We would like to thank the referee for the review and comments to improve the quality of this work. Following these comments, we have made major modifications and point-by-point responses. Please find our response below (in blue).

The manuscript "Portable system for characterizing greenhouse gas analyzers for sensitivity to water vapor" by Liu et al. claims to describe a "lightweight, low-cost, and portable system" for performing water vapor corrections of greenhouse gas analyzers.

The basic premise of their system is forming a "wet" stream of calibration gas using a Nafion-based humidity exchanger, then combining it with a "dry" stream at set ratios using mass flow controllers and valves. This is a description of common laboratory practice that has been used for decades in air pollution / air quality and VOC laboratories and in instrument development and testing, in commercially available instruments (e.g. 2B Technologies DewLine), and presumably also used already in the greenhouse gas community.

We thank the reviewer for this comment highlighting the need to clarify our contribution and we changed Lines 70-77 in the introduction to read "Using Nafions to condition air samples (either stabilizing the humidity, drying, or equilibrating to a reference humidity) is certainly not new and has been used in many fields. For example, the 2B Technologies DewLine (https://2btech.io/dewline/) uses a Nafion tube to equilibrate humidity of the sample and reference gas streams. In our instruments deployed in the field (e.g. Richardson et al., 2017; Liu et al., 2025), we also use a Nafion to equilibrate the humidity (in our case to low levels, around 1000 ppm H_2O). Because of the extreme compatibility goals for greenhouse gases, this procedure of drying the air sample still leaves some humidity effects on the final results. In this paper, we introduce a new portable and automated system that enables water-vapor calibrations to be performed directly in the field. We use this system to develop calibration curves to adjust for those relatively small effects in order to improve overall compatibility."

The concept of mixing humidified (using a Nafion tube submersed in water) and dry split flows using mass flow controllers has also been documented in the greenhouse gas community literature (see Lines 61-63 in the introduction of the paper). Another example is Dembowski et al. (2021), which documented the use of this concept to investigate the role of relative humidity in determining which aluminum mineral forms under various conditions. They use feedback from a humidity sensor to ascertain the humidity levels remain constant – this would be a potentially valuable improvement for future iterations of our system. We added this point to the discussion (Lines 306-310).

References

Dembowski, M., Loring, J. S., Bowden, M. E., Reynolds, J. G., Graham, T. R., Rosso,

K. M., and Pearce, C. I.: The controlling role of atmosphere in dawsonite versus gibbsite precipitation from tetrahedral aluminate species, Dalton Trans., 50, 13438–13446, https://doi.org/10.1039/d1dt02081a, 2021.

The claim to novelty, then, comes from the weight, cost, portability, and apparent stability of their system. They provide dimensions (30.5 cm x 30.5 cm x 20.3 cm, 5 kg) but it is not clear whether that includes all parts of the system (including the water and a battery) or not. They claim "low power requirements" but do not specifically test or report the power draw or make any claims for how long a typical battery-operated system could run in the field. They also provide no head-to-head comparisons of their system to existing, commercially available systems to provide an equivalent comparison.

We agree that using mass flow controllers and a permeation-based water exchanger is not novel and have clarified Lines 70-71 in the manuscript. As noted by the reviewer, the novelty of the system is the weight, cost, portability, and stability. We have also clarified specifics as noted below.

We have revised the manuscript to clarify what is included in the reported 30.5 × 30.5 × 20.3 cm dimensions and 5 kg system mass, and to distinguish the present configuration from an optional battery-powered configuration (Line 100).

We thank the reviewer for this comment and agree that the original manuscript did not provide sufficient quantitative information regarding power consumption and battery operation. We have now measured the steady-state power consumption of P-WAVES during normal operation and found it to be approximately 10 W. Thus, a 24-h field deployment would require approximately 240 Wh of delivered energy. As a representative example, a commercially available 12-V, 30-Ah LiFePO4 battery provides sufficient capacity for more than 24 h of operation while allowing reasonable reserve for conversion losses. Such a battery is approximately 16.5 × 16.3 × 8.6 cm, weighs approximately 3.2 kg, and costs less than US$300. It could readily be accommodated within the enclosure currently used for P-WAVES. We added these important details to Lines 107-112 in the manuscript.

We also agree that the original discussion could be interpreted as implying a direct performance advantage over commercially available humidity-generation systems. A head-to-head experimental comparison was not part of this study, and we have revised the text accordingly. Our intended novelty claim is not that P-WAVES outperforms commercial humidity generators in absolute metrological performance in terms of accuracy, but rather that it provides a small, low-cost, low-power, automated system specifically designed for portable field characterization of greenhouse-gas analyzers

across controlled stable H_2O levels. Commercial systems may provide substantially greater environmental control, traceability, and absolute humidity accuracy, but generally address a different use case. We have therefore moderated the wording of the novelty claim and clarified that the principal advantages of P-WAVES are its portability, low cost, low power requirement, automation, and suitability for analyzer-specific field calibration (at Lines 75-77). RH sensors are accurate enough to hold H_2O to ~500 ppm and chilled mirror sensor to within approximately 20-50 ppm H_2O, so the stability of the P-WAVES system (2-3 ppm below 10000 ppm and 7-10 ppm up to saturation) is an improvement. While the P-WAVES system is very stable, its accuracy is untested, beyond using the Picarro and Aeris water vapor output which is itself not calibrated.

The reviewer is correct to point out that we do not have any head-to-head comparisons of our system to an existing humidity measurement (beyond that of the instruments to be calibrated). Using a sensor, such as a chilled mirror hygrometer inline would be a potentially valuable addition, particularly if the system is to be used for other applications. We added this point to the discussion (Lines 292-294). However, for our application, the key point is the value of water reported by the instrument, since, in practice, that is what we will be using in the field to correct for the remaining water vapor after using a Nafion dryer to remove most of the water. The actual value of the water and its calibration is not crucial (unless the calibration of the water reported by the instrument changes dramatically in the time between calibrations or during individual steps of the calibration).

The authors make a claim of "high stability" (3 ppmv H_2O). However, in the manuscript they appear to only take the average of the last 60 seconds of a 3.5 minute step. It would be much more convincing if they reported a response time or equilibration time to humidity change and provided the long-term (e.g. 8 hour, 24 hour, etc.) stability of the system instead of 60 second transients. A 60-second transient does not show "high stability". Depending upon the type of tubing used (polymer, stainless steel) and specific instruments used, the system response time to water vapor could be on the order of 60 seconds or longer. The apparent high stability could be a system-induced broadening kernel. The authors do not state what materials, tube diameters and lengths, etc. they used, nor do they provide any residence time distribution data to prove that their 60-second stability is actually the stability of their humidity generation and not a system artifact.

We agree that a 60-s interval characterizes short-term within-step stability rather than long-term stability, and have revised the terminology accordingly. For our application, within-step stability is more important but other applications may need long-term stability so we additionally performed a 16-h fixed-setpoint experiment to characterize

the P-WAVES system for applications in which long-term stability is required. Slow changes in generated H_2O were strongly correlated with the Aeris enclosure-inlet temperature (T2; R^2=0.99 for 5-min averages after the first hour) and were consistent with the expected temperature dependence of water-vapor saturation (see figure below, now Figure SX in the supplement). During the thermally stable final 8 h (by coincidence) of the 16 h test, H_2O drift was approximately -5 ppm h^{-1}, while median 60-s variability was approximately 5 ppm (1σ). We added the plots in the supplement (Figure S6) and the text describing this test in the discussion (Lines 326-334). We have therefore revised the manuscript to distinguish short-term within-step stability from long-term absolute stability and to note the importance of temperature measurement or control for extended fixed-humidity operation.

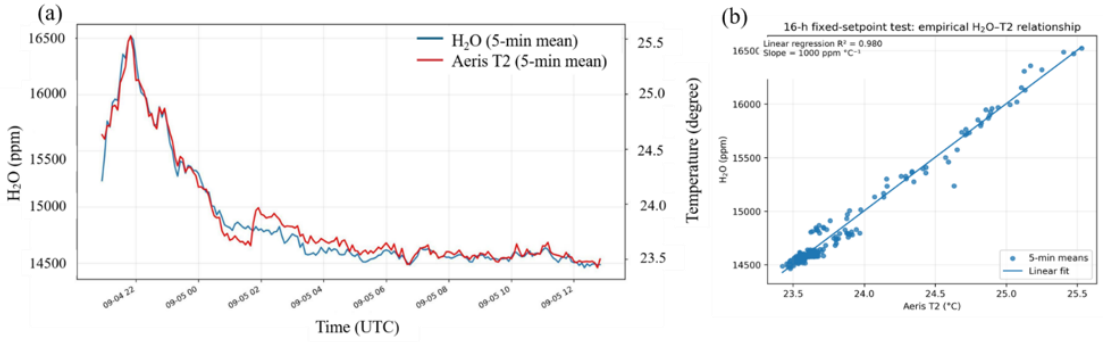


Figure S6. (a) Water vapor as a function of time during a 16-h fixed setpoint test. Enclosure inlet temperature (T2), used as a proxy for water bath temperature, is shown on the right-hand y-axis. (b) Water vapor as a function of enclosure inlet temperature for the test shown in (a).

The system uses 1/8'' OD stainless steel tubing (length is estimated to be about 50 cm of stainless tubing internal to the calibration unit) and we apologize for the oversight in not including this important detail. We do not in general consider longer time steps because our primary goal with the water calibrator system is to calibrate the Aeris sensors for water vapor and these instruments drift dramatically in time (Liu et al., 2025), particularly for ethane. Our aim is to keep the time steps short in order to isolate water vapor effects from instrument drift.

Additionally, we added a figure (below) to the supplement (Figure S1) to more carefully address the water vapor variability at various levels during the steps used to develop water calibration curves for the instruments. The variability increases with increasing water vapor for the completed tests. Across 29 humidity-step tests comprising 368 individual steps for two Aeris analyzers, the median final-60-s H_2O variability was 2.8 ppm (1σ), with typical variability of ~2-3 ppm below 20,000 ppm H_2O and ~7-10 ppm at the highest H_2O levels.

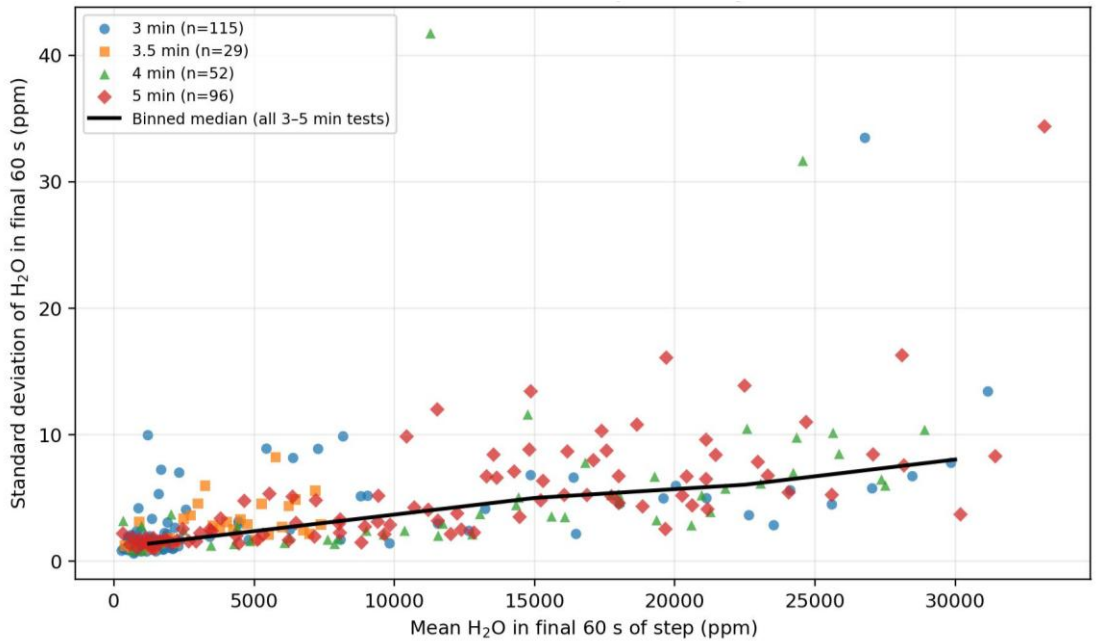
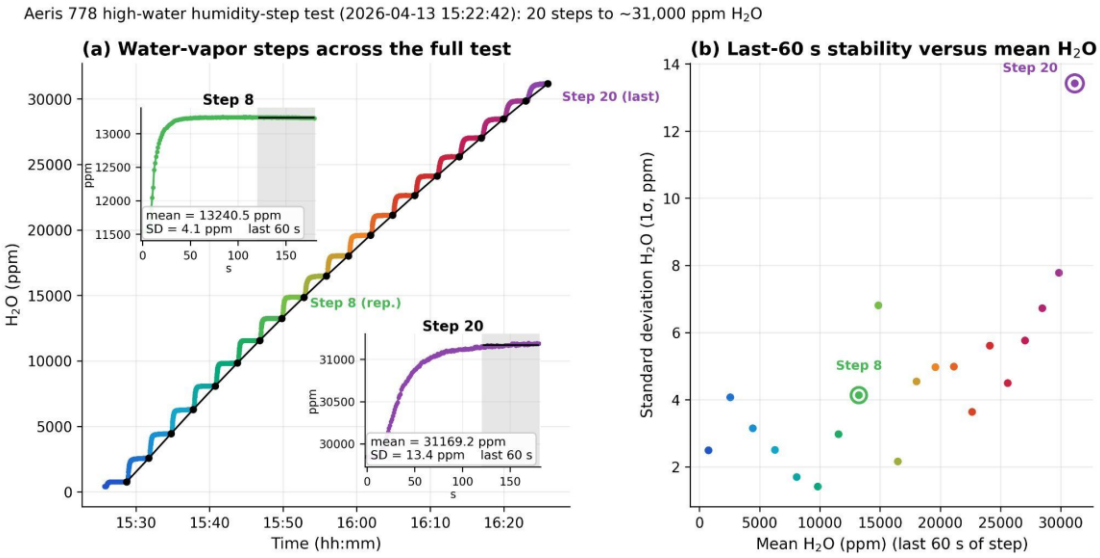


Figure S1. Variability of generated H_2O as a function of water vapor concentration and step duration. Each symbol represents the standard deviation of H_2O during the final 60s of an individual concentration step. Different symbols indicate step durations of 3, 3.5, 4, and 5 min. The black line represents the binned median across all 3-5 min tests.

Across the 3-5 min humidity-step tests, final-60-s H_2O variability showed no clear systematic dependence on step duration.

Based on this reviewer's concern regarding the ability to see details within the steps (also noted by reviewer RC2), we added insets to Fig.2 in order to better visualize the stability of the steps. The stability is quantified by the reported standard deviation.



The authors claim "stable and repeatable humidity control…" but provide no statistics

or demonstration of long-term stability and either short-term or long-term repeatability. As far as I can tell, their experiments were purely done in a climate-controlled laboratory environment. This performance might not transfer to field environments. In particular, the stability of the humidity generation will be highly sensitive to the temperature of the system, so ambient temperature fluctuations will adversely impact the stability of the system. It is not clear that the authors have thoroughly evaluated and addressed these impacts.

Indeed the tests were all done in a climate-controlled laboratory and the temperature of the laboratory (more specifically the water bath) dictates the saturation water pressure (i.e., the maximum water vapor attainable). As also noted by the reviewer, changes in the water bath temperature throughout each step would affect the stability of the generated water vapor level. An improvement to the system could be to measure the water bath (or use a chilled mirror sensor to measure the water vapor) and potentially automatically adjust the mass flow controllers to compensate. But without these complications, the laboratory stability of the water vapor within the steps is 4.8 ± 2.9 ppm H_2O, sufficient for our laboratory characterizations purposes. (add to discussion Lines 310-312). Vacuum-sealed enclosure should be used for field tests.

Additionally, we have now examined whether temperature variability could account for the observed short-term H_2O variability during the calibration steps. We compared the standard deviation of the H_2O during the calibration steps to the standard deviation of the enclosure inlet temperature (T2, see below). The correlation coefficient (R^2) was 0.006 for the 313 windows from the 3-5 min tests. Thus, variation in enclosure inlet temperature explains essentially none of the observed short-term H_2O variability during the final minute of humidity steps. The short-term H_2O variability may be caused by Aeris H_2O measurement noise, MFC flow-control variability, and/or analog setpoint-voltage stability. We added this figure to the supplement (Figure S4) and the text to Lines 314-318 in the discussion.

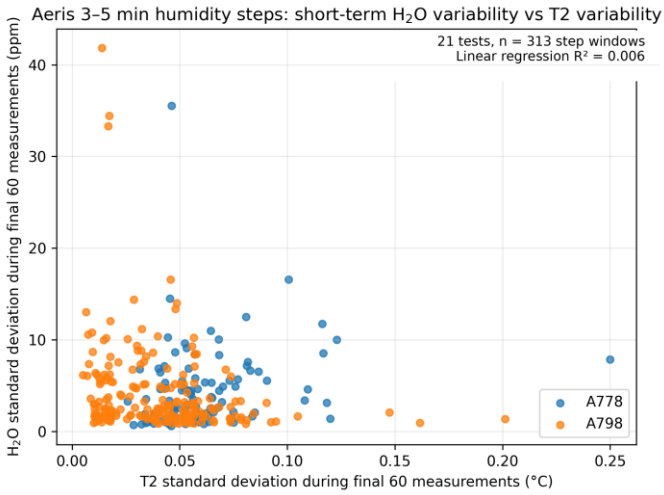


Figure S4. Standard deviation of the H_2O during the calibration steps to the standard deviation of the enclosure inlet temperature (T_2) of 21 tests (313 windows) for Aeris MIRA analyzers (serial number A778 and A798). The correction coefficient (R^2) was 0.006 for these 3-5 min tests.

For our application, the key is the reproducibility of the calibration curves for water vapor. Figure 3 and 5-6 show CO_2 and CH_4 repeated tests for the Aeris and Picarro systems. Table 1 and 2 document the related statistics and its repeatability is discussed in Lines 319-325 in the manuscript.

The authors do not use any independent humidity or dewpoint sensors (such as a chilled mirror or capacitive sensor) to confirm the accuracy or stability of their humidity generation. Instead, they rely on instrument-reported water vapor lines. However, they provide no calibration or traceability for these measurements. As a humidification system or humidity generator it has no inherent humidity measurement capability.

As noted above, the actual value of the water vapor is not crucial for our particular application. The reviewer is correct that humidity measurement capability should be added in order to use this system more generally and we have noted this in the discussion (Lines 290-294).

As mentioned earlier, much of the system (split flows, Nafion humidification, mass flow controllers to adjust humidity, etc.) are already common practice in the field. If the authors want to make a novel claim about the superiority of their system they need to provide additional details about the other factors (mechanical designs, CAD files, Python code, parts lists, etc.) to make it reproducible and independently testable. If the authors can address some of these concerns and provide more convincing data about the value of their system I think it would be a useful contribution to the field, especially if the technology is made available (public domain or commercially).

Thank you for the suggestion. We have made the Python control software used to operate P-WAVES publicly available on GitHub (https://github.com/srichard13/P-WAVES). The repository includes the production controller used for the Aeris MIRA Ultra analyzers and a representative Picarro controller (CFKADS2025). Separate Picarro versions were used for individual instruments because of instrument-specific communication/configuration differences. The archived source files are provided unchanged from the versions used for the experiments, together with documentation of the hardware assumptions and logged variables.

I do not address the Aeris and Picarro testing here, as those aspects are already well-covered in the scientific literature and well known to be dependent upon instrument, firmware / software version, etc. and not necessarily transferrable (although the calibration process and applied correction models could be).

We believe our system to be a significant improvement on the systems currently documented in the greenhouse gas literature (Lines 75-80 in the introduction) primarily because of the size, stability, and automation.