Review of the revised version of 'The "Golden Points" and nonequilibrium correction of high-accuracy frost point hygrometers' by Poltera et al.

Reply to the editor

Dear Marc von Hobe

Thank you for evaluating our manuscript and giving us the opportunity to revise it further. We have completed a second round of peer review and have taken all comments into account. Most of them help to further improve the quality of our manuscript.

However, we do not agree with the second reviewer, who states that the text could be shortened by 50%. This article addresses a complex topic and draws conclusions that could have far-reaching implications. We understand that Rev. 2, who appears to be a highly experienced expert, would have written this article shorter, though certainly not by 50%. The reviewer also believes that the article has become longer rather than shorter, which is objectively incorrect. Please note our response to the reviewer.

At the same time, the reviewer asked us to include additional material, which we did in the previous round. Now he is again requesting that we add a new section to evaluate the quality of the entire NDACC dataset, which clearly goes beyond the scope of this article. Such an undertaking would likely merit a separate publication, would significantly lengthen the present manuscript, and should best be carried out by NDACC scientists.

Therefore, we can only reject these two comments. However, we have addressed all other points and hope for a positive decision.

With best regards,

Beiping Luo for all authors

Report Editor - Marc von Hobe, 04 Mar 2026

Dear Yann Poltera and coauthors,

Both original referees have now reviewed the revised version. While one of them is mostly content with the revisions made and only recommends technical corrections, the other referee still recommends extensive revisions. As far as I can see, the bottom line is that the true added value of the new methodology and of this manuscript still does not become fully visible, because it is to some extent hidden between the many details in this very long paper.

Please read the review carefully and see how you can make the added value clearer, following the recommendations of the reviewer.

Best regards,

Marc

Review of the revised version of
‘The “Golden Points” and nonequilibrium correction of high-accuracy frost point hygrometers’ by Poltera et al.

Author response to reviewer comments

We would like to thank the two anonymous reviewers for their constructive and critical evaluation of the revised manuscript. Some of their comments were challenging, but they also helped to improve the quality of our manuscript. The reviewers' comments are listed below in black font, and our responses in italicized blue font.

Report #1 - Anonymous Referee #1, 24 Dec 2025

Thank you for your responses to my comments and suggestions. All of them look very reasonable to me. I spotted one thing: The reformulated text in response to my comment to the original Line 434 has duplicated "also". Other than this, I think that the manuscript can be published (once the very interesting and informative discussions with the other reviewer have been settled, of course.)

Thank you, we appreciate this comment. We have removed the duplicated ‘also’ in line 323 of the revised manuscript (line 434 in original manuscript).

Review of the revised version of
‘The “Golden Points” and nonequilibrium correction of high-accuracy frost point hygrometers’ by Poltera et al.

Author response to reviewer comments

We would like to thank the two anonymous reviewers for their constructive and critical evaluation of the revised manuscript. Some of their comments were challenging, but they also helped to improve the quality of our manuscript. The reviewers' comments are listed below in black font, and our responses in italicized blue font.

Report #2 - Anonymous Referee #2, 2 Mar 2026

Review of the revised version of “The “Golden Points” and nonequilibrium correction of high-accuracy frost point hygrometers” by Polterra et al.

The authors clearly have a good point that uses the mirror reflectance in a correction of frostpoint measurements, for which the conventional uncertainty estimation is not appropriate. This is an important result that may eventually be implemented in the processing of frostpoint soundings. The manuscript should therefore eventually be published.

Thank you, we appreciate this comment.

However, the prolix style and still frequent repetitions make it difficult to follow exactly where the value of this manuscript lies. To some extent the loquacity hides where this method adds value and where it does not. I would therefore recommend to give the authors time for major revisions before publishing this manuscript.

We have made every effort to avoid unnecessary repetitions. We believe the examples and case studies are needed to prove that the method works under various conditions. We further believe the level of detail is valuable information for the community. We have made efforts to emphasize the key points, and highlight where this method adds value and where it does not.

“Golden Points”

There are still fundamental aspects that the authors do not yet describe appropriately. My initial review focused on the issue of looking at the “golden points” as points yet requiring smoothing to properly identify them. The problem I see is the interpretation of “equilibrium.” Their definition of “golden points” is dU/dt = 0. I agree that averaging around these points is needed, but the question is why. The authors attribute this to electronic noise, even though the purely electronic noise determined prior to cryogenic cooling is a factor of ten smaller as they describe.

We do not attribute the noise solely to the electronics, but state quite clearly: “white noise from the electronics and from the condensate itself (such as specular reflections of individual crystal faces), but above all spectral noise from problematic and unwanted signals that correspond to some eigenfrequency of the condensate-substrate-electronics system” (see Section 3.1 “Definition of Golden Points”). In fact, we had clarified this based on the reviewer’s objections during the first review round and added Figures D1 and D2, which make this unmistakably clear. (As a side effect, this more comprehensive treatment of system noise and the associated smoothing requirements, based on the reviewer’s comments, means that the manuscript does not appear to have been shortened as much as is actually the case.)

In my opinion, the challenge is that the “equilibrium” is only a partial equilibrium in the mirror reflectance, but not an equilibrium of the mirror temperature. In fact, as pointed out the rate of change of the mirror temperature is near a maximum at the “golden points.” Under this condition, there is a question of what the temperature measurement truly represents. The mirror temperature is measured at the location of the thermistor. Since its rate of change is at a maximum, it cannot be assumed that the temperature across the entire mirror disk and that of the exposed frost layer is in equilibrium. There are likely temperature gradients across the mirror, despite fast heat transport. This implies that the equilibrium of the frost layer with the gas phase above is not a local equilibrium at any point on the mirror. Rather, the optical sensor averages out the local close-to-equilibrium signals over the entire mirror surface. This is the most likely source of noise in the mirror reflectance and the justification of why further averaging around the “golden point” is required. The question now is what the temperature at the location of the thermistor represents. It is a snapshot at one location and not an average across the mirror disk during a period when the mirror temperature is changing rapidly. Thus this is a fundamental limitation in which the optical detector intrinsically averages across the mirror disk and the temperature sensor does not. Thus, the timing mismatch between mirror temperature and mirror reflectance is not due to the 1 Hz resolution of the data system, but rather a fundamental limitation due to the different methods of how both are measured especially when one is changing rapidly. Increasing the data resolution would not solve this issue.

Good point! The uncertainty regarding the position of the thermistor is reported to be < 0.1 K for FPH and CFH (Oltmans 1985; Vömel et al., 2007; Hall et al., 2016) and is considered the major contribution to the mirror temperature uncertainty which is not related to frost control.

Using the thermal diffusivity of copper of 115 mm^2 / s and the mirror radius (3.5 mm) as representative length scale, we get an e-folding time of 3.5 ^2 / 115 ≈ 0.11 s for redistributing heat evenly across the mirror.

Using the thermal diffusivity of ice of 1 mm^2 / s and the ice thickness (< 0.1 mm) as representative length scale, we get an e-folding time < 0.1 ^2 / 1 ≈ 10 ms for heat transport through the thin ice layer.

Therefore, the temperatures of the mirror surface and the ice film thus equilibrate within ≈ 0.12 s when the total heat load changes. In the original manuscript we had only taken the second of these two thermal diffusion processes into account. We have now corrected this in Section 4.1.3. Since 0.12 s × dTm/dt is the time-lag error, dTm/dt > 1.8 K/s would be required to increase the uncertainty of the Golden Points to more than 0.2 K. . Such conditions hardly ever occur. Even in the most extreme case of the Hunga Tunga eruption (Section 5.6.2), we find dTm/dt < 1 K/s.

In other words: Your argument is correct, and we thank you for bringing it to our attention. However, we believe that it is not relevant to CFH in practice and this does not alter either the interpretation or the quantitative analysis in the paper.

Ideally, an instrument should reduce temperature inhomogeneities on the mirror as much as possible and account for any timing mismatch between Um and Tm. We agree that the inherent timing mismatch cannot be eliminated by higher-frequency measurements, and we have never claimed that this would be possible. However, if the timing mismatch is quantitatively less than the present 1-s intervals of common instruments, then higher frequencies could indeed reduce the temporal uncertainty at the Golden Points. It is up to the instrument developers to enable such conditions.

The authors recommend an improved system with a control generating more “golden points” and a slow rate of change of mirror temperature at the golden points. I believe these are conflicting requirements and that the observations of the authors with Skydew are inherent. While this may work under stationary conditions, a fast response instruments must be able to achieve fast temperature changes. Controllers with better implementation than currently done in the CFH may be able to do so. However, it is questionable whether they would lead to more “golden points” and whether more of these is indeed desirable.

The generation of more Golden Points together with slow changes in mirror temperature are not necessarily conflicting requirements. There are measurement scenarios, these conditions are compatible with one another, as demonstrated by the examples in Fig. 7A for Lindenberg and 7B for Nainital for CFH. Here, the amplitudes of the controller oscillations remain below 0.2 K with Golden Points every ~150 m. In the revised version of the article, we now explicitly refer to these two examples. As for SKYDEW and its aggressive controller, with amplitudes of the temperature oscillations typically above 1 K and a Golden point every ~15 m, it would be desirable to reduce the amplitude, which should be possible without running the risk of no longer keeping the true frost point within the oscillation interval.

I agree with the authors how they implement their “golden points” but would suggest writing the equation as dU_bar/dt = 0, where U_bar is the averaged mirror reflectance. This would describe what is really done.

If we were to write U_bar, we would probably also have to write T_bar. We appreciate the suggestion and, based on the reviewer’s earlier comments, have added extensive text and figures on the topic of averaging, smoothing, and filtering. However, the focus of this article is to draw attention to the use of U, which has so far been largely neglected or viewed merely as a simple household variable. Ideally, in the absence of noise, no filtering of U would be necessary. We apply a shape-preserving noise filter to U. U_bar might give the impression that we are applying long boxcar averages to U, which would not preserve the shape of the underlying U time series and does not correspond to our approach. For these reasons, we refrain from writing U_bar.

Brewer and Dobson as foundation of the “golden points”

I still disagree with the overinterpretation of the work by Brewer and Dobson as the foundation of the “golden points” (discussed in section 1 and again in section 3). I reread the Bakerian lecture Dobson gave in 1945, and it does not support the claim of the authors. Dobson emphasized that “the only necessary condition being that the amount [of deposit] should not be changing” and he may be the first to explicitly state this. This is in contrast to the original dew point hygrometers going back all the way to John Frederic Daniell, where the onset and disappearance of deposit was taken as the dew point. To complete the quote by Dobson from his Bakerian lecture: “… two temperatures are found at which the deposit is seen just to increase and just to decrease. The mean of these temperatures will be very near the true frost-point." They realized the difficulties of observing a constant deposit and resorted to averaging two manually read temperatures. They did not suggest using one well-timed reading at constant deposit.

We have a different interpretation of Dobson's work, in line with the definition of the Golden Points. From the paragraph describing the visual observation method ("two temperatures are found at which the deposit is seen just to increase and just to decrease. The mean of these temperatures will be very near the true frost-point"), we understand that Dobson et al. measured the temperature just before and just after the net-zero-flux condition was met (i.e. a Golden Point), arising as the frost condensate transitioned from growth to sublimation. They had to rely on two readings because for a human observer this is a more robust way to estimate the temperature at the net-zero-flux state (i.e., record T as long as condensate is still seen to increase, and record the last T before the condensate starts to decrease; then take the average / linear interpolation of the last two T-readings as the frost point estimate). At their time, this was the best possible approach to a Golden Point.

Regarding the automatization of visual readings, Dobson et al. went on to write: "As with most instruments, increase in elaboration brings with it disadvantages as well as advantages, and one cannot be so certain that errors may not have crept into an observation which would have been noticed at once if the measurement had been made by a good observer using the eye-reading instrument". This almost resembles a prophecy regarding the neglect of the Golden Points and nonequilibrium states of the mirror following the advent of automatic instruments.

Thornthwaite (1940) writes at the end of the description of his dewpoint hygrometer: "In fact, by means of radio tubes it should be possible to build a circuit such that the current to the heater varies inversely as the reflective quality of the mirror. Then as moisture begins to appear on the mirror the heat applied to it will increase and the temperature of the mirror will be maintained continuously at the dew point." This effectively describes a proportional controller, which has been used by many groups since. His use of "maintained" may already imply constant condensate. Nevertheless, Dobson spells it out exactly.

The mirror condensate in their device evaporated during its cycles around the dewpoint, and the recordings occurred "just at the incipient condensation of water". We agree that the two sentences of outlook in Thornthwaite and Owen (1940) could be interpreted as the idea of a P-controller maintaining constant dew on the mirror, although this is not explicitly stated. To our knowledge, Dobson and colleagues were the first to clearly articulate the idea of constant frost on the mirror for upper-air measurements.

Derivation of the “golden points” non-equilibrium correction

It took a little while to understand that equation 5 allows the correction of the mixing ratio and equation 8 the correction of the mirror temperature. Both correction methods are basically equivalent. What is the benefit of going through the derivation of equation 8?

In the atmosphere, the frost point decreases more or less linearly with altitude, whereas the mixing ratio varies over several orders of magnitude. Also, frost point is the primary measured quantity of chilled mirror hygrometers and the nonequilibrium equation in frost point space (Eq. 8) relates closely to frost control considerations.

For applications in the stratosphere, the mixing ratio formulation (i.e., dividing Eq. 5 by pressure) might be more appropriate, or in case of large (>1-2 K) nonequilibrium errors, where the linearization error in frost point space becomes significant.

The correction using eqn. 5 is used in more examples than eqn. 8. When should the user choose one or the other and why? Why not use eqn. 5 in all cases?

Typically, Eq. 8 is sufficient, but for maximum applicability, Eq. 5 should be used.

Importance of the “golden points”

In the review I struggled with understanding the extent to which the work of the authors improves the frostpoint observations. In many places the manuscript reads as if the errors of the CFH are large and that the “golden points” method are a required approach to improving the measurements. Figure 12 provides a distribution for the reduction of error using the “golden point” method in 70 soundings. This Figure is a little hard to interpret. I would read it as if in 99% of all data points the error reduction is minimal, and in the other 1% the error reduction can be as large as 100%. The text explaining that Figure is not consistent with my reading and states for example “18% of all data points have errors larger than 0.2K”. Be that as it may, the main result would be that more than 80% (or 99% in my reading) of all data points already have errors as small as what is assumed for the “golden points”. This means that this method can help in some cases, but the majority of all data points is not improved. This should be better explained to give the reader a better context. Another example is the end of the abstract, which states that “… the nonequilibrium correction removes over 80 % of the nonequilibrium error …”. The abstract does not mention here that this applies only to somewhere between 1% and 18% of all data and that there is no significant error reduction in the other > 80% of the data. There are more places like this, which I do not highlight here, and I leave it to the authors to check for appropriate representation of the results.

This is a valid point of criticism. To avoid the false impressions of an excessively high percentage of measurement points with serious errors, we complement the abstract as follows: "Based on 70 nighttime CFH-RS41 tandem flights, the nonequilibrium analysis suggests that 18 % of the about 23'000 measurement points in the UT/LS are associated with errors in the frost point of more than 0.2 K, 10 % with > 0.3 K or more, 4 % with > 0.5 K, and 1 % with > 1 K, but most profiles (68 %) have at least 5 % of their measurements with errors > 0.2 K ."

To improve the comprehensibility of Fig. 12B, we add the sentence "Each bar in the histogram represents the occurrence probability in a 0.1-K interval" to the figure caption. We had assumed – obviously incorrectly – that this was self-evident.

In several places, the authors point out that the data points that can be improved are those where the mirror reflectance changes fast. Can they show which data points are most likely to be suspect, i.e., can they give a threshold for the rate of change of the mirror reflectance above which the "golden point" method would provide significant benefit? This would be most helpful in interpreting the existing data.

Specifying a threshold is difficult, because the mirror sensitivity (A') is not the same for all instruments (nor for all CFHs). Recalling Eq.8:

$$T_m - T_f = B/A \; dU_m/dt \;\rightarrow\; dU_m/dt = A * (T_m - T_f) / B = A * (T_m - T_f) * T_{air}/T_m^2 * X_{H2O}$$

Nevertheless, assuming a 'low sensitivity case' scenario (low ice film sensitivity A = 1000 V/s), a rough threshold may be $|dU_m/dt| > 0.06$ *mV/s for detecting* $|T_m - T_f| > 2$ *K errors in the lower stratosphere (where* $B = T_m^2/(T_{air} X_{H2O}) \sim 30\,000\,000$ *K). But again, there is no universal value.*

Another major drawback of the "golden points" method seems to be that it requires a second instrument. This would severely limit its applicability to correct existing data where no such instrument is available, which is the vast majority of all existing soundings. This would also need to be clearly described including in the abstract.

Sorry, but for large nonequilibrium errors of CFH using a reliable second instrument (which may have worse accuracy but should have a high vertical resolution) is an optimal possibility to correct the nonequilibrium error between the Golden Points. This is not "a major drawback" but often a necessary requirement.

In addition, we clearly describe two other possibilities: (i) using ascent and descent data in the lower stratosphere might be useful to obtain the morphological sensitivity of the second ice film without the use of an additional sensor (see Sections 5.3 and 5.6.1); (ii) *instruments with a high frequency of Golden Points may be self-corrected (see Section 5.7).*

Following the reviewer's recommendation we have added the following information to the abstract: "Alternatively, under suitable circumstances the morphological sensitivity can also be derived from matching ascent and descent data."

Uncertainties

The uncertainty analysis for the "golden points" involves a statistical analysis for all 70 profiles used in this study. However, this says little about the uncertainty of individual profiles. Since there is a large instrument to instrument variability, this rough estimate is not directly comparable to the uncertainties estimated in the NDACC data for each individual profile.

Indeed, Section 6 gives an overview of the general performance of CFH. In the first revision of the manuscript, we had added a section to describe the uncertainty on a per-profile basis (section 5.5).

I had suggested making a comparison with the NDACC-derived uncertainties. The only place I found such a comparison was in Figure 6, with the implication that the NDACC derived uncertainties were insufficient. In the profile shown here the uncertainties are indeed insufficient and the data can be improved using the "golden points" method. However, this Figure and discussion is not suitable for a larger statistical comparison of how often the NDACC provided measurements are outside the combined measurement uncertainty in the comparison with the "golden point" processed data.

We appreciate the suggestion, but believe it would go far beyond the scope of this paper, extend the length of the paper by several pages, and be best implemented by the NDACC processors.

Length of the manuscript

The authors have restructured their manuscript to some extent, and some reordering of sections has been helpful. However, rather than shorting it, the manuscript is now even longer. This was not helpful for this review and will not be helpful to the reader. Frequent cross references (such as the phrase "see section xx") and reference to an Appendix are also not helpful. Instead of going through the paper again in detail and suggesting more places where paragraphs can be shortened or cut altogether, I leave this task to the authors. I still feel that cutting 50% of the text without much loss of relevant content is easily achievable and would improve the overall clarity.

Just as a simple example, the description of the basics of frostpoint hygrometry in section 2.3 can be dramatically shortened and appendix A deleted. A reference to the relevant literature, where needed, would be sufficient rather than summarizing this literature.

This article addresses a complex topic and draws conclusions that could have far-reaching implications. We understand that Rev. 2, who is apparently a highly experienced expert, would have written this article shorter, though certainly not by 50%.

It is simply not true that the revised manuscript "is now even longer". The total length (text, bibliography, Table A1) was 73 pages in the initial submission; this was reduced to 60 pages, i.e., by 18%, without compromising clarity. In addition, we shortened the main text by moving material to the appendices. However, Reviewer 2 asked us to discuss noise and uncertainty in the measurement, which added additional material, which resulted in an expansion by 7 pages, so that the final page count is 67. We agree with the reviewer that this is important information and it improved the manuscript, but at the same time led again to an increase in size, but still shorter than originally.

Minor points

Line 109f: This sentence reads as if no such long-term program existed. Clearly several stations already conduct routine measurements for several decades with the station at Boulder being the pioneer with well over 40 years.

Good point, we have replaced the sentence with: "Several stations already conduct routine balloon-borne measurements of stratospheric water vapor for several decades, with the station at Boulder being the pioneer with well over 40 years".

Line 167f: The authors have already identified that electronic noise is at least an order of magnitude smaller than the variation seen in flight. They should not state here that this noise is constraining the measurements.

Good point, we have replaced "the electronic noise of its detection unit (in mV)" with "the signal-to-noise ratio of its detection unit (in mV)"

Line 238: I could not find such a quote in the Bakerian lecture, although Dobson would probably not have disagreed.

We appreciate the comment, we have deleted this sentence.

Line 287: This is one of several places that point to a timing issue between “golden point” mirror reflectance and “golden point” mirror temperature. See my comments above that I do not believe this to be a timing issue, but rather a more fundamental issue.

Thank you for your comment. We are aware that apparent “timing” issues could be attributed to the heat distribution time across the mirror and through the ice layer, which we estimate together to be approximately 120 ms, and that this could lead to time lag effects for very large dTm/dt values. We assume this is what you mean by “fundamental issues” and have now included this in the text (Section 4.1.3).

Equation 4 and Line 417f: I question the validity of this assumption at the golden points, since dT/dt is at a maximum at these points.

As we described in our response to your major comment on “Golden Points”

above, the heat distribution across the mirror is slower and, thus, more relevant than the thermal diffusion through the ice layer (and Eq. 4). We have updated our heat equilibration time estimate in Section 4.1.3.

Line 486ff: What electronic delays might be involved here other than what is interpreted as delay between mirror reflectance and mirror temperature. What data transmission problems? Missing data points? Electronic noise is not an issue. I believe this sentence can be deleted.

Thank you for pointing that out. We have removed points (i) and (ii) (i.e., delay and noise in Um).

Line 498: That the CFH measurements are excellent at the “golden points” is an assumption, quite reasonable in fact, but still an assumption. Here it is overstated as a fact. The uncertainties are based not on “golden points” but on regular control.

Sorry, we do not understand what you mean by ‘regular control’. If you mean good PID control, then indeed, the 0.2 K uncertainty for CFH is due to the +-1 s uncertainty together with fast heat transport across the mirror disk with generally small |dTm/dt|. We have added the following sentence to Section 5.5:

“This is because $|dT_m/dt|$ is generally well-constrained by CFH’s PID controller, and heat equilibrates rapidly (~ 120 ms) across the CFH mirror disk.”

Line 1078 ff: This is a truly important result, i.e., some layers of extreme low-bias can be identified and corrected. The following statement is important for trend detections. This should be highlighted in the abstract. It does point to a systematic weakness for some layers in some publicly archived soundings.

Thank you for this comment.

We added the following sentence to Section 6.2 and updated the abstract:

"In particular, some layers of extreme-low bias can be identified and corrected, which is important for trend detection."

Line 1123ff (and discussion about time lag correction for frostpoint hygrometers): I am not aware that this is done anywhere.

There are examples of time-lag correction and time constants >> 1 s for CFH in the literature (e.g., Hasebe et al., 2013; Vömel et al., 2016) and WMO documents (e.g. CIMO guide 2023). We believe it is important to point out that time-lag correction is generally not appropriate, even if, we as well, are not aware that this is done elsewhere, especially on routine measurements.

We have reformulated the sentence as:

In contrast, attempts to consider chilled mirror instruments as time-lagged (i.e. following a first-order response) as found in literature are generally inappropriate.