

We thank the reviewer 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 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, 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 our 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 literature has reported catalytic activity of Pt surfaces for water dissociation (Fridell et al., 1994), 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 hardware 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. This information has been added to the caption of Figure 8.

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 " $\text{kcounts s}^{-1} \text{mW}^{-1}$ " and " $\text{counts s}^{-1} \text{mW}^{-1}$ " are present.

Answer: We thank the reviewer for pointing this out, the unit " $\text{counts s}^{-1} \text{mW}^{-1}$ " appears three times in the main text, all within Section 3.3. In the revised manuscript, they have been modified to " $\text{kcounts s}^{-1} \text{mW}^{-1}$ " 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., 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 $\text{H}_2\text{O}/\text{O}_2$ and $\text{H}_2\text{O}/\text{H}_2$ Mixtures. *Langmuir*, 1994. 10(3): p. 699-708.