeling plots for the C) North, F) Central, and I) East towers. The black box in each plot indicates the approximate scale of the corresponding isotopic ratio difference and methane mole fraction enhancement plots. The median and standard deviation of the isotopic ratios at each tower are indicated on the plots. Note that the Allan deviation for 10-min means at ambient mole fractions was 0.4 ‰ and this decreases with increasing mole fraction.

Deleted: enhancement above

Deleted: time scale

Deleted: enhancement

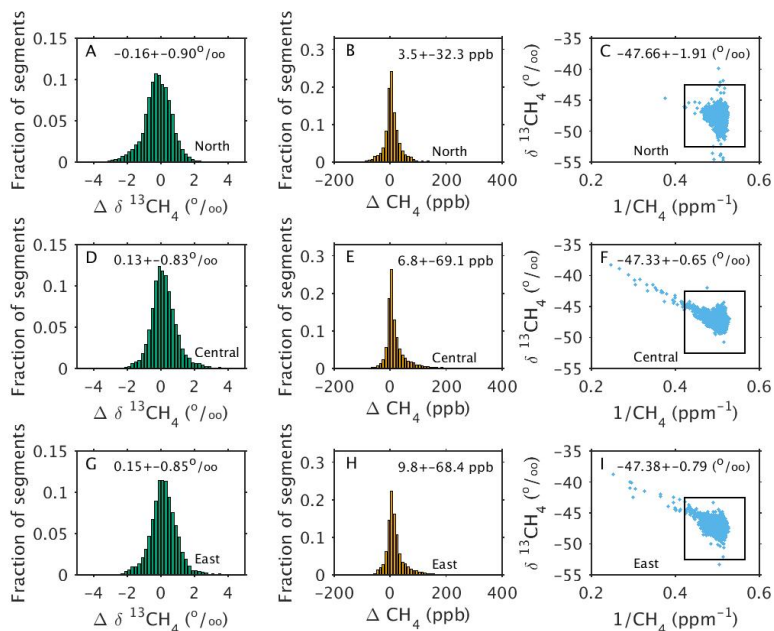
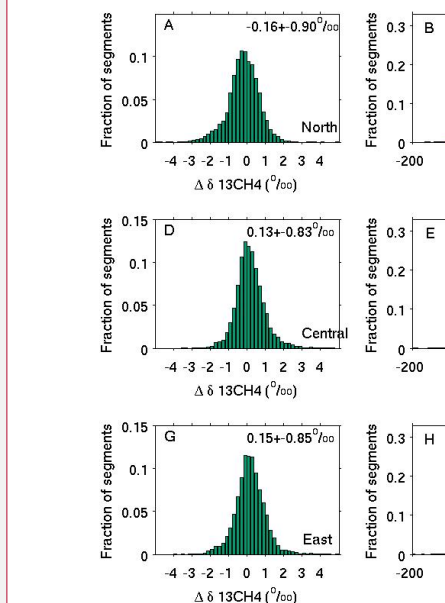


Figure 15. Probability distribution function of measured isotopic ratio differences from the background South tower ($\Delta \delta^{13}\text{CH}_4$) for the A) North, D) Central, and G) East towers for all times of data excluding the afternoon hours shown in Fig. 14. The averaging interval of the individual data points for all plots is 10 min and the time period is January – May 2016. The bin size for A), D) and G) is 0.2 ‰. The median and standard deviation of the differences are indicated on the plots. Probability distribution function of methane mole fraction enhancements (ΔCH_4) for the B) North, E) Central, and H) East towers. Note that the scale for B), E), and H) has been truncated to focus on majority of the data points. The bin size is 10 ppb CH_4 . Keeling plots for the C) North, F) Central, and I) East towers. The black box in each plot indicates the approximate scale of the corresponding isotopic ratio difference and methane mole fraction enhancement plots. The median and standard deviation of the isotopic ratios at each tower are indicated on the plots. Note that the Allan deviation for 10-min means at ambient mole fractions was 0.4 ‰ and this decreases with increasing mole fraction.



Deleted:

Deleted: 14

Deleted: enhancement above

Deleted: 13

Deleted: time scale

Deleted: enhancement

Formatted: Caption, Justified

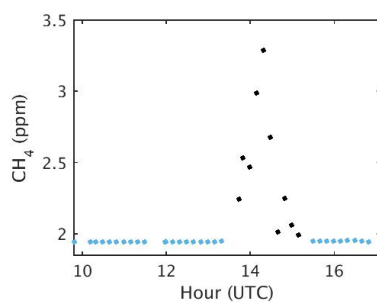
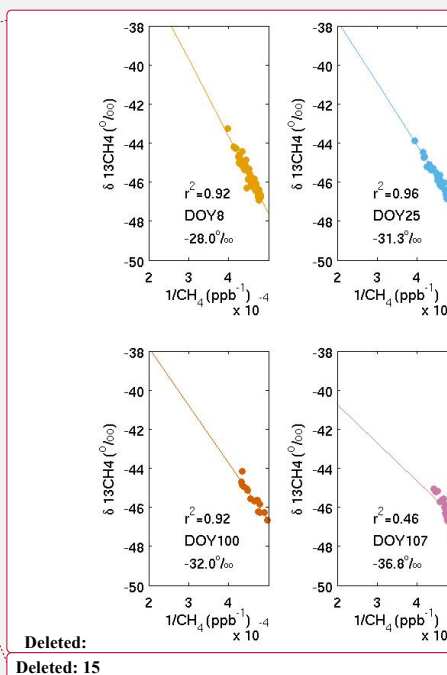


Figure 16. Time series of CH_4 encompassing one of the eight peaks in CH_4 at the Central tower (DOY 55) for which the Keeling plot approach was applied. The averaging interval of the individual points was 10 min, and periods during which field tanks were sampled were excluded from the plot. The linear fit was calculated using the points clearly within the plume (black dots).



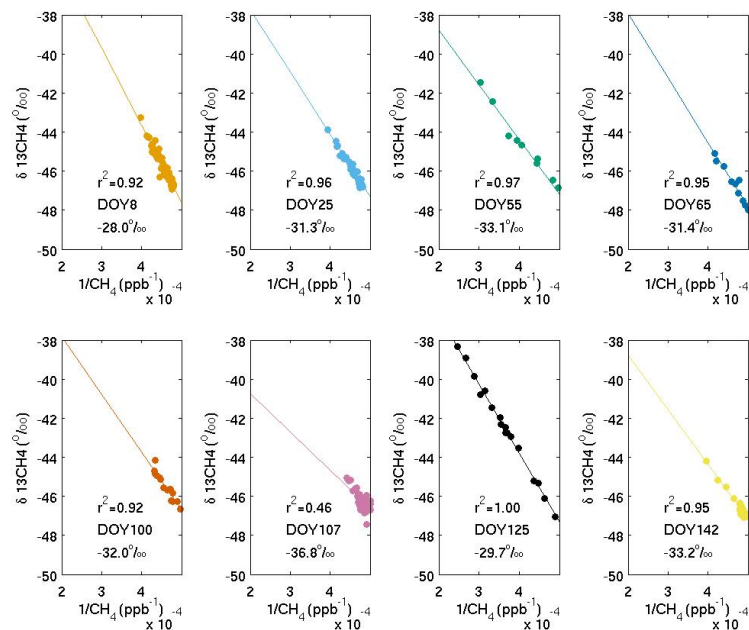


Figure 17. Keeling plots for the Central tower for the eight largest peaks in the non-afternoon methane time series. Black lines indicate the best-fit lines. Correlation coefficients (r^2), day of year (DOY) and y-intercepts are indicated in the plots. FIX subscripts.

Deleted: peak

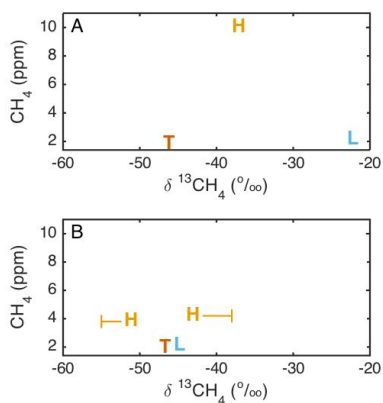
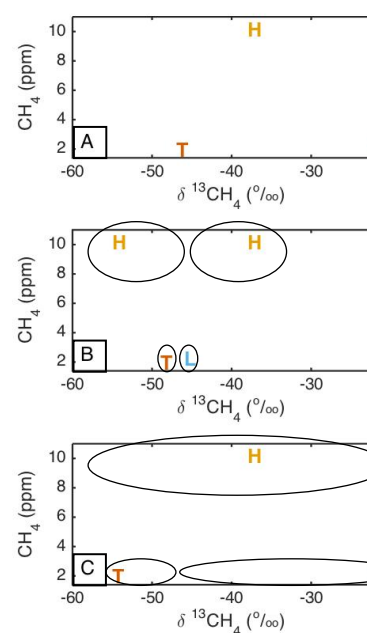


Figure 18. Graphical representation of the field tanks used in the present study (A), and for an improved strategy (as in Table 4) (B). Orange ‘H’ symbols indicate high mole fraction tanks, blue ‘L’ symbols indicate low mole fraction tanks, and red ‘T’ symbols indicate target tanks. Lines in (B) indicate range of isotopic values desirable for the high tanks. Ideally, calibration tanks are near the range of ambient values to be measured, but in this case, specific values are more easily obtained (–54.5 ‰, –38.3 ‰, –23.9 ‰, from Isometric Instruments, Inc.)



Deleted:

Deleted: 16

Deleted: calibration

Deleted: alternate

Deleted: 5

Deleted:), and for a strategy for reduced compatibility requirements (C

Deleted: , as

Deleted: Table 5. In A), the low tank is independent, but in

Deleted: and C), the target tank is independent. A “Target” tank is commonly used to

Deleted: an independent tank. The ovals around the tanks in B) and C) indicate

Deleted: ranges of values.

An alternative calibration approach is

2017).

Here we

In this paper, we describe a network of four tower-based atmospheric observation locations, measuring CH₄ and CO₂ dry mole fractions and $\delta^{13}\text{CH}_4$ using CRDS (Picarro, Inc., model G2132-i) analyzers in the Marcellus shale region in north-central Pennsylvania

3 Methods: Laboratory testing

3.1

.1 Study area

Four CRDS isotopic CH₄ analyzers (G2132-i, Picarro, Inc.) were deployed on commercial towers 46–61 m AGL in northeast Pennsylvania (Fig.

2012), measuring a suite of > 55 gases (including greenhouse gases, hydrocarbons, and halocarbons) and $\delta^{13}\text{CH}_4$.

4.2

The sampling scheme and procedure for using these field calibration tanks at each tower to correct the ambient $\delta^{13}\text{CH}_4$ measurements is described in Sections 3.2.2 and 6.2.

4.3

Page 17: [8] Moved from page 8 (Move #10)	Microsoft Office User	1/24/18 10:53:00 AM
---	-----------------------	---------------------

.1 Study area

Four CRDS isotopic CH_4 analyzers (G2132-i, Picarro, Inc.) were deployed on commercial towers 46–61 m AGL in northeast Pennsylvania (Fig.

Page 17: [9] Moved from page 8 (Move #11)	Microsoft Office User	1/24/18 10:53:00 AM
---	-----------------------	---------------------

). The South and North towers were located on the southern and northern edges of the unconventional gas well region, respectively, and were intended to measure background values depending on the wind direction.

Page 17: [9] Moved from page 8 (Move #11)	Microsoft Office User	1/24/18 10:53:00 AM
---	-----------------------	---------------------

). The South and North towers were located on the southern and northern edges of the unconventional gas well region, respectively, and were intended to measure background values depending on the wind direction.

Page 17: [10] Deleted	Microsoft Office User	1/24/18 10:53:00 AM
-----------------------	-----------------------	---------------------

5

Page 17: [10] Deleted	Microsoft Office User	1/24/18 10:53:00 AM
-----------------------	-----------------------	---------------------

5

Page 17: [10] Deleted	Microsoft Office User	1/24/18 10:53:00 AM
-----------------------	-----------------------	---------------------

5

Page 17: [10] Deleted	Microsoft Office User	1/24/18 10:53:00 AM
-----------------------	-----------------------	---------------------

5

Keeling plots (Keeling 1961; Röckmann et al., 2016) are often used to infer the isotopic ratio of the methane source as the intercept of the best fit line of the isotopic ratio as a function of the inverse methane mole fraction. In Section 6.7 we used this approach to estimate the source isotopic ratio of peaks observed during non-afternoon hours at the Central tower.

5 Methods for evaluating compatibility of in-situ tower measurements

5.1 Independent low tank

While the low tank was planned to be used in the calibration of the isotopic ratio, the optimized calibration scheme given the deployed tanks instead utilized the target tank in the calibration and kept the low tank as independent (Section 6.2). The low tank was thus treated as an ambient sample. To evaluate the noise in the calibrated ambient samples that results from noise in the calibration, we calculated the standard deviation over the period September 1 – December 2 of the individual calibration cycles (6 min each), of the calibration cycles averaged over 1 day (81 min total), and of the calibration cycles averaged over 3 days (4.1 hours total). These results are a proxy for the noise in the calibrated ambient samples over those sampling periods. The same calculation was performed for the period December 3 – December 31, a period during which an improved calibration tank sampling scheme was utilized.

5.2 Round-robin testing

Post-deployment round-robin style tests were completed in the laboratory in March 2017 for the analyzers previously deployed at the North, Central and East Towers, in order to assess the compatibility achievable via our calibration method.

The analyzer deployed at the South tower was not included in these tests, as it was still in the field. Two NOAA/INSTAAR tanks (JB03428: -46.82‰ $\delta^{13}\text{CH}_4$, 1895.3 ppb CH_4 and 381.63 ppm CO_2 ; and JB03412: -45.29‰ $\delta^{13}\text{CH}_4$, 2385.2 ppb CH_4 and 432.71 ppm CO_2) were

sampled for 70 min, with 8 min ignored after each transition, and treated as unknowns. Additionally, high, low, and target tanks were sampled, with the calibration applied as in the field for ambient samples (as described in Section 6.2). The high mole fraction tank was sampled for 20 min and the low and target mole fraction tanks were sampled for 70 min, with 8 min ignored after each gas transition.

Four to six tests were completed for each analyzer. We used these tests as a means of evaluating the compatibility of the analyzers, in terms of both mole fractions and the isotopic ratio.

5.3 Side-by-side testing

The precision and drift characteristics are not optimized for CO₂ for the G2132-i analyzers, compared to the G2301 analyzers, which measure CO₂, CH₄ and H₂O mole fractions. To test the performance of the G2132-i analyzers for consideration of the data for use as part of the continental-scale CO₂ network, G2301 and G2132-i (Picarro, Inc.)

analyzers were run side-by-side for one month (June 2016) at the South tower. The sampling system for the G2132-i was as described in Section

4.3. A separate ¼" (0.64 cm) tubing was used for the G2301 analyzer and an intercept calibration using the target tank is applied daily. The sample air for the G2301 analyzer was not dried and the on-board water vapor correction was used. This testing was used to evaluate the mole fraction compatibility, particularly for CO₂, of the G2132-i analyzers compared to the G2301 analyzers.

5.4 Flask measurements

Flask measurements were used for independent validation and error estimation of the continuous CO₂, CH₄ and $\delta^{13}\text{CH}_4$ in-situ measurements. In addition, the flasks were measured for a suite of species including N₂O, SF₆, CO, H₂ (Conway et al., 2011

), halo- and hydro-carbons (Montzka et al., 1993) and stable isotopes of CH₄ (Vaughn et al., 2004).

The flasks were filled over a 1-hour time period in the late

afternoon (1400–1500 LST), thereby yielding a more representative measurement compared to most flask sampling systems, which collect nearly instantaneous samples (e.g., ~10 sec).

Samples were measured only when winds were blowing steadily out of the west or north (~45–225°) to ensure that the samples are sensitive to and representative of the broader Marcellus shale gas production region that is the focus of this study.

6 Results

6.1 Allan standard deviation results

Page 17: [22] Moved to page 13 (Move #14)	Microsoft Office User	1/24/18 10:53:00 AM
---	-----------------------	---------------------

As

Page 17: [23] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
-------------------------	-----------------------	---------------------

Font color: Auto, English (US)

Page 17: [24] Deleted	Microsoft Office User	1/24/18 10:53:00 AM
-----------------------	-----------------------	---------------------

described in Section 3.1, two tanks were sampled for 24 hours each to determine the Allan standard deviation as a function of averaging interval for the G2132-i analyzers.

Page 17: [25] Moved to page 6 (Move #4)	Microsoft Office User	1/24/18 10:53:00 AM
---	-----------------------	---------------------

The resulting Allan standard deviations for $\delta^{13}\text{CH}_4$, CH_4 and CO_2 are shown in Fig.

Page 17: [26] Moved to page 6 (Move #5)	Microsoft Office User	1/24/18 10:53:00 AM
---	-----------------------	---------------------

. For the high tank, the Allan deviation for $\delta^{13}\text{CH}_4$ (Fig.

Page 17: [27] Moved to page 6 (Move #6)	Microsoft Office User	1/24/18 10:53:00 AM
---	-----------------------	---------------------

) was $< 0.2 \text{ ‰}$ (our target compatibility) for an averaging interval of 2 min (the averaging interval used each field calibration cycle of the high tank). To reduce the noise to $< 0.1 \text{ ‰}$, an averaging interval of 4 min is sufficient (in addition to the time required for the transition between gases). For the low tank, in order for the Allan standard deviation to be $< 0.2 \text{ ‰}$, 32 min were required and 64 min for 0.1 ‰ noise.

Page 17: [27] Moved to page 6 (Move #6)	Microsoft Office User	1/24/18 10:53:00 AM
---	-----------------------	---------------------

) was $< 0.2 \text{ ‰}$ (our target compatibility) for an averaging interval of 2 min (the averaging interval used each field calibration cycle of the high tank). To reduce the noise to $< 0.1 \text{ ‰}$, an averaging interval of 4 min is sufficient (in addition to the time required for the transition between gases). For the low tank, in order for the Allan standard deviation to be $< 0.2 \text{ ‰}$, 32 min were required and 64 min for 0.1 ‰ noise.

Page 17: [28] Deleted	Microsoft Office User	1/24/18 10:53:00 AM
-----------------------	-----------------------	---------------------

7B), both the high and low tank Allan deviation was $< 1 \text{ ppb}$ for even a 1-min averaging interval.

Page 17: [29] Moved to page 6 (Move #8)	Microsoft Office User	1/24/18 10:53:00 AM
---	-----------------------	---------------------

The CO_2 levels in the high and low tanks were similar ($\sim 400 \text{ ppm}$), and an averaging interval of 6 min corresponded to Allan standard deviations of 0.3 ppm , and 64 min were necessary for 0.1 ppm (Fig.

Page 17: [30] Deleted	Microsoft Office User	1/24/18 10:53:00 AM
-----------------------	-----------------------	---------------------

7C).

While in some cases, analyzer drift may contribute to increased Allan standard deviations as the averaging interval increases, we do not see evidence of this for the averaging intervals shown.

6.2 Optimized calibration scheme determination

As described in Section 4.4, several calibration strategy “experiments” were conducted using the data from October 2016 for the South tower to determine the optimized procedure for utilizing the field calibration data for isotopic methane. The results from these experiments (EXPTs) are shown in Table 3. In EXPTs A–G, the daily average of the 13.5 calibration cycles per day for the high tank and low tank was used (if applicable), and the target was independent of the calibration. Of the EXPTs A–G, EXPT D (laboratory calibration, and high and low tanks for mole fraction correction) gave the results with the lowest bias and standard deviations of the difference from the known value: -0.3 ± 0.4 ‰ for the target tank. The standard error for EXPT D was 0.1 ‰. Therefore, the mean was significantly improved in EXPT D, compared to the EXPTs A–C and G, with biases ranging from -2.2 to -0.5 ‰. EXPTs E and F showed comparable bias, but larger standard deviations.

With only the factory calibration (EXPT A), the bias of the values averaged over the month was -0.6 ‰ ± 1.3 ‰ for the target tank. Here the standard deviation was calculated for all of the calibration cycles over the month, e.g., 67 % of the calibrations cycles yielded differences from known values between -1.9 and 0.7 ‰ for the target tank using EXPT A. Without applying any slope calibration either in the laboratory or in the field, but applying a mole fraction correction (EXPT B), the standard deviation was improved to 0.4 ‰, but the bias for the target tank was increased to -2.2 ‰. With a linear calibration in the field, but no laboratory or mole fraction calibration (EXPT C), the mean bias for the target tank was improved to -0.5 ‰, but the standard deviation of the calibration cycles was high (1.4 ‰).

EXPT E used only the laboratory calibration (completed prior to deployment of the analyzers). The bias for the target tank was the same as for EXPT D, but the variation of the errors was larger, with a standard deviation of 1.2 ‰. We thus answer question 1) in Section 4.4: field calibration of the isotopic ratio significantly improved the variation in the tank differences from the known value over the month.

In EXPT F, neither the low nor the target tank were used in the field calibration of the isotopic ratio, and a field calibration of the mole fraction correction was not applied. The bias and standard deviation of the calibration cycles for the target tank were larger (-0.4 ± 0.9 ‰) than for EXPT D, indicating that field calibration of the mole fraction correction (and thus having both high and low field calibration tanks) was beneficial (question 2 in Section 4.4).

EXPT G, with no laboratory calibration, using both the high and low tanks to apply a linear calibration, and to apply the mole fraction correction yielded results that indicated larger bias and standard deviation in the target tank (-0.5 ± 1.4 ‰) compared to EXPT D. Thus, using the low tank in the application of the linear calibration was not beneficial in this case compared to using the laboratory calibration (question 3 in Section 4.4). It did, however, significantly improve results compared to the case with no laboratory calibration of the linear term (i.e., EXPTs A and B). Following EXPT G but using two, instead of one, high mole fraction tanks (with different isotopic ratios) for the linear calibration would likely yield similar results to EXPT D, and would require an additional field calibration tank (~ 10

ppm and -23.9% , for example), but has the advantage of not requiring a laboratory calibration. We note that in the present case, we used two measurements to solve for three unknowns: p_1 , c_0 and χ . They were thus not independent.

In EXPT D1, only the first high and low tank calibration cycle of the day was used, rather than the average of the 13.5 daily calibration cycles. The results, in terms of the one-month mean bias and standard deviation were similar to EXPT D, using the mean of all 13.5 daily calibration cycles. In EXPT D-T, only the high and low tank calibration cycle immediately preceding the target tank sampling was used. Again, the results were similar to EXPT D, indicating that the mean results over the month are not significantly affected by using only one calibration cycle per day for the field calibration, thus answering question 4 in Section 4.4 – the analyzers did not appear to drift significantly over several hours and thus multiple calibration cycles within each day are not required.

We now expand the analysis of the best performing calibration scheme (EXPT D and variations) to the other towers (Table 4). First, we show the results for EXPT H, with only the laboratory calibration prior to deployment, for comparison. The bias of the differences from known values varied amongst the towers, with magnitudes ranging from 0.0 to 1.9 ‰. For EXPT D, the bias of the one-month data of the target tank varied between -0.3 and -0.8% among the towers. The target tank contained typically ambient levels of isotopic methane, and we want to minimize biases in this range.

Since we found that the difference between EXPT D and its variants was not significant, only one calibration cycle per day was necessary (although longer sampling would be preferable based on Allan deviation results). So instead of using the low tank for the calibration and keeping the target tank independent, we instead used the target tank for the calibration and keep the low tank independent. EXPT D-HI-PRE-HI-T was similar to EXPT D, but used only the high tank value immediately preceding the target tank sampling for the day and used the target tank for the mole fraction correction. The results indicate that the center of the maximum bias was shifted from the target tank (about -47.2%) to the low tank (about -23.9%). This method is preferable since the ambient sample is near the target tank values. Using the daily average for the high instead of just one calibration cycle (EXPT D-HI-T) indicates similar results. We thus chose to use the EXPT-D-HI-T protocol, but sampling the high and low tanks every 90 minutes does not appear to be necessary – using the daily average is sufficient, i.e., the analyzers do not appear to drift significantly in the period of one day.

As an example, the tank results (differences from known values) using only the laboratory calibration for isotopic ratio and following the EXPT-D-HI-T protocol are shown in Fig. 8

Page 17: [31] Moved to page 13 (Move #15)	Microsoft Office User	1/24/18 10:53:00 AM
---	-----------------------	---------------------

for the period September – December 2016. For the results using only the laboratory calibration, analyzer drift is apparent for all three tanks. Without a field calibration, the isotopic ratio was biased by up to 2 ‰.

Page 17: [32] Deleted	Microsoft Office User	1/24/18 10:53:00 AM
-----------------------	-----------------------	---------------------

The target tank was sampled only once per day and the resulting measurement is used in the calibration; hence the apparent drift following final calibration is necessarily zero. On the other hand, the high tank was sampled 13.5 times

per day and the average is used in the calibration. The low tank was independent of the isotopic ratio calibration. Prior to 3 December 2016, the low tank was also sampled 13.5 times per day, for 6 min each time (excluding transition time between gases). Thus, the noise apparent in Fig. 8B prior to that date is at least partially due to insufficient sampling times.

On 3 December 2016, an improved sampling strategy was implemented, in which the target tank sampling time was increased from 6 min/day to 54 min/day (excluding transition times), achieved by sampling for 20 min every 420-min cycle (3.4 times/day, on average). The calibration times were achieved with multiple cycles in order to avoid not sampling the atmosphere for long periods. The calibration data for each day are averaged and applied to the ambient data. The low tank was sampled using an identical strategy (20 min every 420-min cycle), with the total amount of sampling time per day changing from 81 min to 54 min. The high tank was sampled on average 1.7 times per day (every 840 min) for 10 min. Excluding the transition times, the high tank sampling time was thus reduced from 26 min/day to about 10 min/day. This strategy reduced the noise apparent in Fig. 8B, but is not expected to affect the long-term (e.g., scales of several days or longer) bias

Page 17: [33] Moved to page 13 (Move #16)	Microsoft Office User	1/24/18 10:53:00 AM
---	-----------------------	---------------------

The relative effects of the calibration terms are illustrated in Fig.

Page 17: [34] Moved to page 13 (Move #18)	Microsoft Office User	1/24/18 10:53:00 AM
---	-----------------------	---------------------

) in Eq. (3) are time-dependent drift terms. These terms vary because of spectral variations in the optical loss of the empty cavity (c_0), and because of errors in the temperature or pressure of the gas, or changes in the wavelength calibration (χ). Recall that the parameters c_0 and χ were calculated following Eq. (15) in Rella et al. (2015). The calculation of the parameter c_0 used measurements from the high and target tank. The calculation of the parameter χ used measurements of the high tank and was not independent from p_0 . The largest calibration effect was from the c_0 term, which increased the calibrated isotopic ratios by -0.5 to 4 ‰ during September to December 2016. The χ term increased the final calibrated isotopic ratios by a smaller amount, -0.6 to 0.2 ‰.

Page 17: [35] Deleted	Microsoft Office User	1/24/18 10:53:00 AM
-----------------------	-----------------------	---------------------

There is variability in the calibration effect of these terms, although no software or hardware changes were applied during this period.

6.3 Noise in independent low tank as a function of averaging interval

In the optimized calibration scheme, the low tank was independent, and was treated as an ambient sample.

Page 17: [36] Moved to page 13 (Move #19)	Microsoft Office User	1/24/18 10:53:00 AM
---	-----------------------	---------------------

The low tank differences from known values, averaged over differing intervals, are shown in Fig.

Page 17: [37] Deleted	Microsoft Office User	1/24/18 10:53:00 AM
-----------------------	-----------------------	---------------------

10. As described in Section 5.1, the standard deviations of the low tank differences are a proxy for the noise in the calibrated ambient samples over those averaging intervals. The standard deviation of 13.5 calibration cycles per day, each of 6 min length, over the period September 1 – December 2 is 0.62 ‰.

Page 17: [38] Moved to page 14 (Move #20)	Microsoft Office User	1/24/18 10:53:00 AM
---	-----------------------	---------------------

During this period, the calibration used 6 min/day measurements of the target tank. The standard deviation of the low tank calibration cycles was similar to expectations based on the Allan standard deviation (Fig.

Page 17: [39] Deleted	Microsoft Office User	1/24/18 10:53:00 AM
-----------------------	-----------------------	---------------------

7). Averaging over the cycles in one day (a total of 81 min of data) yielded a standard deviation of 0.40 ‰.

Page 17: [40] Moved to page 14 (Move #21)	Microsoft Office User	1/24/18 10:53:00 AM
---	-----------------------	---------------------

Based on this result, differences in the hourly average between towers of less than 0.40 ‰ were likely not significant.

Page 17: [41] Deleted	Microsoft Office User	1/24/18 10:53:00 AM
-----------------------	-----------------------	---------------------

For 3-day means (a total of 4.1 hr), the standard deviation over the three-month period was 0.26 ‰. For the period after the calibration tank sampling scheme was improved (primarily by sampling the target tank for 54 min/day instead of 6 min/day), December 3 – December 31, the standard deviation of the individual cycles reduced substantially, to 0.25 ‰, and that of the one-day (three-day) mean of the cycles was 0.18 ‰ (0.11 ‰).

Page 17: [42] Moved to page 14 (Move #22)	Microsoft Office User	1/24/18 10:53:00 AM
---	-----------------------	---------------------

Therefore, according to this metric, after the improved calibration scheme was implemented, differences in the hourly average between towers of greater than 0.18 ‰ were significant.

Page 17: [43] Moved to page 14 (Move #25)	Microsoft Office User	1/24/18 10:53:00 AM
---	-----------------------	---------------------

The results for the round-robin style laboratory testing

Page 17: [44] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
-------------------------	-----------------------	---------------------

Font color: Auto, English (US)

Page 17: [45] Deleted	Microsoft Office User	1/24/18 10:53:00 AM
-----------------------	-----------------------	---------------------

of two NOAA/INSTAAR tanks are shown in Fig. 11. The mean of the errors (measured – NOAA known value) calculated from the results of four to six tests for each analyzer were –0.08 to 0.04 ppm CO₂, within the 0.1 ppm WMO compatibility recommendation for global studies of CO₂ (GAW Report No.

Page 17: [46] Moved to page 14 (Move #26)	Microsoft Office User	1/24/18 10:53:00 AM
---	-----------------------	---------------------

229, 2016). The standard error, indicating an estimate of how far the sample mean is likely to be from the true mean, for the means of the CO₂ tests were 0.03 – 0.10 ppm. The mean difference was –0.03 to 0.02 ppm CO₂ for the analyzers, averaged over the two round-robin tanks (analogous to averaging over the entire range of CO₂ during the flask comparison, for example). For CH₄, the means of the errors were 0.03 – 0.07 ppb CH₄, for the NOAA/INSTAAR tank measuring 2385.2 ppb, and –0.83 to –0.70 ppb CH₄ for the NOAA/INSTAAR tank measuring 1895.3 ppb CH₄. Therefore, there was a slight error in the slope of the linear calibration, possibly attributable to tank assignment errors.

However, the error was well within the WMO recommendations for global studies of 2 ppb CH₄ (GAW Report No. 229, 2016), and the range of NOAA/INSTAAR tanks encompassed the majority of the CH₄ mole fraction observed during the study. We also note that the standard error for the means of the CH₄ tests were 0.07 – 0.12 ppb. Averaging over the two round-robin tanks, the mean difference was –0.40 to –0.32 ppm CH₄ for the analyzers. For $\delta^{13}\text{CH}_4$, the mean errors for each analyzer/tank pair were –0.33 to 0.24 ‰ for these tanks within the range of ambient isotopic ratio and the standard errors were 0.05 – 0.10 ‰. The mean errors were –0.14 to 0.03 ‰ for each analyzer.

Page 17: [47] Deleted

Microsoft Office User

1/24/18 10:53:00 AM

6.5 Side-by-side testing

Side-by-side testing of a G2301 (CO₂/CH₄/H₂O) analyzer and a G2132-i analyzer (CH₄/ $\delta^{13}\text{CH}_4$ /CO₂) for June 2016 at the South tower

Page 17: [48] Moved to page 15 (Move #28)

Microsoft Office User

1/24/18 10:53:00 AM

resulted in mean differences of 0.06±0.41 ppm CO₂ and 0.9±1.5 ppb CH₄, with the G2132-i analyzer measuring slightly lower for both species. Here the standard deviation was based on the 10-min average calibrated values for the month for all times of the day. The standard error of the differences was 0.01 ppm CO₂ and 0.02 ppb CH₄. These results indicate that the performance of the G2132-i is similar for CO₂ and CH₄ mole fractions, at least in terms of the long-term mean. In terms of utilizing the mole fraction data in atmospheric inversions, the multi-day mean afternoon differences are

Page 17: [49] Deleted

Microsoft Office User

1/24/18 10:53:00 AM

more appropriate. The five-day mean afternoon difference for the month was 0.05±0.08 ppm CO₂ and –0.7±0.1 ppm CH₄.

Page 17: [50] Moved to page 15 (Move #29)

Microsoft Office User

1/24/18 10:53:00 AM

The G2132-i analyzers are thus appropriate for use in the atmospheric inversions and in the global network where 0.1 ppm CO₂ and 2.0 ppb CH₄ have been identified as criteria. For these results, recall that the target tank was

Page 17: [51] Moved to page 15 (Move #30)

Microsoft Office User

1/24/18 10:53:00 AM

To optimize results on a daily time scale, sampling the target tank for 60 min per day would be preferable for improving CO₂ results. We also note that round robin testing of these instruments requires 60 min sampling per tank.

Page 17: [52] Moved to page 16 (Move #33)

Microsoft Office User

1/24/18 10:53:00 AM

For January – December 2016, the mean flask to in-situ CH₄ difference at the East tower was –1.2 ± 2.2 ppb CH₄, and at the South tower was –0.9 ± 1.4 ppb CH₄ (Fig.

Page 17: [53] Deleted

Microsoft Office User

1/24/18 10:53:00 AM

12A). Recall from Section 5.4 that flasks were sampled in the late afternoon, integrated over one hour, and only when the winds were steadily from the west or north. Here the standard deviation reported is that of the hourly flask to in-situ differences for the year.

Page 17: [54] Moved to page 16 (Move #34) **Microsoft Office User** **1/24/18 10:53:00 AM**

Thus, at the South tower, for example, on 67% of the sampled afternoons indicated differences for CH₄ within 1.4 ppb of the mean of -0.9 ppb.

Page 17: [55] Deleted **Microsoft Office User** **1/24/18 10:53:00 AM**

For CH₄, data points with high temporal variability (standard deviation of raw ~2 sec data within the hour > 20 ppb) were excluded, on the basis that the ambient variability was large, making comparisons difficult.

Page 17: [56] Moved to page 16 (Move #35) **Microsoft Office User** **1/24/18 10:53:00 AM**

The standard error was 0.24 ppb at the East tower and 0.14 ppb at the South tower. Thus, there is high confidence that the difference between the in-situ and flask measurements at both towers is more compatible than the WMO recommendation. As for the side-by-side testing, the G2132-i analyzers were slightly lower than the “known”, in this case, the flask results. The difference, was however, less than the target compatibility, and the flasks could in theory be biased.

Although CO₂ is not the focus of this paper, the differences were -0.21 ± 0.31 ppm for the East tower and 0.21 ± 0.35 ppm for the South tower (Fig.

Page 17: [57] Moved to page 16 (Move #36) **Microsoft Office User** **1/24/18 10:53:00 AM**

). The standard error was 0.03 ppm at the East tower and 0.04 ppm at the South tower. The magnitude of CO₂ differences was somewhat larger in the growing season.

Page 17: [58] Moved to page 16 (Move #37) **Microsoft Office User** **1/24/18 10:53:00 AM**

thus larger than the WMO recommendation of 0.1 ppm, but at the extended compatibility goal of 0.2 ppm CO₂ (GAW Report No. 229, 2016).

For the isotopic ratio of methane, the mean flask to in-situ differences were 0.08 ± 0.54 ‰ and 0.02 ± 0.38 ‰ at the East and South towers, respectively (Fig.

Page 17: [59] Deleted **Microsoft Office User** **1/24/18 10:53:00 AM**

12C). The standard error was 0.06‰ and 0.04‰ at the East and South towers, respectively. Thus, there is high confidence that these differences are less than the target compatibility of 0.2 ‰. The standard deviation reported is that of the hourly flask to in-situ differences for the year.

Page 19: [60] Deleted **Microsoft Office User** **1/24/18 10:53:00 AM**

We found that field calibrations (including both a linear calibration and a correction for mole fraction dependence of $\delta^{13}\text{CH}_4$) significantly improved the compatibility of the measurements (as seen by comparing EXPTs H and F with EXPT D in Table 3). There was, however, no significant drift within a single day (as seen by comparing variations

of EXPT D in Table 4). Using these findings, we developed an optimized calibration strategy, given the tank sampling strategy

Page 19: [61] Deleted **Microsoft Office User** **1/24/18 10:53:00 AM**

3 December 2016 (Table 5). Instead of using the target tank as an independent assessment of compatibility, we used the target tank in the calibration scheme and evaluated performance with the low tank at each site. An improved tank sampling strategy was implemented on 3 December 2016.

Prior to the improvement in the tank sampling strategy, averaging over a period of 4.1 hours within a 3-day time period (a proxy for the noise within 4-hour afternoon averages of ambient data) yielded a standard deviation of the independent low tank of 0.26 ‰. After the improvements in the tank sampling strategy were implemented, averaging over the same time period was sufficient to achieve standard deviation of 0.11 ‰.

Furthermore, an ethane rough calibration was performed using the two field tanks scrubbed of ethane and one field tank with ambient levels of ethane. These roughly calibrated ethane values were subsequently used to correct for cross-interference with the

Page 19: [62] Deleted **Microsoft Office User** **1/24/18 10:53:00 AM**

of methane.

The round robin results using NOAA/INSTAAR tanks treated as unknowns in Fig. 11 showed mean error (averaged over the two round robin tanks) of -0.03 to 0.02 ppm CO₂ and -0.40 to -0.32 ppb CH₄ for the

Page 19: [63] Deleted **Microsoft Office User** **1/24/18 10:53:00 AM**

. These results were within the WMO recommendations (GAW Report No. 229, 2016). For $\delta^{13}\text{CH}_4$, the mean errors of the tests for each analyzer were -0.14 to 0.03 ‰, with larger errors for individual tanks. Earlier

Page 19: [64] Deleted **Microsoft Office User** **1/24/18 10:53:00 AM**

with reduced sampling times of about 10 min per tank (not shown) indicated increased errors for CO₂ and $\delta^{13}\text{CH}_4$, consistent with the Allan standard deviation results (Fig. 7).

The

Page 19: [65] Deleted **Microsoft Office User** **1/24/18 10:53:00 AM**

to in-situ comparison for $\delta^{13}\text{CH}_4$ (Fig. 12) showed that the field calibration tank sampling strategy and calibration protocol (Table 4) used in this study was sufficient for producing isotopic methane results with low long-term bias (0.02 to 0.08 ‰ averaged over one year). The CO₂ differences were -0.21 to 0.21 ppm and the CH₄ differences were -1.22 to -0.87 ppb.

Recall from Table 1 that the calibration of the G-2132i

Note that the high mole fraction tanks are ideally near ambient isotopic ratio, but the effect of being further from ambient isotopic ratios is likely small. The isotopic ratios of the high tanks are thus listed as values provided by Isometric Instruments bottles, making the tanks logistically less challenging and less expensive to acquire. On 3 December 2016, we implemented an improved sampling strategy, primarily by increasing the sampling time for the target tank (Table 5). An alternate possible calibration tank sampling strategy is to sample an additional high tank at a different isotopic ratio (Table 5, third column, and Fig. 16B). With this strategy, a laboratory calibration prior to deployment is not necessary. Also, both the slope and intercept of the linear calibration can be adjusted in field, rather than just the intercept, which may improve results. In this case, however, the calibration is over-constrained, using two tanks to solve for two variables. Additionally, it would also be preferable for both low and target tanks to be near -47‰ and sampled for about one hour per day, either all at once or spread out over the day. The ranges of isotopic ratios and CH_4 mole fractions desirable are indicated in Fig. 16B.

$\delta^{13}\text{CH}_4$ for tower measurements as described in this paper are demanding.

If compatibility of 0.5‰ is sufficient for a specific application, the sampling strategy listed in fourth column of Table 5 and Fig. 16C may be applicable.

Table 3. Results for the South tower for October 2016 using multiple possible isotopic ratio calibration schemes. In all of these experiments, the target tank isotope value was independent of the analyzer calibration, as planned prior to the deployment. This differs from the optimized calibration scheme and tank uses listed in Table 1, as discussed in the text. In EXPTs A – G, the daily average of the 13.5 calibration cycles per day for the high tank and low tank was used (if applicable). In EXPT D1, only the first high and low tank calibration cycle of the day was used. In EXPT D-T, only the high and low tank calibration cycle immediately preceding the target tank sampling was used. The high tank at each tower contains 9.7–10.5 ppm CH_4 and about -38.3‰ $\delta^{13}\text{CH}_4$, the low tank contained 1.9–2.1 ppm CH_4 and about -23.9‰ $\delta^{13}\text{CH}_4$, and the target contained ~ 1.8 ppm CH_4 and about -47.2‰ . The experiments yielding the results with lowest bias and standard deviation are highlighted in blue (EXPTs D and variants). EXPTs that answer the four specific questions raised in Section 4.4 are indicated. Tanks used in the calibration are indicated as such and are not independent for that calibration scheme.

	Lab calibration (linear calibration, mole fraction correction) p_1 p_0 c_0 χ	Daily average of high tank for intercept calibration p_0	Daily average of high and low tanks for linear calibration p_1 p_0	Daily average high and low tanks mole fraction correction c_0 χ	High tank error (‰) mean \pm standard deviation for one month (standard error)	Low tank error (‰) mean \pm standard deviation for one month (standard error)	Target tank error (independent) (‰) mean \pm standard deviation for one month (standard error)
EXPT A	-	-	-	-	0.5 ± 0.4 (0.0)	1.5 ± 1.3 (0.1)	-0.6 ± 1.3 (0.2)
EXPT B	-	-	-	✓	Used in cal	Used in cal	-2.2 ± 0.4 (0.1)

EXPT C	-	-	✓	-	Used in cal	Used in cal	-0.5±1.4 (0.3)
EXPT D	✓	-	-	✓	Used in cal	Used in cal	-0.3±0.4 (0.1)
EXPT E (question 1)	✓	-	-	-	0.1±0.4	0.0±1.2	-0.3±1.2 (0.2)
EXPT F (question 2)	✓	✓	-	-	Used in cal	-0.1±0.9 (0.0)	-0.4±0.9 (0.2)
EXPT G (question 3)		-	✓	✓	Used in cal	Used in cal	-0.5±1.4 (0.3)
EXPT D1 (question 4)	✓	-	-	✓	Used in cal	Used in cal	-0.3±0.5 (0.1)
EXPT D-T (question 4)	✓	-	-	✓	Used in cal	Used in cal	-0.3±0.6 (0.1)

Table 4. Results for the four Marcellus towers using multiple possible calibration schemes. Results for October 2016 are shown for the South, East and North towers. Note that the sample size (used for the calculation of the standard error) was larger for the high and low tanks compared to the target tank, as those tanks were sampled 13.5 times per day. Results for the Central tower are shown for May 2016 (analyzer at manufacturer for repairs during October 2016). EXPT H (only lab calibration) and EXPT D (lab calibration, daily mean of high tank for intercept calibration and daily averages of high and low tanks for mole fraction correction) is as in Table 3. The target tank was independent in these cases. EXPT D-HI-PRE-HI-T is similar to EXPT D, but uses only the high tank value immediately preceding the target tank sampling for the day. The low tank was independent in this case (and only the low tank sampling immediately preceding the target was considered in the results for this case). EXPT D-HI-T used the high tank for the intercept calibration (as in EXPT D), but used the daily average High and the Target tanks for the mole fraction correction. The low tank was again independent in this case. EXPT D-HI-T (highlighted in blue) is the calibration scheme applied to the entire dataset.

Page 31: [70] Deleted		Microsoft Office User			1/24/18 10:53:00 AM
EXPT H	South	0.1±0.4 (0.0)	0.0±1.2 (0.1)	-0.3±1.2 (0.2)	
EXPT H	East	0.3±0.4 (0.0)	-0.1±0.8 (0.0)	-0.9±0.9 (0.2)	
EXPT H	Central	0.0±0.2 (0.0)	0.3±0.7 (0.0)	-0.1±0.8 (0.1)	
EXPT H	North	-0.6±0.4 (0.0)	-1.5±1.3 (0.1)	-1.9±0.6 (0.1)	
EXPT DSCHEME A	South	Used in cal	Used in cal	-0.3±0.4 (0.1)	

Page 31: [71] Formatted Table	Microsoft Office User	1/24/18 10:53:00 AM
Formatted Table		

Page 31: [72] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
Font color: Text 1, Pattern: Clear (White)		

Page 31: [73] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
Font color: Text 1		

Page 31: [74] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
Font color: Text 1, Pattern: Clear (White)		
Page 31: [75] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
Font color: Text 1		
Page 31: [76] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
Font color: Text 1, Pattern: Clear (White)		
Page 31: [77] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
Font color: Text 1		
Page 31: [78] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
Font color: Text 1, Pattern: Clear (White)		
Page 31: [79] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
Pattern: Clear (White)		
Page 31: [80] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
Font color: Text 1		
Page 31: [81] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
Font color: Text 1		
Page 31: [82] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
Font color: Text 1		
Page 31: [83] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
Font color: Text 1		
Page 31: [84] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
Font color: Text 1		
Page 31: [85] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
Font color: Text 1		
Page 31: [86] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
Font color: Text 1		
Page 31: [87] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
Font color: Text 1		
Page 31: [88] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
Font color: Text 1		
Page 31: [89] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
Font color: Text 1		
Page 31: [90] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
Font color: Text 1		
Page 31: [91] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
Font color: Text 1		
Page 31: [92] Formatted	Microsoft Office User	1/24/18 10:53:00 AM

Font color: Text 1

Page 31: [93] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
-------------------------	-----------------------	---------------------

Font color: Text 1

Page 31: [94] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
-------------------------	-----------------------	---------------------

Font color: Text 1

Page 31: [95] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
-------------------------	-----------------------	---------------------

Font color: Text 1

Page 31: [96] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
-------------------------	-----------------------	---------------------

Font color: Text 1

Page 31: [97] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
-------------------------	-----------------------	---------------------

Font color: Text 1

Page 31: [98] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
-------------------------	-----------------------	---------------------

Font color: Text 1

Page 31: [99] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
-------------------------	-----------------------	---------------------

Font color: Text 1

Page 31: [100] Formatted	Microsoft Office User	1/24/18 10:53:00 AM
--------------------------	-----------------------	---------------------

Font color: Text 1

Page 32: [101] Deleted	Microsoft Office User	1/24/18 10:53:00 AM
------------------------	-----------------------	---------------------

HIGH (10 ppm,

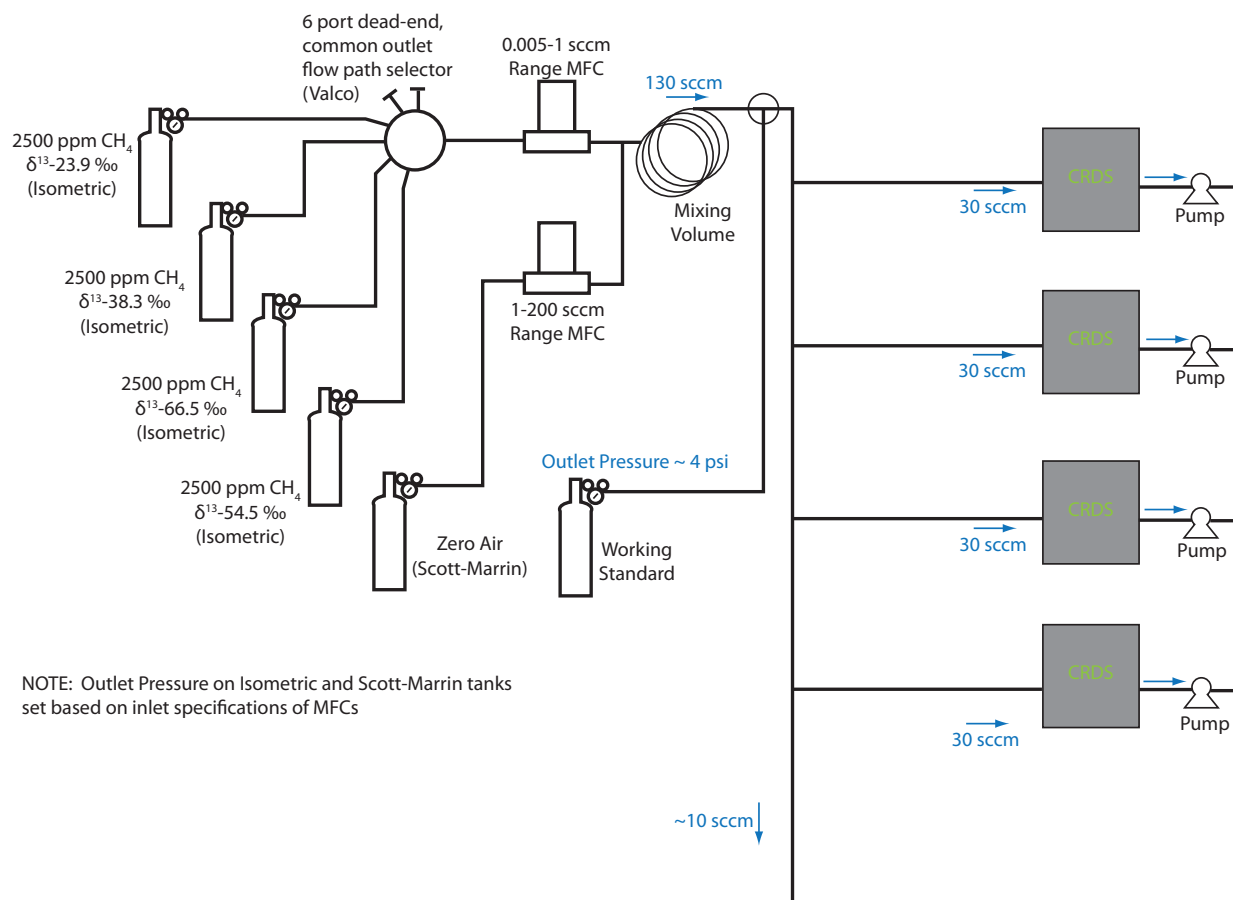
–38.3‰, 4 min/day)

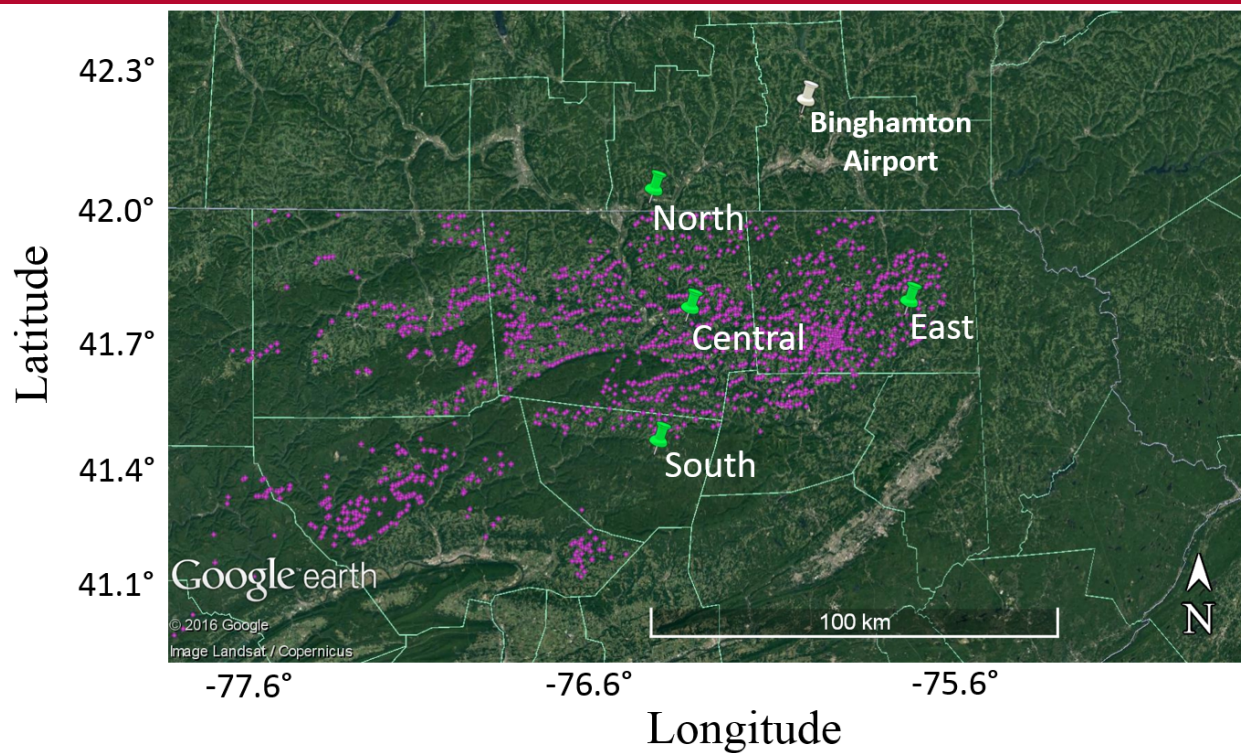
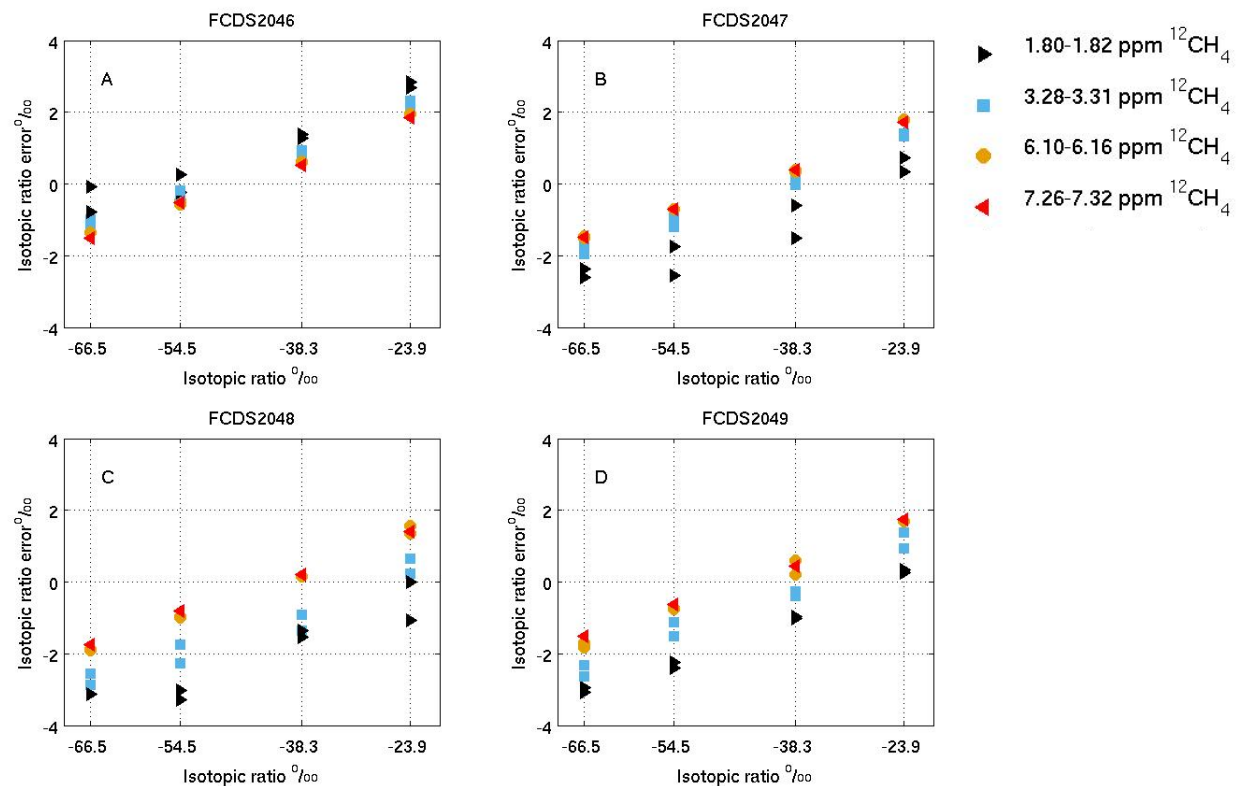
Page 32: [102] Deleted	Microsoft Office User	1/24/18 10:53:00 AM
------------------------	-----------------------	---------------------

LOW (2 ppm,

–23.9‰, 6–32 min/day)

Page 35: [103] Deleted	Microsoft Office User	1/24/18 10:53:00 AM
------------------------	-----------------------	---------------------





Map of Pennsylvania with permitted unconventional natural gas wells (magenta dots) and network of towers with methane and stable isotope analyzers (Picarro G2132-i). The East and South towers were also equipped with NOAA flask sampling systems. The Binghamton Airport is also indicated.

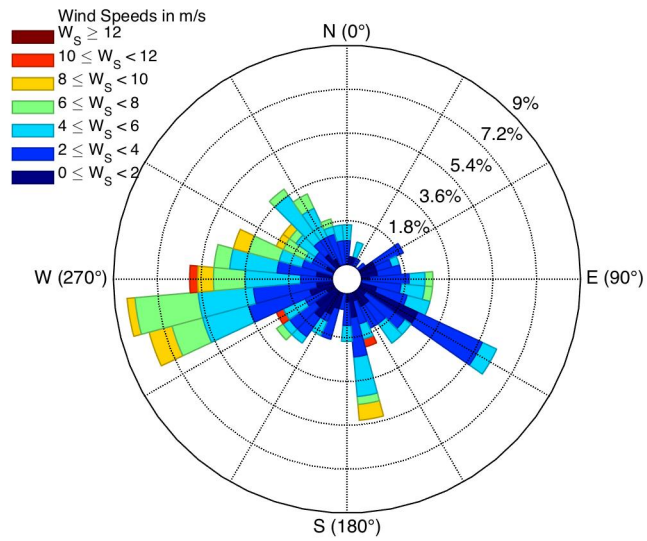


Figure 5.

m s⁻¹) are not categorized by direction.

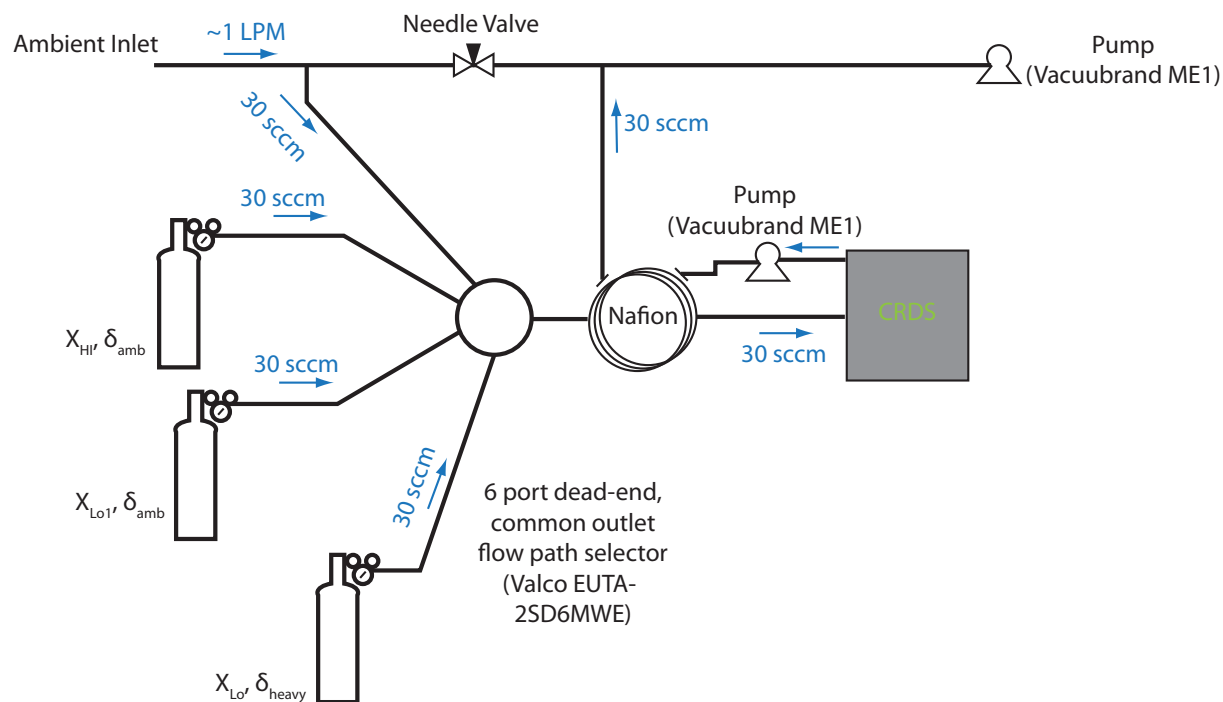


Figure 6

Allan standard deviation for (A) $\delta^{13}\text{CH}_4$, (B) CH_4 , and (C) CO_2 for a high CH_4 mole fraction tank (9.7 ppm CH_4 , ~ 400 ppm CO_2 , -38.3 ‰ $\delta^{13}\text{CH}_4$) (orange) and a low (1.9 ppm CH_4 , ~ 400 ppm CO_2 , -23.7 ‰ $\delta^{13}\text{CH}_4$) tank (blue). The x-axis is truncated to focus on minimum averaging times required to achieve the desired compatibility goals.

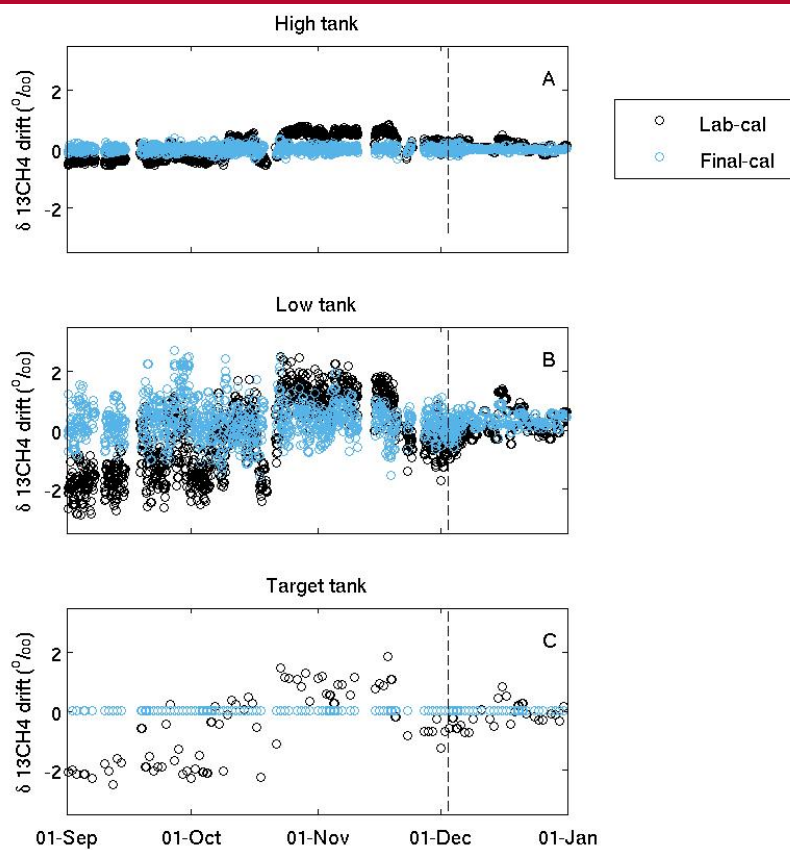
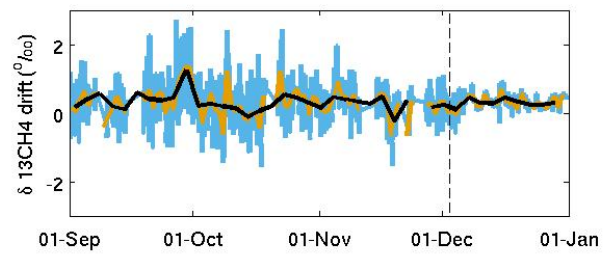
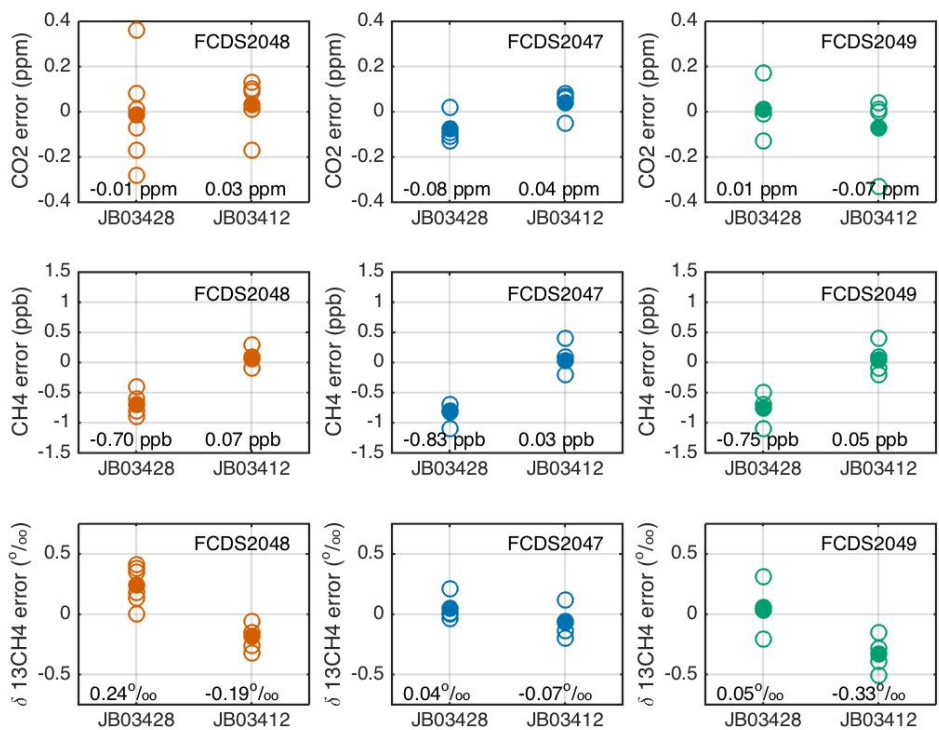
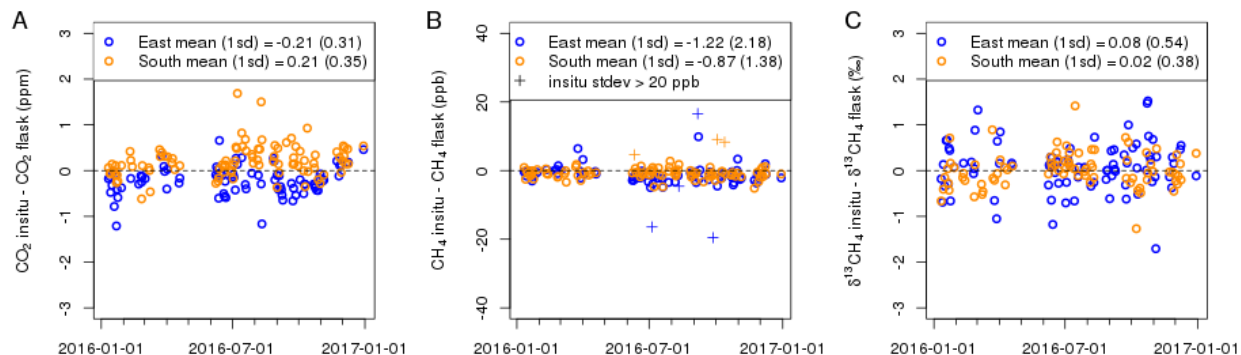


Figure 8.







The standard deviation of the in-situ to flask differences are shown in parentheses on each plot. The standard errors, indicating an estimate of how far the sample mean is likely to be from the true mean, is 0.24 ppb CH₄, 0.03 ppm CO₂ and 0.06 ‰ at the East tower and 0.14 ppb CH₄, 0.04 ppm CO₂ and 0.04 ‰ at the South tower.

