Author's responses to reviewer's comments follow. A copy of the reviewer comment is given (with comment 'number') followed by a response (blue font).

Response to referee 2

1. General comment

The clear explanation of the set-up of the measurements at the 3 stations, scale propagation and uncertainty would be valuable to others developing and refining greenhouse gas measurement systems. The publication of the high-quality data from the stations, and discussion of the regional context is good. I think further work should be done on the source identification (section 3.5). This may change the overall conclusion of the work – i.e. whether the changes in the region are a relative increase in biogenic sources or not. Before publication some edits are required.

Thank you for your comments on the paper's value. We also appreciate your helpful comments to improve our manuscript. According to your specific comments, we revised our manuscript, especially isotope analysis results. Also, we tried to explain PSS analysis more detailed.

2. Isotopic signatures of source regions

The section on the isotopic signatures for identifying the predominant sources is interesting but needs some more work: The Keeling plot technique (δ 13C against 1/CH4) is only appropriate for a constant background. Some detail is required about how these Keeling plots were constructed. Is a constant background appropriate or are you plotting data over a period when there will be seasonal variability or interannual differences in the background? If you are plotting data over several years and the global background methane mole fraction and δ 13C are changing, then Keeling plots should not be used. If the constant background assumption cannot be made then Miller-Tans plots could be used instead to identify the source isotopic

composition, e.g. Miller and Tans, 2003 https://doi.org/10.1034/j.1600-0889.2003.00020.x; Al-Shalan et al., 2022 https://doi.org/10.1016/j.atmosenv.2021.118763; Varga et al., 2021. Figure 9 In (c) and (d) we see that source signatures for CS and KL increased in 2016-2020 compared with 2006-2010. This seems to contrast with line 13 on page 21 which talks of a decreasing trend in δ 13C What are the uncertainties in the trends in (e) and (f). Can you really say there is a trend?

Agree. According to your comments, we analysed the Miller-Tans plot and could get more clear results for source changes. So we use PSS analysis as a tool to show the representative area where affected atmospheric CH_4 at AMY. Through the HYSPLIT cluster analysis based on TAP flask sampling dates, we selected the samples from the source regions where PSS analysis indicated. Miller-Tans plots were analysed every five years to see its changes.

Finally, we revised whole description in section 3.5. like below and we also added new graph in the manuscript and supplementary:

Section 3.5 from P21 L26: To understand the source regions affected AMY CH_4 level, we analysed PSS with hourly CH_4xs from 2006 to 2020. CH_4xs did not vary much and was 49±74 ppb during 2006–2010 and 50±70 ppb during 2016–2020. According to the PSS analysis, affecting major source regions were CN, CS and KL sectors (Fig. 9 (a)). Sources affecting CS and KL are paddy and livestock fields and that for CN was reported to be fossil fuel emissions mainly (Zhang et al., 2011, Ito et al., 2022, Chen et al., 2022).

Through the HYSPLIT cluster analysis from 2006 to 2020, we categorized the TAP $\delta^{13}C_{(CH4)}$ data and select the samples only affected by each source regions, CN, CS, and KL, respectively

(section 2.6). Using TAP $\delta^{13}C_{(CH4)}$ long-term data from 2006 to 2020 affected by CN, CS and KL, Miller-Tans plots indicated that emissions from CN were mainly related to fossil fuel or biomass burning (–44.3±1.8‰), while CS (–56.1±1.5‰) and KL (– 54.6±1.2‰) were affected more by biogenic sources during 2006–2020 (Fig.9 (b)). Sherwood et al. (2017) reported unweighted global mean $\delta^{13}C$ of –44.8±10.7‰ from fossil fuel use, –26.2±4‰ from biomass burning, and –61.7±6.2‰ from microbial sources. Even though the uncertainty of isotopic source signature is quite large, CH_4 formed at high temperature such as combustion is enriched in the heavier isotope while CH_4 from wetland, rice paddies and livestock is depleted. Therefore, our isotope analysis was well matched to reported source regions.

On the other hand, isotope signatures were shifted slightly in China (CN and CS) while for Korea (KL) it was steady in the uncertainty range from 2006 to 2020. When we analyze the Miller-Tans plots in every 5 years (Fig. S7), for CN the slope was -38±3‰ in 2006/10 but it became depleted -45±2.4‰ in 2016/20 while those value was enriched from -59.8±1.5‰ to -51.9±2.5‰ in CS. KL showed the quite constant values from -55 to -54‰ in the same period. This suggested that CH_4 growth rate in East Asia was affected not only biogenic but also pyrogenic sources, unlike global. The recent global accelerated increase in atmospheric CH_4 was more related to biogenic sources such as agriculture and wetland (Jackson et al., 2020, Lan et al., 2021).

Since the CH_4 emissions from agriculture and livestock accounted for 30% and 36% in China and Korea respectively in 2020 (Crippa et al., 2022), CH_4 might be increased by temperature impacts on biogenic CH_4 source. However, the fast urbanization and energy consumption strategy also can affect these regions. Especially the coal emissions decreased from 2010 in China (Liu et al., 2021) but the coal to gas policy lead natural gas consumptions increase again in China (Wang et al., 2022).

Overall, AMY and global growth rates were renewed in 2006 and during 2006–2020; the increasing trend could be linked to mixed biogenic and fossil fuel sources in East Asia while

global to more biogenic sources.

Regarding this, we revised the abstract and section 4. Summary and conclusion as well.

P1 L2: From the long-term records at AMY, we confirmed that growth rate increased 3.3 ppb·yr^{-1} during 2006/2010 and by 8.3 ppb·yr^{-1} from 2016 to 2020, which is similar trend to global. It is reported that the recent global accelerated CH_4 growth rate was related to biogenic sources. However, isotopic signature using $\delta^{13}CH_4$ explained that CH_4 sources are becoming mixture of not only biogenic but also fossil fuel sources in East Asia from 2006 to 2020. We confirmed that long-term high-quality data can help understand changes in CH_4 emissions in East Asia.

P24 L10: From the long-term analysis of CH_4 data at AMY, average CH_4 growth rate was 3.3 ppb·yr^{-1} during 2006–2010, but increased to 8.3 ppb·yr^{-1} in 2016–2020 as similar to the global trend. Through the source distributions with our PSS analysis using CH_4xs data, CN, CS and KL sectors were main regions to affect atmospheric CH_4 observed at AMY. Isotope signature based on Miller-Tans plots at CN represents fossil fuel or burning activities while CS and KL biogenic sources during 2006-2020. However, we infer atmospheric CH_4 drivers changes in air masses arriving from China sector, CN and CS. For East Asia the increasing trend could be linked to mixed biogenic and fossil fuel sources while global to more biogenic sources (e.g. agriculture and wetland). Through this study, we confirmed that long-term high-quality data can help understand changes in CH_4 emissions in East Asia. Also, further studies are necessary based on observations to understand sources changes in East Asia since there is a discrepancy between reported inventory and observations (Wang et al., 2022).

3. References

The reference lists need editing. Some of the references were missing from the reference list: Watanabe et al., 2000; Remann et al., 2004; Remann et al., 2008; Li et al., 2017; Turnbull et al., 2015. The reference Shuang-Xi Fang et al. (2013) should be deleted as this is already listed as Fang et al., 2013. Kim et al., 2014 on page 2 should be Kim et al., 2015 to match the reference list.

Corrected

4. Other questions:

4.1 As AMD is affected by local sources it would be helpful to use an inventory to suggest quantitatively what the anthropogenic emissions sources are (e.g. EDGAR, or UNFCCC) in the introduction.

We added the information of anthropogenic source in Korea in the introductions with the reference of EDGAR.

P 2, L10: China has also the largest anthropogenic CH_4 emissions in the world mainly from solid fuel (34%), rice cultivations (20%) and enteric fermentation (10%), respectively (Janssens-Maenhout et al., 2019; Crippa et al., 2022).

P 2, L14: South Korea major CH_4 emissions are derived from wastewater treatment (40%), enteric fermentation (22%) and then rice cultivations (14%) respectively (Crippa et al., 2022).

4.2 Figure 1 – add a scale bar to this map.

Corrected

4.3 What was the reason for drying the air rather than using the water correction built into Picarro software (see Rella et al., 2013)?

We do not use the water correction function provided by CRDS since we believe that H_2O value also should be calibrated when we apply the correction algorism to our data and general water correction cannot be applied to each instrument in different environment. In this context, many networks and stations don't use dry mole fraction reported by Picarro and apply post-processing water correction (Hazan et al., 2016, Zellweger et al., 2016).

In our case, we don't use post-processing since we dry our samples through cryogenic method (detailed in section 2.2). Also, even when we compress air into cylinder for working standards, we dry them. However, it should be considered the bias resulted from the differences of H_2O values between samples and standards or working and laboratory standards. To avoid the confusion, we added the sentence.

P 4, L20: Despite the Picarro provides built-in dry correction algorism, this is not applied to our data since generic water correction cannot be applied to each instrument in different environment. Though we dry our samples with this system, the biases resulted from the different H_2O values between samples and standards or working and laboratory standards are considered here. This is described in section 2.3 and 3.1.

[Reference]

Hazan, L., J. Tarniewicz, M. Ramonet, O. Laurent, and A. Abbaris.: Automatic processing of atmospheric CO2 and CH4 mole frations at the ICOS Atmosphere Thematic Centre, Atmos. Meas.Tech., 9, 4719-4736, 2016

Zellweger, C., L. Emmenegger, M. Firdaus, J. Hatakka, M. Hemann, E. Kozlova, T. G. Spain, M. Steinbacher, M. V. van der Schoot, B. Buchmann.: Assessment of recent advances in measurement techniques for atmospheric carbon dioxide and methane observations, Atmos. Meas.Tech., 9, 4737-4757, 2016

4.4 Page 2, line 31. A large ratio, CH4/C2H6 – explain what that means. Is high methane but low ethane indicative of a biogenic source?

We explicitly explained what it means.

P2 L32: The ratio, CH_4/C_2H_6, was observed 53 ppb·ppb^{-1} during KORUS-AQ campaign from May to June 2016 which seems to be associated fossil fuel in Seoul and Busan while it was 150 to 250 ppb·ppb^{-1} related to biogenic emissions such as rice paddies in southern western part of South Korea (Li et al., 2022).

4.5 Page 3 – lines 29/30. Were there 2 garbage incinerators or one? This part needs clarifying.

Thank you for the correction. We revised the sentence.

P3 L31: In the southwestern area, there is a small brickyard 200 m from the station and a garbage incinerator within 100 m. The garbage incineration facility was moved to the north side of island in December 2016.

4.6 Page 7 – it's not clear how the filtering was applied to the data using HS, CD and MS. Are data outside of 1 s.d. of the mean filtered out?

Since this method was published in Seo et al.(2021), we did not handle this process precisely. However, we tried to explain our method with more clear ways here to avoid the confusions.

P7 L14 There are three steps to select the background levels (L3 hourly data) from valid L2 hourly data:

Step 1) HS(t) ≤ A

Step 2) $|HA(t) - HA(t-1)| \leq B$ or $|HA(t) - HA(t+1)| \leq B$

Step 3) $|HA(t) -$ 30 days moving median of of $HA| \leq C$.

Where HS represents CH_4 hourly standard deviation, HA is CH_4 hourly means and t represents time in hours. In step 3), t is the middle of the time window. A, B and C are criteria determined empirically for each step, as given in Table 2. C is the standard deviation of 30 days moving average multiplied by α and here $1.8\sigma_{30d}$ is applied to all three stations as C.

Even though the data were selected by step 1) and 2), high CH_4 levels remained because of long-lasting stagnant conditions (e.g. over 6 days). Therefore, we also apply step 3). This process retained 21–52% of the data at each station, which were defined as L3 hourly on observations (Fig. S2). To get L3 daily/monthly data, the method developed by Thoning et al. (1989) was used to fit smooth curves to the daily averages computed by L3 hourly data. The methods reduce noise induced by synoptic-scale atmospheric variability, fill measurement gaps, and are used to represent the regional baseline. The details were described in Appendix B the supplementary. Finally, we can get the L3 daily data, L3 monthly data, long-term trend and seasonal amplitude after applying Thoning et al. (1989). The detailed definitions are in Appendix C, supplementary.

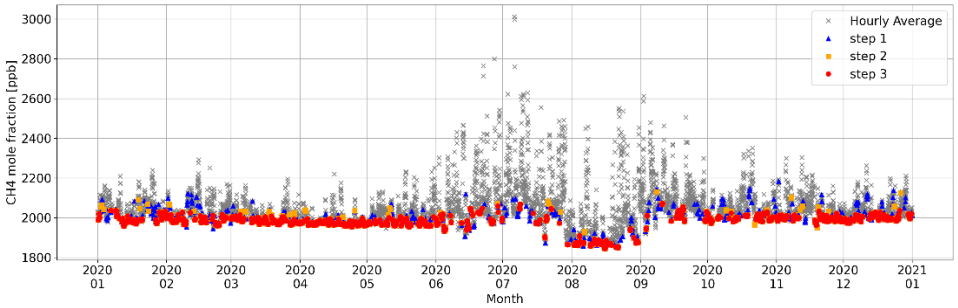
We also revised the table 2 and added the supplementary figure (Figure S2) to show our method.

Table 2. Criteria and percentage of selected background levels from observed data at each station.

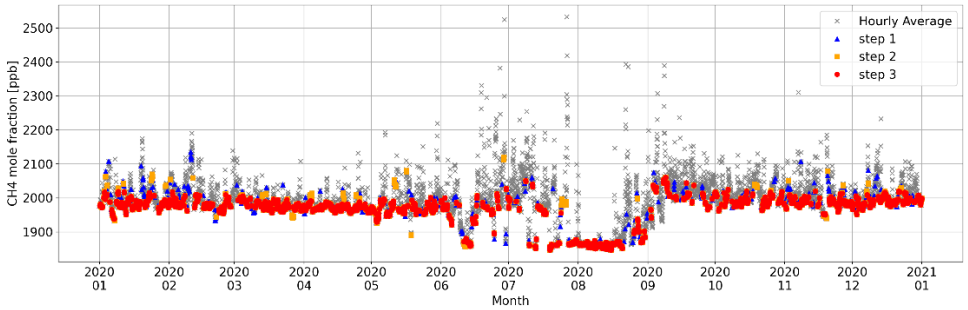
Station ID	AMY	JGS	ULD
Data period	1999 to 2020	2012 to 2020	2012 to 2020
A [ppb]	2.1	2.1	2.8
B [ppb]	4.9	5.2	3.6

C [ppb]		$1.8\sigma_{30d}$ (for all three stations)	
Spring, MAM [%]	29.1	46.6	57.9
Summer, JJA [%]	11.0	33.5	37.6
Autumn, SON [%]	16.9	30.9	53.2
Winter, DJF [%]	28.4	49.1	58.9
Total [%]	21.3	40.64	52.2

(a)



(b)



(c)

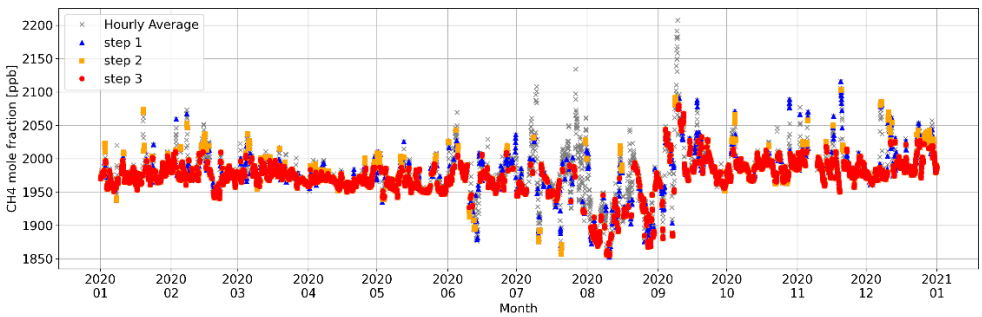


Figure S2. The time series of hourly CH_4 data through our selection method from step 1 to 3 at (a) AMY (b) JGS and (c)ULD in 2020.

4.7 Page 8 – line 15 – typo in HYSPLIT

Corrected

4.8 Figure 7 – The growth rates for AMY, JGS and ULD aren't shown for 2018. I think this is because they are negative (-1 ppb) but this is still a result so they should be shown.

Corrected

4.9 Why are some of the numbers in Table 4 written in bold?

Corrected

4.10 Figure 8 – 2009 was unusual because there was no seasonal cycle – can you comment on why this was?

When we reviewed the raw, L1 and L2 data, other period captured the summer drop but 2009 had an instrumental issue in summer. We commented in the manuscript.

P21 L20: In 2009 there is no clear seasonal cycle. There was an instrumental issue in summer while other period captured summer drop by observations.

4.11 I didn't fully understand the PSS analysis – I think that needs some more detail

To explain PSS analysis in detailed we revised the section 2.7.

Section 2.7: To identify and illustrate the potential source distributions for regional pollutions, we calculated the PSS using the trajectory statistics approach, which has often been applied to estimate the potential source areas of greenhouse gases (Reimann et al., 2004, 2008; Li et al., 2017). The trajectory statistics approach was introduced first by Seibert et al. (1994). The underlying assumption of the method is that elevated concentrations at an observations site are proportionally related to the air mass residence time on a specific grid cell over which observed air mass has been passing. Thus, this method simply calculates the airmass residence time weighted mean concentrations for target compounds (CH_4 in this study) for the domain with 0.5 × 0.5 grids using the following formula (Eq. 2):

$$\overline{C_{(i,j)}} = \frac{\sum_{a=1}^{M} T_{(i,j,a)} C_a}{\sum_{a=1}^{M} T_{(i,j,a)}} \quad \text{(Eq. 2)}$$

where $C_{(i,\,j)}$ represents the potential source strength of the grid cell i, j as a potential source region of the target compound (CH_4); a is the index of the trajectory; M is the total number of trajectories that passed through cell i, j; Ca is the enhanced mole fraction (difference from background mole fractions mentioned in section 3.2 below) measured during the arrival of trajectory a; and $T_{(i,j,a)}$ is the residence time of trajectory a spent over grid cell i, j and which were calculated using the method described by Poirot and Wishinski (1986) as following formula (Eq.3)

$$T_{(i,j,a)} = \sum_{n=1}^{N} \frac{S_{(i,j,n,a)}}{V_{(n,a)}} \quad \text{(Eq. 3)}$$

where $S_{(i,j,n,a)}$ is the length of that portion of the n^{th} segment of the a back-trajectory which falls over grid cell i,j. $V_{(n,a)}$ is the average speed of the air parcel as it travels along the n^{th} segment of the a back-trajectory. Backward trajectories were calculated using the HYSPLIT

model of the NOAA Air Resources Laboratory (ARL) using meteorological information from the Global Data Assimilation System (GDAS) model. For the trajectory reliability, we used only the 4-days (96h) backward trajectories at an altitude of 500 m above the mean sea level. To consider the influence of air masses on emissions at ground level, the air masses passing above the boundary layer height (BLH) were excluded. The BLH can was obtained from the HYSPLIT model. To exclude the influences of emission sources surrounding AMY stations, enhanced CH_4 data with wind speeds lower than 2 $m \cdot s^{-1}$ were excluded from the PSS analysis. When we compare PSS analysis among AMY, JGS, and ULD using CH_4xs data from 2016 to 2020, they showed the similar source regions while the coverage and CH_4xs are different (Fig. S5.).