Referee comments in black. *My responses in green italics.*

Anonymous Referee #3

Received and published: 15 January 2021

This paper describes the first dataset of atmospheric propane (C3H8) retrieved from remote sensing measurements. A large amount of data obtained at 12 different ground- stations over several years is presented. The authors give convincing explanations for the enhanced propane amounts observed at two sites, namely losses during exploitation of natural gas and pollution by liquified petroleum gas in large cities. They substantiate their conclusions by correlation with C2H6 and with CO, tracers for gas losses and anthropogenic pollution. The manuscript is well structured and its objective is clear. Therefore I recommend publication in ACP. *Thank you!*

Major comments:

The strong correlation with C2H6 for, e.g., the Ft. Sumner data is convincing evidence that the C3H8 VMRs are real, maybe with an unknown bias. Nevertheless, I would be interested in a plot similar to Figure 2 but for retrievals using all other fit-parameters except of C3H8 to see, if the residuals become significantly larger. Maybe this could be shown in the Appendix.

Good idea. Fig.2 has been expanded to include an extra panel showing the residuals that would arise if there were no atmospheric C3H8. Residuals increase from 0.36% to 0.39%

A table containing the name, acronym, geographical coordinate, altitude and country of each of the 12 sites at the beginning of Section 3 would be helpful. *Done.*

Figures 3-5: The authors state that Fig. 3 contains C3H8 column amounts from 12 sites. If so, the sites should be made better distinguishable. I can detect 9 different lines or 7 different colours in the upper panel at the most. The same applies for Figures 4 and 5. Especially the two different "greens" are not easily to distinguish. Maybe it would help if one "green" was a bit darker? I concede that it might hardly be possible to present the data using 12 clearly distinguishable colours. Maybe subdivision into low-and high-altitude stations would help, by which the lines in the upper panel of Fig. 3 would be easier to distinguish as well. Further, every station should be listed with its associated colour code in the figure captions. Which station is, e.g., blue (0 km)?

Yes, 12 colors are too many to be distinguishable. The easiest solution is to drop the four sites with fewest observations, namely JPL Mesa, CA, Palestine, TX, and Mountain View, CA, and Lynn Lake, Manitoba. This significantly reduces the color ambiguity, without noticeable changes to the figure. Terse sites contributes only 1% of the total observations.

L260ff: I would appreciate a little more information how the authors derive 0.72E+19 molecules cm-2 in total and 3E+16 molecules cm-2 of propane from 15 billion cu. ft. produced per day.

I expanded the text as follows. "The Permian basin currently produces 16 billion cu.ft./day of NG (https://www.eia.gov/petroleum/drilling/pdf/permian.pdf) over an area of 220,000 km^2. The molar volume of an ideal gas at STP is 22.4 liters. One cu. ft. is 28.3 liters. So 16 billion cu. ft. is 20 billion moles of NG or 120x10^{32} molecules per day. Over an area of 220,000 km^2 or 2.2x10^{15} cm^2, this represents an average areal production of 55x10^{17} molec./cm^2/day. Assuming that the Permian basin of 480 km long, at an average low-level wind speed of 15 km/hour, an air parcel will take 32 hours (1.33 days) to traverse the Basin, during which time 73x10^{17} molecules/cm^2 will have been extracted. Of this, 10% will be C_3H_8 (Howard et al., 2015), so if all this production were released into the atmosphere we would expect a C_3H_8 column enhancement of 73x10^{16}.

In airmasses with trajectories from the SE, we see maximum C_3H_8 column enhancements of only 3x10^{16} molecules/cm^2, which suggests that only 4% of the NG escapes into the atmosphere and that 96% of the NG is successfully captured (or burnt by flaring). Of course, this analysis assumes that the Permian basin is a uniform emitter and that the back trajectory wind speeds are accurate. There are likely hot spots with higher-than-average emissions, and regions with little NG production."

Minor comments:

L10: Aren't rather the high C3h8 amounts than the variations "correlated" with back trajectories from SE New Mexico ...? *Agreed and fixed.*

L55: "the entire 650–5650 cm-1 range" instead of "the entire 650–5650 cm-1" Fixed.

L63/64: "(Irion et al., 2002)" instead of "(Irion et al., 2003)" *Fixed.*

L71: I do only count 9 simultaneously-fitted scalars. Can you help me?

We retrieve 5 gases (C_3H_8, H_2O, CH_4, C_2H_6, HDO), two frequency stretches (solar & telluric), two continuum parameters (a straight line is defined by two coefficients), and a zero-level offset.

L75: The acronym TCCON should be introduced in line 65. *Done.*

Further, what is GGG2014? L78:

Added sentence "The entire package including spectral fitting software, spectroscopic linelists, and software to generate a priori VMR/T/P profiles, is denoted GGG."

What is GGG2020?

Added "an updated version of GGG with improved a priori VMR/T/P profiles, spectroscopy, and software (Laughner et al., 2021)"

L79: I think "less than 10%" is more appropriate than "less than 10% rms", because specifications in percent are dimensionless. *Agreed. Removed "rms".*

Further, "shown in Fig. A2" more specifically indicates, where the differences can be seen, than "shown in appendix A". *Done.*

L82: Which infra-red lab? *Added "Pacific North-West National Laboratory (PNNL)" and also a reference to Sharpe et al. (2004)*

Figure 1: The labelling of the axes and the legend is rather small and blurred and should be represented clearer. *Done.*

L124: Here it says: "Table Mountain Facility at 2.2 km" but in the captions of Fig. 3 "orange=2.25 km (TMF)". Please adjust the heights. *All are now 2.26 km.*

L130: I believe the sentence "So C3H8 has only ..." would become clearer, if it would be exchanged with the preceding sentence. *Agreed. Done.*

Figure 2: The colors for H2O and HDO are hardly to distinguish. *Agreed. I have re-made the figure with more color separation between H2O and HDO.*

L155: The upper row and not the left hand panels show XC3H8. Please change into "The upper row of Fig.4 shows the XC3H8 time series plotted versus year (left) and versus day of the year (right)."

Changed to "The upper and lower rows of Fig.4 show the XC3H8 and XC2H6 time series, respectively, plotted versus year (left) and versus day of the year (right)."

Figure 4, captions: It should be mentioned that the colour coding is the same as in Fig. 3. This information could then be removed in lines 170/171. *Done.*

L166: "The Antarctic measurement (blue) are even lower than they appear because ..." sounds strange. What about "The Antarctic measurements (blue) (of C2H6?) are very low (0.2-0.3 ppb) and most probably even lower during the rest of the year, because ..."? *Agreed and fixed.*

L172: The sentence "In fact, the highest VMRs of C2H6 were seen from there, even more than from JPL ..." seems to be contradictionary to L167/168: "The highest ever C2H6 was measured from JPL (cyan) in late 2015 ..." *Agreed. The latter sentence now says "The highest C2H6 ever measured from JPL (cyan) was in late 2015 ..."*

Figure 5: "... by site altitude like in Fig. 3." (?) *Yes. Fixed.*

L197: The authors state: "... but only when the wind direction is from the SE quadrant (green/lime colors)." On the other hand it says green = 180 deg, lime = 220 deg in Fig. 6. Shouldn't the SE quadrant be in the direction 135+-45 deg? Or is the wind direction not counted clockwise, beginning in the north?

This is a simple typo: lime was written instead of cyan. The text now states ""... but only when the wind direction is from the SE quadrant (green/cyan) colors". Lime is now listed with the other wind directions in the next sentence.

Figure 7: Do the authors apply a standard regression analysis (minimization of the squared vertical distances)? If so, wouldn't it be better to correlate XC3H8 versus XC2H6, because the X2H6-errors are much smaller?

Given the measured data $[x_i \pm ex_i, y_i \pm ey_i]$ *the regression minimizes*

$$\chi^2 = \Sigma_i (x_i - x')^2 / ex_i^2 + \Sigma_i (y_i - y')^2 / ey_i^2$$

with respect to A and B, where $y' = A + Bx'$ *is the fitted straight line. This is equivalent to the method described by (York, 1969; Wehr and Saleska, 2017) but assuming no correlation between the x- and y-uncertainties. So the error bars of both C2H6 and C3H8 are taken into account. Upon switching X and Y, the same the exact same PCC is obtained and the gradient simply reciprocates.*

L228: Why was a trajectory altitude of 0.4 km over Ft. Sumner selected?

I thought that this was the altitude with the most transport of short-lived gas molecules released at the surface. At lower altitudes there is more C3H8, but the wind speed is less, especially at night. At higher altitudes above the PBL the winds are stronger but the gas concentration drops off rapidly.

L280: Why is 25% low only half the problem? Because of the rest of the profiles above the PBL or due to other reasons? Please explain.

Paragraph has been re-written as " A puzzle in our findings is that when both C_3H_8 and C_2H_6 are elevated, we measure 22% more C_3H_8 than C_2H_6 (see fig.9). Yet independent essays of well-head wet NG find 33% more C_2H_6 than C_3H_8 in the Permian basin (Howard et al., 2015). So we have a 55% discrepancy. We note that the C_2H_6 averaging kernel is 0.7 at the surface versus 0.9 for C_3H_8 (see Appendix B). So when these gases exceed their priors in the PBL, which is likely at high enhancements, both will be under-estimated, but C_2H_6 more so than C_3H_8. So this effect would cause the C_3H_8/C_2H_6 ratio to be 28% high, which explains half the 55% problem."

Figure B.2: The axes-units are missing. *Added units.*

Technical comments:

L9: "shows" instead of "show"? *Fixed.*

L16: "losses" instead of "loses"? *Fixed.*

L32: "is therefore is": one "is" should be removed *Fixed.*

L40: "show a large" instead of "show large a" *Fixed.*

L65/66: "(Toon et al., 2016; 2018a; 2018b)" instead of "Toon, 2016; 2018a; 2018b)" *Fixed.*

L124: "in red" and "in orange" instead of "= Red" and "= Orange" *Fixed.*

L166: "measurements" instead of "measurement" *Fixed.*

L363: "shows" instead of "show a" *Fixed.*

L364: "when back-trajectories from SE New Mexico and West Texas ...": I think a verb is missing here. *Fixed.*