

Overview:

Yirdaw berhe, et al., have submitted a manuscript comparing ground-based MIR-FTS measurements of atmospheric CH₄ and N₂O at Addis Ababa, Ethiopia to that of three satellite (MIPAS, MLS and AIRS) data products. The manuscript details the Addis Ababa site, the measurements made and the spectral processing procedure (including retrieval uncertainty estimates). A brief overview of the satellite data products used are given, then coincident comparison criteria and lastly analysis of the profile and partial column comparison results.

The novelty of this manuscript is that this is the first time Addis Ababa FTIR N₂O and CH₄ measurements are compared to satellite measurements.

The manuscript content is in the scope of the AMT journal. This research will be a welcome addition to already published literature concerning FTIR data from the Addis Ababa station, and also in the wider context of atmospheric ground-based trace gases measurements (including *in situ*) situated on the African continent (a data sparse region of the globe). Unfortunately, the manuscript is let down in multiple critical areas and I do not recommend publication until the issues listed below are addressed; either fixed or with a sufficient logical rebuttal.

Specific comments:

S1/ AMT English guidelines and house standards: A draw-back of the submitted manuscript is that I do not believe the grammar meets the standard required for publication in AMT. The authors are referred to AMT guidelines: https://www.atmospheric-measurement-techniques.net/for_authors/manuscript_preparation.html. There are instances of incorrect grammar use, ambiguous statements (most likely a consequence of improper grammar) and repetition of statements. All such instances need to be corrected. This is no reflection on the quality of the science presented and doesn't detract (only distracts and introduces ambiguity) from the novelty and importance of the presented subject matter (along with the effort the authors have already put into the manuscript). I would have expected the more experienced co-authors to have alerted the lead author to many of these grammatical and stylistic errors. For the manuscript review, correction of such grammatical errors will be left out (to speed up the review), and only commented upon if scientific clarity is required.

S2/ Could the authors clarify in the focus of the research. Comparisons are made between three satellite datasets and that of the ground-based FTIR measurements at Addis Ababa, but why? what is the motivation? In section 2.1 the manuscript alludes to why measurements at Addis Ababa are made, but only very broadly in a generic tropical atmosphere context. I gather the motivation is to use satellite measurements to validate the ground base measurements? This is unusual (usually the other way around), but a valid approach to help assess the quality of the ground-based measurements, if there is concern.

The authors state that the comparisons at the "Addis Ababa station is good to study tropical atmospheric processes" (Pg 19, L12). 'Good' in what context? Given the comparison results, will the ground-based CH₄ and N₂O measurements capture seasonal cycles and multi-year trends? Will biomass burning or other episodic events most likely be seen, and from what part of the tropics (the tropics is a large place)?

S3/ Pg 3, L19. As, in S2, the authors give a generic/broad scale reason for the importance of trace gas measurements in the tropics. I recommend that a more specific reason/motivation for Addis

Ababa measurements be stated in the context of physical (or chemical) processes (emissions) related more specifically to Addis Ababa and the atmospheric footprint it 'sees'.

S4/ Pg 4, L3-24. Retrieval information is incomplete, see comment T31 below. After this sentence the authors start describing the retrieval specifications (spectral Microwindows and model atmosphere layer scheme), then return to describing the optimal estimation method (L8-L24). It would be better to complete describing the retrieval theory prior to specific retrieval strategies. The information supplied in L8-L24 is ubiquitous and generic, I do not think it needs description. This section could be condensed to a single sentence stating Roger OEM approach is used (referenced) with Tikhonov regularization(reference).

S5/ Pg4, L28: Apriori is mentioned. Are the apriori profiles used static? i.e. unvarying, or are they changing seasonally, yearly, or daily? If the apriori is static, then how is it constructed, a mean over XX years? Is the apriori based on a certain global region?

S6/ Pg 5, Fig 1. 'Tropics' is a big area with a variable atmospheric state. Do the authors mean the apriori over the Addis Ababa region? Could the date of the Apriori temperature profile be put in the figure caption? Also, to show the reader the variability of the atmospheric state, could the 1-sigma SD at each layer be plotted. The authors could also possibly omit figure 1 completely, as information content is minimal.

S7/ Pg 5, L15. The description "with positive and negative signs..." can be removed. This is implicit in the retrieval. The authors should also describe the residuals. Are they dominated by random or systematic uncertainties? For instance, in Fig 2, the CH₄ fit residuals are dominated by systematic spectral error, most likely due to imperfections in the spectroscopic database line parameters.

S8/ Pg 5, L15. The authors mention an "optimised retrieval strategy" but only give a passing mention to the Tikhonov retrieval regularization scheme. This is an important part of the retrieval; influencing overall information content and interlayer correlations of information content. Could the author please describe the Tikhonov regularization parameters. Why was the Tikhonov scheme implemented instead of using apriori uncertainties? What type of smoothing constraint is used (L1, L2 etc..), were the smoothing constraints normalised using layer thickness? what is the alpha parameter used? and how was the alpha parameter selected? is the alpha parameter static? or varies per retrieval?

S9/ Pg 7, L4. Since equation 1 may be eliminated in section 2.2, insert equation 1 here or a reference to this equation in Rodgers, 2000.

S10/ Pg 9, L21. Section 3.3 should end and new section 3.4 "Timeseries" (or something similarly named) should start. The content from L22 onwards (to end of the section) is concerned with the timeseries, not explicitly error estimation. The sentence starting "Concentrations of CH₄..." (Pg 6, L1) should be moved into this new section.

S11/ Figure 6 and Pg 6, L1. Please state the reason why is only data from 2009 to 2013 is analysed? I assume the Addis Ababa station is still currently (up to 2019) taking measurements?

S12/ Section 4: this section details MIPAS, MLS and AIRS satellite-based measurement platforms. It would be more helpful if the focus of this section was on details about Addis Ababa overpasses for each platform (such as the number of 'good' overpasses as a proportion of total). This would also help diagnose if Addis Ababa is a 'good' site (as the authors have stated) for such satellite validation.

S13/ Pg 11, L1. Why was only the period Mar2009 to Dec2010 used in MIPAS Addis Ababa comparisons? Why not longer?

S14/ Pg 12, L3. The last two sentences, starting with "Nitrous oxide derived..." should be omitted as it refers to MLS data version 2.2, not 3.3, unless the authors state (after verifying) that the precision of MLS N₂O v3.3 is the same as v2.2.

S15/ Section 5 details and quantifies comparisons between the satellite data products and the Addis Ababa ground based FTIR data, but (in my opinion) does not elaborate on the results with respect to other ground based FTIR site measurements. Are the biases and spread seen at Addis Ababa like that of other ground-based FTIR sites (most likely also part of the NDACC)? This would help ascertain if the Addis Ababa is a 'good' validation site and is network comparable. All that is required is a literature review, this will help put the results derived in this study in context.

S16/ Equations 4 to 10 all pertain to statistical calculations between the FTIR and MIPAS measurements. I assume the same statistical methods are applied to comparisons with the other satellite data? Maybe make this section more generic, not just MIPAS specific.

S17/ Pg 14, L3. "Hence we will focus on the random uncertainties associated with...". This statement does not connect with the analysis in section 5.2. In section 5.2 dataset biases are quantified, which includes both random and systematic uncertainties (not separated). The standard deviations of the dataset comparisons will also include any systematic uncertainties. Maybe this sentence be retracted or changed to explain what is meant in a clearer manner.

S18/ Pg 14, L4. "However, the residual coincidence and horizontal smoothing errors...". If they are important why are they not investigated? The sentence starting on L19, pg14 ("In addition, the overestimation") also alludes (and offers conjecture) to issues arising around differences in the datasets relating to coincidence criteria but is not quantified. The authors could easily check this by changing the coincidence criteria (spatial and temporal) and see the effect of this in the dataset statistical differences.

S19/ Section 5.2 and Section 5.3. In both these sections there is no mention of how dataset degrees of freedom affect the profile differences. For example, the ground-based retrievals of CH₄ have approx. 2 DOFs. Differences at different altitudes will not be independent pieces of information. At Pg 14, L25 the authors state the bias of FTIR and MLS CH₄ at 18-20km is insignificant, at 17km -1.7%, and between 20-27km below 11%. Are these pieces of information independent? The authors may wish to comment on this fact and its implications.

S20/ Figures 10 and 11. There is no commentary on the large 'RD' differences above 30km (no sensitivity?). Could the authors comment on this?

S21/ Pg 17, L1. There needs to be a new section "section 5.4: Comparisons of partial columns" (or similar) starting at pg17 L1 if the authors are to start discussing partial column comparisons. Currently, the partial column comparisons for both N₂O and CH₄ are under section 5.3.

S22/ Pg 17, L1. Why are only MIPAS partial column comparisons conducted? What about other satellite data products?

S23/ Fig 12. Uncertainty/error bars could be added to all data points. This would help in assessing the comparisons.

S24/ Pg 19, L12. Define 'good', do the authors mean the measurement quality and retrievals are 'good', or the location, or both? Since the focus of the manuscript is on assessing the performance of the Addis Ababa FTIR measurements, explaining 'good' is quite important.

Technical comments (no particular order):

T1/ Title: FTIR should be expanded, not an acronym. There is no need for the chemical formulas. The word 'measurements' should also be added after FTIR. So...I recommend the full title should be along the lines of: "Methane and nitrous oxide from ground-based Fourier transform infrared spectrometer measurements at Addis Ababa: observations, error analysis and comparison with satellite data."

T2/ Pg 1, L2. Possible change: "total column abundances and vertical distribution of various constituents in the atmosphere" to "total column trace gas abundances and vertical distributions".

T3/ Pg 1, L4. The superlative sentence "They reveal the high quality of FTIR measurements at Addis Ababa" is not required. The data and analysis reveal this.

T4/ In the abstract, I do not think it is necessary to specify satellite data product versions, for example 'V5R_CH4_224'. This is done in the main body.

T5/ Pg 1, L12. There are phases throughout the manuscript of the sort "a positive bias of less than 0.14 ppmv (9%) is found in the altitude range of 21 to 27 km". I gather this means there is a maximum positive bias of 0.14 ppmv in the range 21 to 27km? This may be a better way to state it.

T6/ Pg 2, L1. CH₄, N₂O and CFCs are also stratospheric species...;)

T7/ Pg 2, L7. "ENIVSAT" to "ENVISAT satellite".

T8/ Pg 2, L9. Remove the word 'recent' from "The recent increasing..." (also replace 'to the' with 'on').

T9/ Pg 2, L10. Merge the sentences to read: "The recent increasing impact of CH₄ and N₂O to the global warming has also been assessed by the last AR4 IPCC report (IPCC, 2007; Sussmann et al., 2012), additionally N₂O will become the dominant ozone depleting substance emitted in the 21st century (Ravishankara et al., 2009)."

T10/ Pg 2, L11. What is IASI? Expand to say: "IASI instrument aboard the MetOp-2 satellite". MetOp-1 or MetOp-2...I can't remember.

T11/ Pg 2, L18. Rephrase first sentence: "In the tropics two important..."

T12/ Pg 2, L25. Replace 'launched' with 'taken'.

T13/ Pg 2, L27. Replace "The quality of ground based FTIR measurements" with "The Addis Ababa FTIR measurements".

T14/ Pg 2, L33. Replace 'confirm' with 'show'.

T15/ Pg 2, L34. Replace 'biased high and provided +14% as the most likely bias' with 'biased 14% high'.

T16/ Pg 2, L33. The reference Kenea throughout the manuscript should be replaced with Takele Kenea (2013)?

T17/ Pg 2, L33. The quoted references of Laeng (2015) and Plieninger (2016) refer to MIPAS comparisons with other satellite products. The paragraph starting at Pg 2, L25 concerns Addis Ababa FTIR measurements. There is a jump in topic. Reading as is, it could easily be taken that Addis Ababa measurements were used in these studies. These sentences should be removed or moved to a different part of the manuscript.

T18/ Pg 3, L3. The sentence “In this study, the previous work on intercomparison is extended to source gases CH₄ and N₂O from ground-based FTIR” is quite ambiguous. Either remove or make more specific to Addis Ababa.

T19/ Pg 3, L7. “approach” can be replaced with ‘strategy’.

T20/ Pg 3, L13. Is Addis Ababa part of the Network for the Detection of Atmospheric Composition Change (NDACC)? I suspect so, if this is the case it should be stated. The Takele-Kenea paper should be used as a site reference paper.

T21/ Pg 3, L15. Could the Addis Ababa site altitude (MASL) also be added?

T22/ Pg 3, L15. How is ‘suitability’ defined? Why is it suitable? Could possibly mention the amount of cloud free days a year.

T23/ Pg 3, L18. The superlative “extremely” can be removed, not needed.

T24/ Pg 3, L23. The superlative “very successfully” can be removed, not needed.

T25/ Pg 3, L24. Replace ‘sun’ with ‘solar’

T26/ Pg 3, L27. The sentence “This technique...” should be moved to precede the sentence “The high resolution...”

T27/ Pg 3, L28. “Using seven narrowband filters”. Assuming Addis Ababa is an NDACC site, do the seven filters meet NDACC specifications?

T28/ Pg 3, L29. It is mentioned an InSb detector is used to take measurements in over the range: 1500-4400cm⁻¹. There is no mention of detectors used to measure down to 750cm⁻¹, as mentioned in the prior sentence. Are measurements taken below 1500cm⁻¹?

T29/ Pg 3, L31. Replace “we used PROFFIT V...algorithm”, with “we used the retrieval code PROFFIT (Ver95)”.

T30/ Pg 4, L2. As the sentence reads, PROFFIT was developed to only retrieve CH₄ and N₂O. Could this sentence be corrected to reflect the fact PROFFIT was developed to retrieve multiple species.

T31/ Pg 4, L3. The sentence “This algorithm...” only tells half the information. Once a forward model calculation is completed, what happens next?

T32/ Pg 4, L4. At the end of the sentence “The vertical profiles...N₂O respectively” change to “N₂O respectively (see table 1 for spectral regions).” and could possibly be moved to section 3.1.

T33/, Pg4, L6. Could the bottom (base) layer height be stated.

T34/ Pg4, L27. replace ‘setup’ with ‘strategy’

T35/ Pg 4, L29. This is the first time the ‘NDACC’ and ‘IRWG’ acronyms are used, please state in full.

T36/ Pg 5, L4. I think the sentence “The vertical variability...” is not required, or more information is required, i.e. define ‘large’.

T37/ Pg 5, L8. “The micro-windows have been adopted from different sources”. Why, and could the sources please be referenced.

T38/ Pg, 5, L9. Modified Microwindows: why is this?

T39/ Pg 5, L7. After the end of the sentence “lines are presented”, the sentence Pg 4, L4 “The vertical profiles...” should be inserted.

T40/ Pg 5, L12. Remove the word “example”. Also, could the authors detail the Signal to Noise ratio (SNR) of the Addis Ababa spectra. All this information is part of the ‘optimization process’.

T41/ Table 1. could “int. Gases” be replaced with “Interfering gases”. In the table legend, could “column amounts” be replaced with “total column amounts”. Just an idea, (authors discretion), could another column be added for DOFs of the 0-27km partial column.

T42/ Figure 2. The usual convention is for measurement spectra data to be displayed as points (usually joined) and simulations as thin lines. The authors have the opposite of this. A minor point, and up to the authors if they would like to change it a more standard convention.

T43/ Figure 2 Legend. Acronyms are not explained prior, so the full name is required. Sorry. SEA is an unusual way to label/present the solar angle. Does SEA mean Solar elevation angle? The standard convention is solar zenith angle (SZA = 90.0 - SEA). I recommend that SZA be presented, not SEA. Spectra time is presented as “101715”, please reformat to 10h17m15s or similar. Is SD the root mean square difference of the measured spectra and forward model? I.e. RMS. Please define.

T44/ Pg 7, L2. To be pedantic, “the most” can be replaced with “an” (as a matter of opinion).

T45/ Pg 7, L11. The authors should state at this point the units of the AVKs displayed, normalised to layer VMR or not? i.e. [VMR/VMR]

T46/ Pg 7, L16. Pedantic point, but any AVK that has non-zero elements has ‘sensitivity’, that is, infer information from the spectra. 0.5 is an arbitrarily defined ‘cut-off’. So, I think a better statement would be “Fig 4 (top panel) shows that the.... has a sensitivity greater than 0.5 over the altitude range 2.45km to 27km”.

T47/ Pg 8, L2. Could the altitude ranges of the two independent partial columns be stated.

T48/ Pg 8, L4. The sentence “The amplitude...” repeats information given earlier so can be removed.

T49/ Pg 8, L5. The sentence “We also ignore”: Sorry I cannot understand this sentence (includes spelling mistakes), can it please be reworded and made clearer. The cited reference (Rinsland, 2005) does not contain any such information pertaining to altitude resolution. Can the authors please check the reference is correct or point out where it is in the paper.

T50/ Figure 4 legend. The datetime stamp of spectra the analysis is performed on should be given.

T51/ Pg 9, L4. Could the acronyms for sources of error be included, i.e. ‘instrument line shape (ILS)’, so that they correspond to the legend labels in fig 5.

T52/ Pg 9, L5. Can the term ‘zero baselines offset’ be ‘zero level baseline offset’

T53/ Pg 9, L5. I assume statistical error means 'random' error, maybe rephrase as "statistical (random) error". Also take out "typical" as it relates directly to an example.

T54/ Pg 9, L17. Change part of the sentence "for the profile", to "for the N₂O profile".

T55/ Pg 10, L6. The sentence starting "Vertical resolution..." needs to be referenced.

T56/ Pg 10, L8. The sentence starting "The analysis of the comparison..." should be moved to section 5.1.

T57/ Pg 11, L17. The starting sentence should be reworded: "...we have used the MLS N₂O (v3.3) product to validate the ground-based Addis Ababa FTIR measurements."

T58/ Pg 11, L18. The altitude levels are given in pressure coordinates. So far in the manuscript, the altitude units have been in kilometres. Could the authors include the geometric altitude as well as the pressure. i.e. "100 and 0.1 hPa (XX to XX km's)"

T59/ Pg 12, L7. This sentence needs a reference.

T60/ Pg 12, L11. Unfortunately, the sentence starting "The spectral resolution..." does not make sense, could this be reworded, rephrased. The following sentence can be omitted as it does not add any information concerning this study.

T61/ Pg 12, L13. Could the data product version of AIRS CH₄ be added?

T62/ Pg 12, L20. The word 'version' can be removed.

T63/ Pg 12, L22. I think the meaning of 'degraded' means smoothed, so maybe the sentence could read "...MLS have been degraded (smoothed) to make a ..." or replaced degraded with smoothed.

T64/ Pg 12, L29. Replace 'parameters' with 'statistics'.

T65/ Pg 13, L11. I assume that Sat_i(z) is smoothed? So maybe state that: "and the corresponding Sat_i(z) smoothed volume mixing ratio is derived from..."

T66/ Figures 7, 8 and 9 could be combined into a single 3x3 figure, so could figures 10 and 11 (2x2)

T67/ Pg 14, L26. "In the tropopause layer", could the tropopause layer be defined, i.e. "In the tropopause layer (~XX-XX km)"

T68/ Figures 7,8,9,10 & 11. The legends in the figures are slightly different. Maybe standardise these as the figure captions all reference back to fig 7 caption. Figure 7 caption should also explain/define the legend captions. For example: "mean difference FTIR minus MIPAS (MAD, blue solid line) ..."

T69/ Pg 16, L2. The first sentence could be abbreviated to "FTIR N₂O mixing ratio MIPAS comparison results are shown in fig 10."

T70/ Pg 16, L8. (-0.02 ppmv) ... include units please.

T71/ Pg 16, L11. Replace 'can' with 'could'.

T72/ Pg 16, L17. Add to end of sentence: "...the value derived from the FTIR is overestimated (relative to MLS)"

T73/ Pg 18, L4. The last part of the first sentence "..., which is a very useful..." is not needed.

T74/ Figure 12. The date label lacks information on the years(s), only months are given. Add information on the year(s) comparisons were made.

T74/ Acknowledgement: Remove full stop at the start of the first sentence. Remove the word 'besides'. Support, not supports.

T75/ References. The authors may wish to take out, or update references concerning an ACPD articles, such articles have not passed the peer review process. Reference formatting differs, so would be good to get it all consistent. Decide on an author convention for Samuel Takele Kenea, as this author is referenced a few times, but referenced differently (Takele Kenea S.).