C1. Did the authors modify the AERIoe code itself or did they develop a new code base from scratch? Please state in the Data availability if/where/how the fast AERIoe code is available. (Proprietary or open source? How does one obtain it?).

R1: Thank you for the comments. The code for the Fast AERIoe algorithm was developed by ourselves written in MATLAB language. The code for recalculating the Jacobians are not publicly available at this time but may be obtained from the authors upon reasonable request.

C2. Given that the main goal is to reduce the computation time, specifics in that regard are needed. Has the code timing been analyzed and what are the bottlenecks? I assume calculation of the Jacobian is the main bottleneck; is that the case?

R2: We appreciate the suggestion. The code timing has been analyzed as follows if we understand correctly:

The codes for the retrieval algorithm are written in MATLAB language and runs on a Lenovo Aircross 510P computer, of which the CPU is Intel Core i7-7700 and the operating system is Ubuntu 14.04. To analyze the code timing of the retrieval algorithm, the code was divided into the following sections: preparation, iteration 1, iteration 2, iteration 3,... and iteration final. The preparation section mainly consists of atmosphere construction, observation vector construction and pre-calculated variables importation. The iteration sections include the recalculation of $\mathbf{K}$ and $F(\mathbf{X})$ and the inversion using equation (1). Note that iteration 1 does not need to calculate $\mathbf{K}$ and $F(\mathbf{X})$ because the a prior profile $\mathbf{X}_a$ is fixed (mean value of the atmosphere), and the $\mathbf{K}$ and $F(\mathbf{X})$ associated with it are pre-calculated. The time consumed by each section was analyzed both for AERIoe and Fast AERIoe, results for an arbitrarily selected case are provided in Table 1. The recalculation of $F(\mathbf{X})$ and $\mathbf{K}$ consumed an immense amount of time in the retrieval process of AERIoe, and the latter is the most time consuming section. Therefore, by reducing the recalculation of $\mathbf{K}$, the retrieval time of Fast AERIoe is greatly reduced compared to AERIoe.

Table 1. List of time consumption (units: s) by the sections of AERIoe and Fast AERIoe. The sections denoted with superscript "*" indicate that $\mathbf{K}$ is not recalculated during Fast AERIoe retrieval process.

Sections		AERIoe	Fast AERIoe
preparation		0.29	0.22
iteration 1	inversion	0.29	0.22
iteration 2	recalculation of $F(\mathbf{X})$	17.11	16.69
	recalculation of $\mathbf{K}$	68.76	70.27
	inversion	0.31	0.27
iteration 3	recalculation of $F(\mathbf{X})$	17.18	17.04
	recalculation of $\mathbf{K}$	70.55	0.00
	inversion	0.22	0.22
iteration 4	recalculation of $F(\mathbf{X})$	17.71	16.36
	recalculation of $\mathbf{K}^*$	70.07	0.00
	inversion	0.25	0.21
iteration 5	recalculation of $F(\mathbf{X})$	16.97	17.38
	recalculation of $\mathbf{K}^*$	68.93	0.00
	inversion	0.21	0.25
iteration 6	recalculation of $F(\mathbf{X})$	16.08	15.08
	recalculation of $\mathbf{K}^*$	68.23	0.00
	inversion	0.24	0.24
iteration final	recalculation of $F(\mathbf{X})$	15.91	18.45

	recalculation of $\mathbf{K}^*$	68.11	0.00
	inversion	0.28	0.23

C3. Please begin with a sentence that more clearly gives the background - something like: "Two methods for retrieving … are physical and statistical retrieval algorithms …"

R3: Thanks for the constructive suggestions to improve our manuscript. The corresponding part in our revised submission is given as follows:

Two methods for retrieving temperature and water vapor profiles from the Atmospheric Emitted Radiance Interferometer (AERI) observations are physical and statistical retrieval algorithms. The physical retrieval algorithm, named AERI Optimal Estimation (AERIoe), outperforms statistical retrieval algorithms in many aspects except the retrieval time, which is significantly increased due to the complex radiative transfer process.

C4. Line 12: Begins with "Further analysis showed…" but no analysis has yet been discussed. What changes were made to the Jacobians and why was that expected to speed up performance (but didn't)?

R4: Thank you for pointing out this problem. We have modified the problem in our revised submission. The corresponding revised part in our revised submission is given as follows:

Analysis of the change of AERI observations' information content with Jacobians revealed that the performance of AERIoe algorithm had little dependence on Jacobians. Thus, the Jacobian matrix could remain unchanged when the variation of iterative profiles is small in the retrieval process. This significantly reduces the amount of computation and thus increases the retrieval speed of AERIoe.

C5. The time estimates are not useful without knowing what type of computing platform was used. Perhaps just give the percent improvement. Also, are 3 significant figures warranted here?

R5: Thank you for the suggestion. We have modified the problem in our revised submission. The corresponding revised part in our revised submission is given as follows:

The retrieval speed was significantly improved compared with the original AERIoe algorithm under the condition that the parameters of the computing platform remain unchanged, resulting in an average retrieval time reduction by 58.82%.

C6. What is meant by "certain impact"? What is meant by "to some extent"? Why not state the convergence rate of the traditional algorithm?

R6: Thank you for pointing out the problems. We have modified them in our revised submission. The corresponding revised part in our revised submission is given as follows:

Results based on synthetic observations revealed that the fast retrieval algorithm reached an acceptable convergence rate of 98%, which is slightly lower than the 99.88% convergence rate of AERIoe for the 826 cases used in this study.

C7. The authors say that "The retrieval accuracy of the fast retrieval model is equivalent to that of the traditional algorithm." However, on lines 346-348 differences indicate that the accuracies are not equivalent.

R7: Thank you for pointing this problem out. We have modified the sentence as "The retrieval results of the fast retrieval model are comparable to that of AERIoe." Reasons for this modification are discussed in detail in the response to comments on lines 346 - 348.

C8. How is the convergence criteria adjusted to give reliable retrieval results? It was previously stated that the results were equally accurate. Do you mean they are equally accurate when they both converge?

R8: Thank you for the comments. The sentence 'However, reliable retrieval results can still be obtained by adjusting the convergence criteria.' has been removed.

C9. Line 115: If the authors are using X0 = Xa, they should replace X0 with Xa in Eqn (1) so it is consistent with Turner and Lohnert. If not, they should explain this change.

R9: Thanks for the suggestion. We have replaced $\mathbf{X}_0$ with $\mathbf{X}_a$ in Eqn (1).

C10. Line 116 says "Y is the observed radiance vector, F(X) is the AERI observed spectrum…" Is it rather that Y is the observed radiance vector (from the observed AERI spectrum) and F(X) is the estimate of Y from the forward model calculation? It would also be helpful to define that the background refers to the a priori atmospheric state, if that is the case.

R10: Thanks for the suggestion. We have replaced '$F(\mathbf{X})$ is the AERI observed spectrum' with '$F(\mathbf{X})$ is the computed radiance for $\mathbf{X}$'. The sentence '$\mathbf{S}_a$ it the background covariance matrix' has been changed to '$\mathbf{S}_a$ is the a priori covariance matrix'.

C11. Eqn. 1: I'm curious why this formulation is used instead of the Levenberg-Marquardt formulation (Rodgers 2000, Eqn 5.36). How is the behavior the same or different? Carissimo et al. 2005 state that their method is almost equivalent to Levenberg-Marqardt. In Levenberg-Marquardt, increasing gamma decreases the step size and makes the retrieval weighted more toward steepest descent. How is the formulation here the same or different?

R1: Thank you for the comments, we have put our thinking caps on this question and the discussion regarding it is presented as follows:

The iterative equation for AERIoe is as follows

$$\mathbf{X}_{n+1} = \mathbf{X}_a + \left(\mathbf{K}_n^T \mathbf{S}_e^{-1} \mathbf{K}_n + \gamma \mathbf{S}_a^{-1}\right)^{-1} \mathbf{K}_n^T \mathbf{S}_e^{-1} \times \left(\mathbf{Y}^m - F(\mathbf{X}_n) + \mathbf{K}_n(\mathbf{X}_n - \mathbf{X}_a)\right) \tag{1}$$

The equation for the Levenberg-Marquardt method is given as follows (Rodgers, 2000)

$$\mathbf{X}_{n+1} = \mathbf{X}_n + \left(\mathbf{K}_n^T \mathbf{S}_e^{-1} \mathbf{K}_n + (1+\gamma) \mathbf{S}_a^{-1}\right)^{-1} \times \left(\mathbf{K}_n^T \mathbf{S}_e^{-1}\left(\mathbf{Y}^m - F(\mathbf{X}_n)\right) - \mathbf{S}_a^{-1}(\mathbf{X}_n - \mathbf{X}_a)\right) \tag{2}$$

We believe that the comparison of the two methods can be analyzed from two aspects: the retrieval process and the retrieval results.

(1) In terms of retrieval process, Carissimo et al. (2005) state that gama in Eqn (1) dumps the width of the step between two consecutive iterates and leads its direction toward the steepest descent of the cost function. The gama in Eqn (2) is chosen at each step to reduce the cost function and also tends to the steepest decent of cost function. Therefore, the role of gama in Levenberg-Marquardt method is equivalent to that of AERIoe.

However, the values of gama in the two formulas are quite different, as the profiles in Eqn (1) are retrieved by adjusting the a prior profile $\mathbf{X}_a$, while the profiles in Eqn (2) are iterated by adjusting the iterative profile $\mathbf{X}_n$. For example, in the work of Foth and Pospichal (2017) the initial value of gama is 2, and increases by a factor of 10 if the cost function **J** has increased and reduces by a factor of 2 if **J** has decreased. Therefore, the value of gama in Levenberg-Marquardt method shows a significant difference from AERIoe [γ = 1000,300, 100, 30, 10, 3, 1,...].

(2) In terms of retrieval results, gama in Equation (1) needs to be set to 1 in the final step to eliminate regularization errors in the solutions, which gives

$$\mathbf{X}_{n+1} = \mathbf{X}_a + \left(\mathbf{K}_n^T \mathbf{S}_e^{-1} \mathbf{K}_n + \mathbf{S}_a^{-1}\right)^{-1} \mathbf{K}_n^T \mathbf{S}_e^{-1} \times \left(\mathbf{Y}^m - F(\mathbf{X}_n) + \mathbf{K}_n(\mathbf{X}_n - \mathbf{X}_a)\right) \tag{3}$$

It can be seen that $\left(\mathbf{K}_n^T \mathbf{S}_e^{-1} \mathbf{K}_n + \mathbf{S}_a^{-1}\right)^{-1} \cdot \mathbf{K}_n^T \mathbf{S}_e^{-1} \mathbf{K}_n$ is equivalent to $I - \left(\mathbf{K}_n^T \mathbf{S}_e^{-1} \mathbf{K}_n + \mathbf{S}_a^{-1}\right)^{-1} \cdot \mathbf{S}_a^{-1}$, thus Eqn (3) can be written as

$$\mathbf{X}_{n+1} = \mathbf{X}_a + \left(\mathbf{K}_n^T \mathbf{S}_e^{-1} \mathbf{K}_n + \mathbf{S}_a^{-1}\right)^{-1} \mathbf{K}_n^T \mathbf{S}_e^{-1}\left(\mathbf{Y}^m - F(\mathbf{X}_n)\right) + \left(I - \left(\mathbf{K}_n^T \mathbf{S}_e^{-1} \mathbf{K}_n + \mathbf{S}_a^{-1}\right)^{-1} \mathbf{S}_a^{-1}\right)(\mathbf{X}_n - \mathbf{X}_a) \tag{4}$$

So the solution becomes

$$\mathbf{X}_{n+1} = \mathbf{X}_n + \left(\mathbf{K}_n^T \mathbf{S}_e^{-1} \mathbf{K}_n + \mathbf{S}_a^{-1}\right)^{-1} \times \left(\mathbf{K}_n^T \mathbf{S}_e^{-1}\left(\mathbf{Y}^m - F(\mathbf{X}_n)\right) - \mathbf{S}_a^{-1}(\mathbf{X}_n - \mathbf{X}_a)\right) \quad (5)$$

Eqn (5) is a special case of Eqn (2) when gama in Eqn (2) is chosen to be 0. Therefore, the Levenberg-Marquardt is a combination of a Gauss-Newton (without gama) and steepest descent minimization technique and equivalent to AERIoe.

References:

Rodgers, C. D.: Inverse methods for atmospheric sounding: theory and practice, World scientific, 119-120 pp., ISBN9814498688, 2000.

Foth, A. and Pospichal, B.: Optimal estimation of water vapour profiles using a combination of Raman lidar and microwave radiometer, Atmos. Meas. Tech., 10, 3325-3344, https://doi.org/10.5194/amt-10-3325-2017, 2017.

C12. Figure 1: This figure needs improvement and explanation. E.g. please define "iterative observations" and "iterative profiles" in the caption. Use of the symbol "Sa" is inconsistent with use of "Jacobians" instead of "K". K_Index has not yet been defined.

R12: Thanks for the suggestion. We have changed "iterative observations" to "**F**(**X**)" and "iterative profiles" to "$\mathbf{X}_n$". In order to be consistent with "$\mathbf{S}_a$", the symbol "**K**" was used instead of "Jacobians". A symbol "compute monitoring index" was used instead of "K_Index" and defined it in the caption as follows:

The monitoring index reflects the variations of $\mathbf{X}_n$.

C13. Line 118: I don't think n is the number of iterations, but rather the iteration number.

R13: Thanks for the suggestion. We have changed 'the number of iterations' to 'the iteration number'.

C14. Line 120: The description of how gamma is used is not clear.

R14: Thanks for the suggestion. The description of how gamma is used is given as follows:

AERIoe begins with scalar γ of large values is to stabilize the retrieval and ends with a unity γ to add more information from AERI observations as n increases. This approach allows the AERIoe algorithm to overcome a poor first guess and achieve a results that have the most information from AERI observation, the detailed description of how γ is used can be found in Turner and Löhnert (2014).

C15. Line 122: Remove "progress".

R15: Thanks for the suggestion. The word 'progress' in Line 122 has been removed.

C16. Line 122: Please change "is not allowed to converge until…" to "Iterations are continued until…" if that is what is meant here.

R16: Thanks for the suggestion. This sentence was modified as "Iterations are continued until γ decreases to 1 and the following convergence criterion is satisfied."

C17. Line 124: Use consistent symbols. You have superscript n sometimes and subscript n other times.

R17: Thanks for the suggestion. We have changed all of superscript n into subscript n.

C18. Line 214-215: It is not true that "what affects IC and DFS lies only in gamma and Jacobian". In fact, when gamma = 1, IC and DFS are determined by Se and Sa, with the purpose of the Jacobian being to transform Se into the state space for Sa, so that they have the same units and size (rows and columns). I think what you mean is that IC and DFS only change with iteration due to changes in gamma and the Jacobian. (But see below).

R18: We appreciate the suggestion. The sentence in line 214-215 is rephrased as follows:

It can be seen from equations (6) and (7) that IC and DFS are determined by $\mathbf{S}_e$, $\mathbf{S}_a$, **K** and γ. However, $\mathbf{S}_a$ and $\mathbf{S}_e$ remain unchanged during retrieval, which makes IC and DFS only change with iteration due to variations in γ and **K**. As γ drops to 1 at the final iteration, the values of IC and DFS are only dependent on **K**.

C19. Lines 214 - 232: I don't understand the logic here. On line 224, it is strange to say that gamma changes with the adjustment of the profile, since gamma is prescribed. Figure 3 is confusing. The x-axis goes in the reverse direction as the retrieval proceeds, the figure caption description seems to be wrong (red is actually K_index), and it isn't stated where K_index starts and ends (starts at the high end, ends at the low end?). It is not surprising that the DFS and IC increase as gamma drops to 1, since gamma weights the retrieval away from the observation and toward the first-guess, which presumably has no information content at all. It is also not surprising that there is not much change in DFS and IC with the Jacobian, since, as stated previously, the purpose of the Jacobian here is to transform Se onto the dimensions of Sa. I don't see how this shows that the change of the Jacobian has less influence on the retrieval ability than gamma. Gamma is not supposed to influence the retrieval ability, but only the retrieval stability. That is why iterations are continued until gamma is 1, whereupon the retrieval equation is equivalent to the Gauss-Newton formulation and the maximum information content is used. In fact, I don't see the point of this paragraph or figure at all. The authors could simply state that if X is not changing much, as evidenced by the K_Index, then the Jacobian is probably not changing much either, and therefore does not need to be recomputed. (Note, however, that this is not necessarily true, and they need to show that it is an ok approximation).

R19: We appreciate the suggestion. We admit that the logic here is indeed a bit confusing. Fig 3 and the analysis of this picture are reworked, which are given as follows:

The K_Index, ranged from 0 to 260, was obtained by multiplying the a prior profile by different scale factors. This range covers most of the K_Index during AERIoe retrieval process (see Fig. 4). The atmosphere dependent **K** was computed by LBLRTM with a prior profiles multiplied by different scale factors, and the SIC and DFS corresponding to the Jacobians above were calculated by equations (6) and (7), respectively. Fig. 3 shows the curve of SIC and DFS changed with K_Index. Both of them change slowly with K_Index, with the variation of SIC within 13.46% (from 13.89 to 16.05), and DFS within 4.38% (from 3.71 to to 3.88) for temperature and within 12.73% (from 1.44 to 1.65) for water vapor, which demonstrates that SIC and DFS remain almost unchanged as K_Index increases on the condition that the value of K_Index is small. This provides an effective means to improve the retrieval speed of AERIoe by recalculating **K** selectively when **X** is not changing much or K_Index is small. This could be achieved by comparing the value of K_Index and its threshold at each iteration to determine whether **K** is recalculated or not.

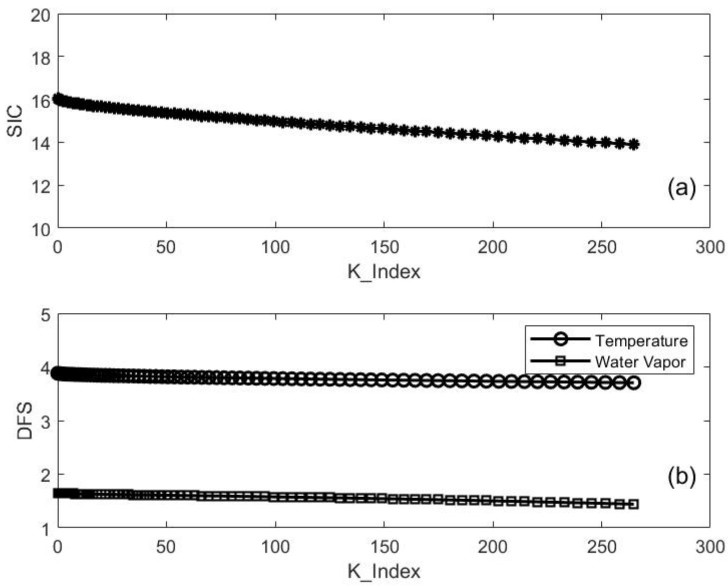

Figure 3. (a) The change of SIC with K_Index. (b) The change of DFS with K_Index for temperature (unfilled circles) and water vapor (open squares), respectively.

C20. Turner and Lohnert state that "Future versions of AERIoe will use the Carissimo et al. (2005) approach in order to more efficiently converge and reduce computational time." Did the authors explore that approach, and how might that change their analysis?

R20: Thank you for the comments. In fact, the first approach we studied to reduce the retrieval time happens to be the L-curve method recommended by Turner and Lohnert (2014). In our study the codes in Regularization Tools developed by Per Christian Hansen were used to calculate the L-curve and locate the corner, and the retrieval results using gama from the L-curve method did not show any superiority over the method used in AERIoe. The reason lies in that the gama obtained from L-curve method does not gradually decrease with iterations in the retrieval process. In order to figure out the impact of gama on the iterations of AERIoe, we find that the change of **K** is negligible during most of the retrieval process, which inspired this study to reduce the retrieval time by recalculating **K** selectively.

C21. Sections 3.1 and 3.2. are unclear. The description of the retrieval forms is confusing. Is the state vector comprised of the temperature and log of water vapor on the 37 atmospheric layers? Why isn't it parametrized, given that there are far fewer degrees of freedom? It continues to be difficult to tell what is new here and what is the same as previous work. Please avoid repeating details where you could reference the previous work. For example, you could say, "The forward model is the same as that described by Turner and Lohnert, except as follows…" Was LBLRTM used here to apply the spectral response function, in contrast to the previous work? It is stated that LBLRTM can be used to calculate the Jacobian. Was it used for this purpose? Again, is this a departure from previous work?

R21: Thanks for the suggestion. We have removed Sections 3.1 and 3.2. from the manuscript and added descriptions of Fast AERIoe configurations that differ from AERIoe in the paragraph between line128 and line136. The new paragraph is reorganized as follows:

Note that **K** depends on **X** used for estimating the Jacobian, which means that **K** must be recomputed for every iteration step. The updating of the Jacobians in the retrieval process requires the calculation of the optical thickness or radiance (intensity) with respect to different atmospheric constituents at each height, which might be computationally expensive depending on the lengths of **X** and $\mathbf{Y}_m$ (Maahn et al., 2020). Owing to the constraints of γ, the decrease of the difference between simulated and observed radiation is not very much in the adjustment of individual iterations to the retrieval profile. At this time, the change in the Jacobian calculated as per the iteration profile is negligible. Backed by the above analysis, a fast iterative algorithm called Fast AERIoe is proposed on the basis of the AERIoe algorithm. The flowchart of Fast AERIoe is shown in Fig. 1, most of the configurations are consistent with AERIoe described by Turner and Löhnert (2014), except some modifications highlighted as follows:

a. atmospheric configurations: The height grid of **X** is consistent with AERIoe, but the maximum retrieval height is limited to 3 km. This is done because the variations of K above 3 km is negligible due to the fact that most of the information content in AERI spectrum lies in the lowest 2 km of the atmosphere for temperature and water vapor profiles (Turner and Löhnert, 2014). The cloud properties were excluded from the state vector **X**, which is beyond the scope of this study. The corresponding priori profile $\mathbf{X}_a$ and the a priori represented by the Sa covariance matrices are modified to be consistent with **X**.

b. observational vector $\mathbf{Y}_m$: Spectral regions that sensitive to cloud properties were removed from the observational vector $\mathbf{Y}_m$ to be consistent with the state vector **X**. Furthermore, additional observations including surface temperature and water vapor were incorporated into the observation vector, details are described by Turner et al. (2019).

c. Jacobian matrix K: K is derived from the line-by-line radiative transfer model (LBLRTM; Clough et al. 1992), which is the same as AERIoe except the version (12.8 instead of 12.1). Another modification is that K is not recomputed to improve the retrieval speed of the algorithm when the variations of the iterative profile $\mathbf{X}_n$ are small.

C22. Line 176: Please remove the statement that LBLRTM is the most accurate forward model or provide a reference for it.

R22: Thanks for the suggestion. The statement that LBLRTM is the most accurate forward model has been removed.

C23. Line 196-197: Please rephrase this: "…determined whether updating or not by monitoring the indicators that can reflect the changes of Jacobian in the iterative process".

R23: Thanks for the suggestion. The sentence in line 196-197 has been modified as follows:

Adaptive updating of **K** is the key to reduce the calculation amount of the AERIoe algorithm. The Jacobians are dependent on the atmospheric constituents, which means that **K** must be recalculated for every iteration step. The question arises as to under what circumstances **K** does not need to be recalculated. Therefore, the dependence of the retrieval capability on Jacobians must be analyzed and the indicators that reflect the changes of Jacobians should be figured out to determine whether **K** recalculated or not.

C24. Line 344: Please change, "with only slight differences in BIAS metrics between 500 m and 1.5 km" to include a quantitative value, such as "with differences within x% to y%"

R24: Thanks for the suggestion. The sentence "with only slight differences in BIAS metrics between 500 m and 1.5 km" was changed to "with maximum differences within 0.06 K in BIAS metrics between 500 m and 1.5 km."

C25. Lines 346 - 348: It seems like "a maximum increase of 0.29 g/kg in BIAS and a maximum of 0.32 g/kg in RMSE" are significant. The bias increase appears to be up to about 40% (0.7 to 0.9), and the RMS increase up to ~12%. This does not seem to me to be comparable retrieval accuracy. Please clarify.

R25: Thanks for pointing this problem out. We've made two modifications to Fig. 9. One is that the radiosonde observations have been smoothed with the averaging kernel **A**, as suggested by Reviewer #1, to reduce the vertical representativeness errors.

$$\mathbf{X}_{sonde}^{smoothed} = \mathbf{A}\left(\mathbf{X}_{sonde} - \mathbf{X}_a\right) + \mathbf{X}_a, \tag{13}$$

Another modification is the transformation of water vapor into the form of log(ppmv). This is due to the fact that the unit of **K** output by LBLRTM is log (ppmv) during the retrieval, which means that the adjustment of the iterative profile is also in the form of log (ppmv). Therefore, we believe that it is more reasonable to compare water vapor profiles of different retrieval algorithms in the form of log(ppmv).

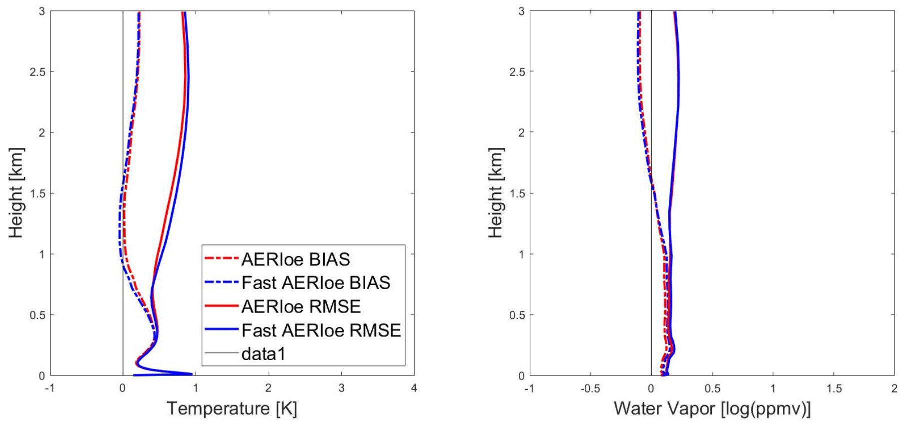

Figure 9. Bias (solid curves) and RMSE (dashed curves) profiles for the comparisons of AERIoe(red curves) and Fast AERIoe (blue curves) retrievals with radiosondes. (Left) Temperature profile, (right) Water Vapor profiles.

The Fast AERIoe retrieved temperature profiles shows a negative deviation of 0.05K between 1.0 km and 1.5 km and a maximum increase of RMSE up to 13.3% (from 0.60 K to 0.68 K) above 1.0 km when compared with AERIoe. For the water vapor profile, the BIAS and RMSE profiles of Fast AERIoe are in good agreement with AERIoe, except for a maximum increase of BIAS up to 25% (from 0.12 log(ppmv) to 0.15 (ppmv)) bellow 1.0 km. Therefore, we agree with the reviewers' comments that Fast AERIoe does not reach a comparable retrieval accuracy. However, we believe that the retrieved profiles rather than the retrieval accuracy of the two algorithms are comparable. This is because the increment of BIAS (within 0.03 log(ppmv)) is much smaller than the value (roughly on the order of 5-10 log(ppmv)) of the retrieved water vapor profile itself, and this increment had little impact on the retrieval results of AERIoe. The comparison of the profiles retrieved by the two algorithms can be demonstrated more clearly by the modified Taylor plots (see Fig. 10). Most of the blue and red symbols ‘ × ’ in Fig. 10 , which indicate the scores for the individual profiles of the two algorithms, are closed to each other. Therefore, the retrieval results of the AERIoe and Fast AERIoe algorithms are comparable both for temperature and water vapor profiles.

To be more precisely, we have rephrased all the sentence like “the retrieval accuracy of Fast AERIoe is comparable to that of AERIoe” in the manuscript to “the retrieved profiles of Fast AERIoe is comparable to that of AERIoe.

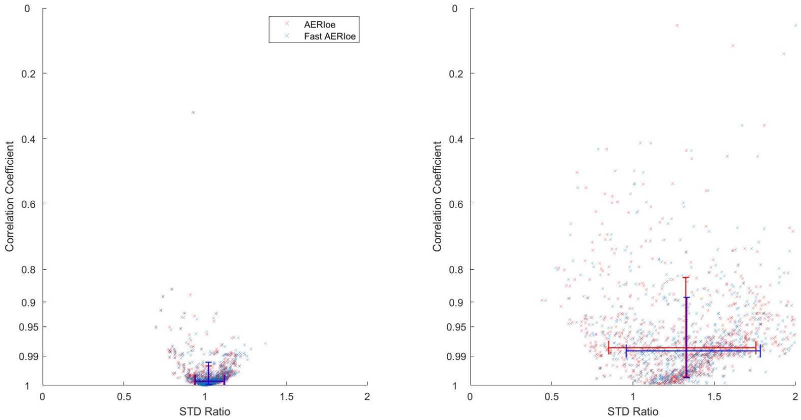

Figure 10. Modified Taylor plots showing the correlation coefficient and standard deviation ratio between the smoothed radiosondes and the retrieved (a) temperature and (b) WVMR profiles using AERIoe (red symbols) and Fast AERIoe (blue symbols). There are 826 cases from the SGP site within 2012. Each symbol indicates the score for an individual profile. The arms of the plotted crosses span the 10th–90th percentiles for the correlation coefficient (vertical arms) and the standard deviation ratio (horizontal arms).

C26. Line 350: More detail is needed about how you calculated "Pearson's correlation coefficient between two datasets on the y-axis and the ratio of the standard deviation on the x-axis", and the caption of Fig. 10 needs to be improved.

R26: The details to calculate the variables in Fig. 10 are given as follows:

Each retrieval/sonde pair is used to derive the correlation coefficient (r) from Eq. (14) and the ratio of the standard deviations from Eq.(15), both are used by Turner and Löhnert (2014).

$$r = \frac{\frac{1}{N}\sum_{z=0}^{z=h}\left[s(z)-\bar{s}\right]\left[a(z)-\bar{a}\right]}{\sigma_s \sigma_a}, \quad (14)$$

$$SDR = \sigma_a / \sigma_s, \quad (15)$$

Within the equations, $s(z)$ and $a(z)$ are defined as the radiosonde observations and retrieved profiles between 0 and 3 km, $(\bar{s}, \bar{a})$ and (σ_s, σ_a) are the mean values and standard deviations at the same height range.

We remove the sentence 'The intersection of the arms represents the location of the median correlation coefficient and standard deviation ratio of the datasets.' in the caption of Fig.10. The caption of Fig. 10 is modified as follows:

Modified Taylor plots showing the correlation coefficient and standard deviation ratio between the smoothed radiosondes and the retrieved (a) temperature and (b) WVMR profiles using AERIoe (red symbols) and Fast AERIoe (blue symbols). There are 826 cases from the SGP site within 2012. Each symbol indicates the score for an individual profile. The arms of the plotted crosses span the 10th – 90th percentiles for the correlation coefficient (vertical arms) and the standard deviation ratio (horizontal arms).

References:

C27. For use of FTIR viewing solar spectra, you could also reference the work of Kimberly Strong's group; e.g.: https://amt.copernicus.org/articles/7/1547/2014/

For retrievals from AERI, please add a reference Rowe et al. 2006, which used constrained linear inversion to retrieve temperature profiles from an AERI instrument: Rowe, P.M., Walden, V.P. and

Warren, S.G., 2006. Measurements of the foreign-broadened continuum of water vapor in the 6.3 μm band at− 30° C. Applied optics, 45(18), pp.4366-4382.

R27: Thank you for the suggestion. We have read the references recommended by the reviewers carefully. The two references have been added into the revised manuscript.

C28. English grammar and clarity: If possible, please have a native English speaker edit your paper throughout, including use of "the" in English, which is very challenging to get right. Examples:

Line 25 and throughout: When you are talking about something in general, omit the word "the" and change the noun to the plural. Examples: change "the observation network" to "observation networks". Change "the convective scale numerical weather prediction system" to "convective scale numerical weather prediction systems". Change "the radiosonde profiles" to "radiosonde profiles". Only use "the" if you are talking about a specific thing, and if you have made it clear which one you are talking about. For example, on line 73, add "the" before "ARM program" since it is clear which program you are talking about (ARM).

R28: Thank you for the suggestion. We have modified the issues identified by the authors in the revised manuscript. Moreover, we have carefully polished our paper to improve the English writing. If there are still writing problems with our manuscript, please let us know, and we will try our best to revise our article.

C29. Line 26: Please define all acronyms (e.g. NWP model)

R29: Thank you for the suggestion. We have checked our manuscript carefully and modified the similar problems in our revised submission.

C30. Line 32: Change "shows" to "has" or "demonstrates". Line 57: remove "to boot".

R30: Thank you for the suggestion. We have modified the problems in our revised submission.