We thank the reviewers for the examination of our work. Here are the point-by-point responses to their comments. In blue are our answers, and the text that has been added to the manuscript, is given in ***blue and italic***.

Reviewer 1:

Comments and answer:

The hydroxyl (OH) radical is the pivotal oxidant in the atmosphere, its accurate measurement is essential for understanding atmospheric oxidation capacity, and thus air quality, and climate change. However, the accuracy of OH measurements using the laser-induced fluorescence (LIF) technique is unavoidably compromised by wavelength drift of the excitation laser. Therefore, a reference system capable of monitoring the laser output wavelength and actively locking it to the optimal OH excitation line is crucial for long-term, reliable measurements. In this manuscript, Chen et al. develop a compact real-time reference system for wavelength locking in LIF-FAGE measurements. Through comprehensive characterization of key parameters, an optimal operational window for stable and high-concentration OH generation and detection is identified. The system's high stability is convincingly demonstrated by continuous measurements over 12 hours, with exceptionally low drift (0.2% per hour) during the first 9 hours. The detailed description of this system will be valuable for other researchers aiming to implement wavelength-locking techniques to improve the stability of their measurement systems. Overall, I recommend its publication after considering the following minor comments:

- In lines 74-75, it mentioned that the AIOFM group used a 282 nm laser as excitation laser, I know that they have updated the excitation laser to 308 nm, please see this paper (*Wang et al., "Development of a field system for measurement of tropospheric OH radical using laser-induced fluorescence technique".)* for detail information, and add the corresponding updated description to the main text.

  Answer: We thank the reviewer for pointing this out. We have added this critical information to the manuscript, and corresponding descriptions has been adjusted:

  The AIOFM group has provided initial characterization data for their earlier version reference cell, however, a 282 nm laser was employed (Xing et al., 2017). In their upgraded system, a 308 nm excitation laser was used, and good stability was reported: a fluorescence fluctuation of approximately 1.2% over 15 minutes at the resonance wavelength (Wang et al., 2019).

- In lines 292-293, it says that, during the 12-hour test, the signal attenuation never exceeded the threshold to trigger the wavelength re-locking program. For a complete performance evaluation, it would be informative to know if and how the system performs when a drift event does trigger the full corrective sequence (wavelength scan and re-lock). Could the authors comment on this? Have they tested the system's response to induced, larger wavelength drifts?

  Answer: We thank the reviewer for this comment. The figure below illustrates a recorded event in which OH signal attenuation triggered the wavelength scan and subsequent re-lock process. It should be noted that this process is currently for display purposes only and was captured by an external standalone program, we can see that the x-axis represents time series rather than wavelength. For this reason, the data are not included in the main text. Nevertheless, this example helps demonstrate the capability of our reference system to effectively re-lock to the optimal OH excitation wavelength.

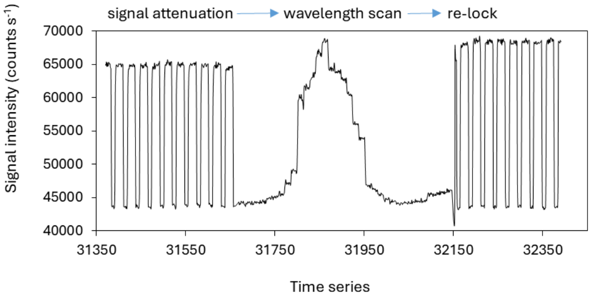

Figure s1. The full corrective sequence captured by an external standalone program.

- In lines 296-298, the suggestion to normalize ambient OH data against the concurrent reference cell signal during wavelength drift should be treated with caution. The physical conditions (e.g., temperature, pressure) inside the reference cell and the ambient sampling cell are different. These differences can lead to slight variations in the OH excitation spectrum. Consequently, the rate of signal attenuation due to laser wavelength drift may not be identical in both cells. Using the reference signal for direct normalization without carefully characterizing and correcting for this potential discrepancy could introduce additional error. Therefore, I recommend either removing this statement or significantly qualifying it with a discussion of the necessary corrections and associated uncertainties.

Answer: Thank you for this important comment. We agree that normalizing ambient OH signals using the reference cell signal requires careful consideration. In response, we have removed this statement from the manuscript to avoid potential misinterpretation.

- In Figure 4, there is a typographical error in the labeling of the OH excitation spectrum. The last characteristic peak within the red frame should be $P_1(1)$, not $P_1(3)$.

Answer: We have corrected the label in Figure 4.

- In Figure 4, an offset appears between the two OH excitation spectral lines, causing them to fail to correspond precisely. What might account for this phenomenon?

Answer: Thank you for pointing this out. Although the two OH excitation spectra show good agreement, a slight offset is observed. This may be attributed to different broadening effects due to slight differences in pressure and a more noticeable temperature difference between the reference cell and the OH detection cell. While both cells operate at approximately 0.4 kPa, minor pressure variations can occur depending on pump performance. More significantly, the reference cell uses a heated filament to generate OH radicals, resulting in a higher internal temperature than the detection cell, leading to different Doppler broadening effects. We therefore conclude that both temperature and pressure differences contribute to the slight spectral offset.

- In Section 3.2, the author did not specify the exact installation location of the THP sensor. Additionally, the filament's position and temperature were not considered in the author's experimental conditions. Would the filament's temperature rise during heating interfere with the humidity measurement results shown in Figure 8?

Answer: The THP sensor is installed on the inlet gas line upstream of the reference cell, providing real time measurements of the incoming gas temperature, relative humidity, and system pressure. Regarding the filament, its position and temperature are indeed critical for OH generation. The filament's mounting position is fixed mechanically. Experimentally, we align the laser beam to pass as close as possible to the filament surface in order to maximize fluorescence from generated OH radicals, however, the exact distance was not measured and is therefore not specified. While the THP sensor is located far from the filament and does not directly measure filament temperature, we could infer thermal behavior from background signal intensity. **Under stable operating conditions at fixed power, the filament temperature reaches thermal equilibrium rather than continuously rising. This is proved by the stability of the offline background signal in long-term tests, as shown in Figure 9, which indicates that thermal radiation, a major component of the background, remains steady.**

In the related humidity experiments, we observed a slight decrease in offline background signal (~1%) at RH above 60%, as shown in Figure s2, possibly because higher humidity slightly enhances heat removal by the pumped gas stream, and thus causes slight filament cooling. This cooling could also influence the OH production, although Figure 8 presents net OH signals increment (signal at online wavelength minus signal at offline wavelength), it is hard to demonstrate that subtle humidity-induced thermal interference is entirely ruled out.

However, in the RH range below 60%, humidity-induced variation in background signal, and by extension, in filament temperature, remains negligible. The observed change in the offline background signal across this range was only 0.154 kcounts s^{-1}, representing a relative variation of 0.23%. Over the same RH interval, the online signal increased by 13.279 kcounts s^{-1}, corresponding to a 7.0% rise in raw signal intensity. Even under the conservative assumption that the entire background decrease results from filament cooling, and that such cooling would proportionally affect the online signal, the corrected online signal increase would be 13.309 kcounts s^{-1}. The resulting increment remains 7.1%, which is statistically indistinguishable from the uncorrected value within typical measurement uncertainties.

In conclusion, within relative low RH regime, the influence of humidity on filament temperature, and consequently on the detected OH signal, is insignificant compared to the direct effect of water vapor concentration on OH production. However, at higher RH levels, multiple effects could slightly influence the OH signal. Water vapor not only contributes to high OH production but also enhances OH fluorescence quenching and may induce filament cooling. These likely contributes to the larger signal variability observed at elevated humidities. We will address this point in the revised manuscript to better support our recommendation that RH is better maintained within the 30–40% range for optimal system performance:

…. Although the nonlinear response suggests that the optimal OH signal occurs around 74% RH, such a high humidity

level would cause slight filament cooling and thus complicate signal interpretation and increase system instability. In addition, high RH levels would also accelerate the aging of electronic components and increase long-term operational risk. Therefore, in practice, RH is maintained within the 30%–40% range, which balances signal intensity, sustained system stability and performance.

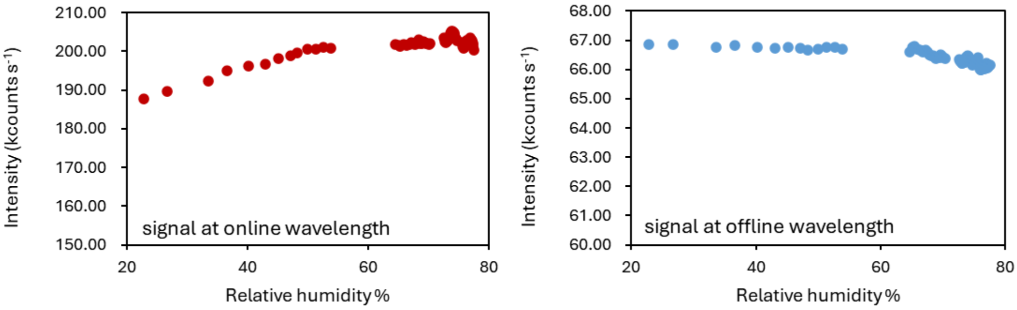

Figure s2. Measured raw signal intensity from the reference cell at online wavelength (left) and offline wavelength (right) as a function of relative humidity%.

Reviewer 2:

Comments and answer:

The manuscript by Chen et al. describes a systematically developed and characterized reference system designed for active wavelength locking in LIF-FAGE instrument. Given that laser wavelength stability is a critical prerequisite for accurate OH radical quantification, this work addresses a significant technical challenge in atmospheric spectroscopy. The authors provide a comprehensive characterization of the thermolysis OH source and demonstrate an impressive stability of 0.2% drift per hour over an extended period. The technical approach is sound, and the results are presented with high clarity. The identification of an "optimal operational window" provides a valuable framework for the community. I recommend publication after the following technical and linguistic points are addressed.

General comments:

1. The authors implement a wavelength locking program that triggers a re-scan once the net signal falls below 95% of the peak intensity. However, the 12-hour stability assessment reveals a systematic decay (1.1% per hour) in the final 3 hours that did not trigger the threshold. This implies that a 5% threshold may be too coarse for high-precision atmospheric measurements. The authors should discuss whether a dual-track strategy—combining amplitude-based triggers with periodic timed re-scans—would further enhance the system's reliability.

Answer: We thank the reviewer for the comment. Regarding our selection of the 5% threshold, the responses are as follows:

1） In comparison with similar systems, the latest AIOFM reference system reported by Wang et al. (2019) adopted a 10% threshold. Our 5% threshold is relatively more stringent, enabling correction at smaller wavelength deviations and thus providing higher measurement accuracy.

2） During long-term field observations, instrument performance inevitably undergoes gradual degradation, such as the alignment of the system, decreased dye efficiency, and declining transmission efficiency of the fiber, and so on. Under such conditions, an overly strict threshold could lead to frequent triggering of the wavelength scan procedure, introducing additional system instability and potentially exacerbating wavelength drift due to excessive mechanical adjustments.

Therefore, the 5% threshold represents a robust choice that maintains measurement accuracy while avoiding frequent corrections. And with the combination of the periodic timed re-scans, it is the optimal decision under current instrument conditions. But we fully agree that, in the future, with the improved instrument performance, stricter thresholds could be adopted to further enhance measurement precision. The above responses have been supplemented to the revised manuscript:

The reason we recommend periodic timed re-scans as a supplement to the 5% signal attenuation threshold is that the system is inevitably subject to various fluctuations during long-term operation. An overly strict threshold (e.g., 1–2%) would increase the risk of frequent triggering of the wavelength scan procedure, introducing undesired mechanical adjustments and potentially exacerbating wavelength drift. Therefore, the 5% threshold combined with scheduled periodic re-scans represents a robust strategy that maintains measurement accuracy while minimizing system instability. As instrument performance continues to improve in the future, stricter thresholds could be adopted to further enhance measurement precision.

2. In Section 1, only a brief comparison with the AIOFM reference system (282 nm excitation) is provided. It is suggested to add comparative data on key performance indicators (e.g., wavelength locking speed, signal stability, detection limit) with other similar reference systems using 308 nm excitation (if any), to more comprehensively highlight the advantages of the proposed system.

Answer: We thank the reviewer for this suggestion. For reference systems employing 308 nm excitation, data on key performance indicators remains scarce in the literature. The only available information we identified is from the AIOFM group (Wang et al., 2019), and this has been incorporated into the revised manuscript, along with the highlight advantages of the proposed system:

The AIOFM group has provided initial characterization data for their earlier version reference cell, however, a 282 nm laser was employed (Xing et al., 2017). In their upgraded system, a 308 nm excitation laser was used, and good stability was reported: a fluorescence fluctuation of approximately 1.2% over 15 minutes at the resonance wavelength (Wang et al., 2019). In this work, we present a comprehensively characterized real time reference system that features compact design, simplified operation, integrated wavelength locking program, and enhanced performance….

3. In Section 2.1.2, it is stated that dissociation into oxygen atoms is a non-negligible pathway for OH production. While the Harvard group's findings are cited, the manuscript would benefit from a more precise discussion of the filament temperature regime. Specifically, is the Fe/Cr/Al/Ni alloy surface temperature sufficient to drive molecular dissociation via thermolysis, or is the catalytic effect of iron the dominant driver?

Answer: We thank the reviewer for this comment. We fully agree that precisely determining whether OH generation in the reference cell is primarily temperature driven or catalytically driven is of great scientific interest. However, this topic requires further dedicated investigation, as catalytic decomposition reactions on metal surfaces are relatively complex. The OH generation mechanism described in the original manuscript, which attributed molecular dissociation to thermolysis, was based on established literature. Regarding catalytic effects, although some studies have reported catalytic activity of Pt surfaces for water dissociation, it would be insufficiently rigorous to directly infer similar catalytic behavior on the surface of our alloy filament without experimental evidence. Therefore, we have revised the manuscript accordingly, considering only thermal decomposition as the OH generation pathway and removing all statements regarding catalytic effects.

4. In Figure 6(a), the linear relationship of inlet flow rate and pressure is presented on a logarithmic scale, which is counterintuitive. It is suggested to supplement a flow rate-pressure relationship graph on the original scale to facilitate readers' intuitive understanding of pressure changes in different flow rate ranges.

Answer: The scale for the x-axis of Figure 6(a) has been adjusted to original scale.

5. Increasing the filament power will accelerate filament aging, but the manuscript does not provide data on the service life of the filament (e.g., continuous operation time under the optimal operating current, number of uses when the signal attenuates to the threshold, etc.). It is suggested to supplement filament lifespan test experiments to provide reference for the long-term maintenance and practical application of the system.

Answer: We thank the reviewer for this suggestion. We acknowledge that filament lifespan information was not provided in the original manuscript. Since the assembly of this reference cell system in June 2024, it has been tested and operated intermittently but stably over 15 months and remains in good working condition under a supply current of 2.5 A. According to the manufacturer's specifications, this heating wire has a service life exceeding three years when operated at a supply current of 2.5 A. This information has now been added into the revised manuscript as a reference for long-term system maintenance:

to balance long term stability with sufficient signal intensity, the filament supply current is optimally maintained within the range of 2.5 ± 0.2 A. *According to the specifications from the manufacturer, this heating wire has a service life exceeding three years when operated at a supply current of 2.5 A.*

6. The results in Figure 8 indicate an optimal OH signal at ~74% RH. Yet, for operational stability, the authors chose 30%–40% RH to prevent electronic aging. This is a pragmatic engineering decision but the authors should clarify if further hardware isolation (e.g., improved sealing of the reference cell channels) could allow for higher RH and thus a better Signal-to-Noise Ratio. Is the red line a polynomial fit of the measurement results? What does it mean?

Answer: Thank you for these comments, we agree that hardwire improvements may enable the reference cell to operate sustainably at higher RH levels, thereby achieving a higher signal to noise ratio. This is a valuable suggestion for future instrument optimization and will be considered in the next generation of our reference system design.

Regarding the red line in Figure 8, yes, it is a polynomial fit to the measurement results, intended to guide the eye and illustrate the overall trend of OH signal variation with relative humidity. The fit is not based on a theoretical model but serves to highlight the nonlinear response observed experimentally, including the plateau and subsequent decline at higher humidity levels.

7. The reasons for the fluctuations in the signal data in Figure 9 are not detailed. It is recommended to add possible factors for short-term fluctuations (e.g., laser pulse stability, minor environmental disturbances) in the figure caption.

Answer: Thank you for this suggestion, we have revised the caption of Figure 9 to include possible contributing factors. The updated contents are:

Figure 9. ... on September 22. The short-term fluctuations observed in the signals are primarily attributed to the instability of the laser.

Minor comments:

1.The term "air-conditioned box" is somewhat colloquial. I suggest replacing it with a more technical term such as "thermostated enclosure" or "temperature-controlled housing".

Answer: We thank the reviewer for this comment. The enclosure helps to control both the temperature and RH of the laser and reference system, thus the term "air-conditioned box" was initially used. It has been replaced with "air-conditioned enclosure" to both accurately describe its function and employ a more technical term.

2. Throughout the text, the word "Therefore" is used with high frequency (e.g., Line 175-176). Utilizing alternatives such as "Accordingly" or "Hence" would improve the flow of the manuscript.

Answer: We thank the reviewer for helping us improving the manuscript. The word "therefore" has been used 6 times throughout the manuscript, some of them has been replaced with alternative words, all of them are highlighted in the revised manuscript.

3. Please ensure consistent use of units for laser power normalization. Both "kcounts s^{-1} mW^{-1}" and "counts s^{-1} mW^{-1}" are present.

Answer: We thank the reviewer for pointing this out, the unit "counts s^{-1} mW^{-1}" appears three times in the main text, all within Section 3.3. In the revised manuscript, they have been modified to "kcounts s^{-1} mW^{-1}" to ensure units consistency throughout the manuscript.

Reference

Wang, F., et al., Development of a field system for measurement of tropospheric OH radical using laser-induced fluorescence technique. Optics Express, 2019. 27(8): p. A419-A435.

Xing, X.B., et al., Study of a laser wavelength correction method applied to the measurement of OH Radical with laser-induced fluorescence. Guang Pu Xue Yu Guang Pu Fen Xi, 2017. 37(3): p. 692-6.

Fridell, E., et al., A laser-induced fluorescence study of OH desorption from Pt in H_2O/O_2 and H_2O/H_2 Mixtures. Langmuir, 1994. 10(3): p. 699-708.