



An update on the uncertainties of water vapor measurements using Cryogenic Frostpoint Hygrometers

Holger Vömel¹, Tatjana Naebert², Ruud Dirksen², Michael Sommer²

5 ¹Earth Observing Laboratory, National Center for Atmospheric Research, Boulder, CO 80303, USA
²Meteorologisches Observatorium Lindenberg, Deutscher Wetterdienst, 15848 Lindenberg, Germany

Correspondence to: Holger Vömel (voemel@ucar.edu)

Abstract. Long time series of observations of essential climate variables in the troposphere and stratosphere are often impacted by inconsistencies in instrumentation and ambiguities in the interpretation of the data. To reduce these problems of long term data series all measurements should include an estimate of their uncertainty and a description of their sources. Here we present an update of the uncertainties for tropospheric and stratospheric water vapor observations using the Cryogenic Frostpoint Hygrometer (CFH). The largest source of measurement uncertainty is the controller stability, which is discussed here in detail. We describe a method to quantify this uncertainty for each profile based on the measurements. We also show the importance of a manufacturer independent ground check, which is an essential tool to continuously monitor the uncertainty introduced by instrument variability. A small bias, which has previously been indicated in lower tropospheric measurements, is described here in detail and has been rectified. Under good conditions the total from all sources of uncertainty of frostpoint or dewpoint measurements using the CFH can be better than 0.2 K. Systematic errors, which are most likely to impact long term climate series are verified to be less than 0.1 K.

1 Introduction

20 Cryogenic frostpoint hygrometers (Vömel et al., 2007a, Vömel et al. 1995) are widely considered as reference instruments for measurements of tropospheric and stratospheric water vapor due to their ability to relate the amount of water vapor to an easier temperature measurement. Frostpoint hygrometry makes use of the equilibrium thermodynamics of the vapor phase of water and its condensed phase. The relation between the vapor pressure and the equilibrium temperature of the condensed phase is described by the Clausius-Clapeyron equation, and the partial pressure of the vapor phase can be calculated using the measured frostpoint (or dewpoint) temperature. Frostpoint or dewpoint hygrometers actively control the equilibrium of a two-phase system consisting of water ice (or liquid) and water vapor by continuously adjusting the temperature of the condensed phase such that the condensed phase neither grows nor shrinks, i.e. the condensed phase is in equilibrium with the gas phase. The temperature variations of the condensed phase needed to maintain the equilibrium are small, provided that the active control is successful, and the mean condensate temperature approximates the frostpoint or dewpoint temperature, which is then used to calculate the partial pressure of the vapor phase.



Implementations of the frostpoint (dewpoint) technique vary significantly and most limitations are a result of the characteristics of the individual implementation (Vömel and Jeannot, 2013). Therefore, not all frostpoint or dewpoint hygrometers are equivalent, and some understanding of the technical realization is needed to properly interpret the reported
5 frostpoint temperature and to be able to estimate the measurement uncertainty.

Here we focus on the Cryogenic Frostpoint Hygrometer (CFH), which has been described in detail elsewhere (Vömel et al., 2007a). This instrument uses the chilled-mirror principle, in which a small mirror is exposed to the ambient air and whose temperature is controlled such that liquid water or ice condenses on the mirror. An optical detector senses this condensate
10 and a digital controller regulates the temperature of the mirror in order to maintain a constant reflectivity of the condensate on the mirror surface. To the extent that the reflectivity is constant, the condensate on the mirror is assumed to be in equilibrium with the gas phase.

These instruments have been used in a large number of studies of upper tropospheric and stratospheric water vapor (e.g.
15 Vömel et al., 2007b; Hasebe et al., 2007; Fujiwara et al., 2010; Selkirk et al., 2010; Shibata et al., 2010). Although they are recognized by many as reference instruments, we refer to the rigorous definition of what constitutes a reference observation given recently by Immler et al. (2010). This paper defines the measurement philosophy of the GCOS Reference Upper Air Network (GRUAN), which aims at providing traceable and well characterized observations of essential climate variables within the troposphere and stratosphere, which may also be applied to other types of atmospheric observations.

20 The present paper discusses the measurement uncertainties of the CFH within the framework laid out by Immler et al. (2010). The work leading up to this paper has resulted in some instrument improvements over the work presented by Vömel et al. (2007a) which are discussed here. The sequence of processing and data quality control steps from raw data to final data product is described in Appendix A.

25 Throughout the paper uncertainties are expressed as one standard deviation, following the recommendations given by the Guide to the expression of uncertainty in measurement (JCGM/WG1, 2008).

2 Sources and type of measurement uncertainty

2.1 Mirror temperature controller

30 Frostpoint instruments do not measure instantaneous water vapor concentrations, but rather average frostpoint temperatures over (short) time periods, during which the mean mirror temperature can be assumed to be in near equilibrium with the condensed phase. Typically a proportional-integral-derivative (PID) controller is used to actively regulate the bulk



reflectivity of the condensate layer, because the stability of the condensate layer is the prerequisite of equilibrium. Mirror temperature readings are taken at frequencies higher than the equilibration time, i.e. typically once per second or faster, but individual temperature readings are a poor measure for the equilibrium temperature between gas phase and condensed phase on the mirror. Figure 1 shows one-second resolution raw data from two instruments on the same balloon launched at
5 Lindenbergl, Germany. This dual sounding was used to test different controller settings in the second instrument and demonstrates the impact of controller settings on the stability of the controller. To interpret these data properly, the dynamic response of the instruments must be considered and appropriate averaging intervals should be used to smooth out the controller oscillations. For high quality instruments, the averaging intervals will be short compared to the rate of change of ambient water vapor.

10

Balloon borne instruments encounter a large range of water vapor concentrations, which may span up to 5 orders of magnitude in partial pressure between the surface and the stratosphere and changes over two orders of magnitude may happen within a few tens of seconds in some soundings. Therefore, instruments need to respond quickly. A slow response may lead to delays in the detection of sharp atmospheric layers and to an underestimation of changes in atmospheric water
15 vapor. Slow-responding instruments may also be susceptible to measurement artifacts and external disturbances in the measurement, which may be impossible to distinguish from true atmospheric water vapor signals. On the other hand, fast-responding instruments tend to show oscillatory behavior, which complicates the interpretation of the raw signals (see e.g. Figure 1). In laboratory settings, mirror temperature oscillations may be as small as a few millikelvin, while under atmospheric conditions oscillations in poorly behaving instrument may be as large as a few kelvin.

20

The response of a PID controller to changes in its process variable, here the mirror reflectivity, depends on the choice of coefficients, which control the strength of the individual feedback process. Choosing the proper coefficients is called tuning and is the most difficult part of PID controller implementation. The tuning of the CFH is in favor of fast response, which may lead to oscillations that are sometimes slightly larger than desirable, in particular in the presence of sunlight. On the
25 other hand, the presence of some oscillations indicates that the instrument responds quickly to changes in water vapor. These oscillations may be smoothed out using appropriate filters, as described below.

Instruments showing no oscillations are likely to suffer from slow controller response in the detection of atmospheric water vapor. If no other measurements are available, it may be difficult to quantify this lag. Therefore, the presence of slight
30 oscillations is often preferred.

The uncertainty due to controller stability has been considered the largest source of measurement uncertainty in balloon borne frostpoint hygrometers (Vömel et al., 2007a) and was estimated at 0.5 K based on experience. In this section we revisit this subject and quantify the controller stability based on measurements. The quantified controller stability is one of the



terms in a vertically resolved uncertainty estimate and allows users a better interpretation of frostpoint observations under all levels of instrument performance.

To quantify the controller stability we use the amplitude of frostpoint oscillations around the mean frostpoint temperature over a time interval. To estimate the oscillation amplitude, the frostpoint profile is first smoothed using a Gaussian filter. The width of the Gaussian filter varies with altitude and depends on the ambient water vapor partial pressure and the performance of the instrument:

$$\bar{\chi}_i = \sum_j c_{i,j} \chi_j \quad (1)$$

10 where the filter coefficients

$$c_{i,j} = \frac{1}{\gamma_i} e^{-\left(\frac{t_i - t_j}{\tau_i}\right)^2} \quad (2)$$

and

$$\gamma_i = \sum_j e^{-\left(\frac{t_i - t_j}{\tau_i}\right)^2} \quad (3)$$

Here χ_j is the original frostpoint temperature or mixing ratio measurement at time-step j and $\bar{\chi}_i$ is the value of the filtered time series at time-step i . τ_i is the width of the Gaussian kernel at time-step t_i and provides a measure for the time interval over which the data are being smoothed. τ_i varies with frostpoint temperature and is provided as part of the data.

The amplitude of the oscillations which have been smoothed out is a direct measure of the residual uncertainty of the smoothed profile and a quantification of the controller-induced measurement uncertainty. This uncertainty is expressed by a weighted standard error of the mean (Gatz and Smith, 1995 and Endlich et al., 1986).

20

$$\sigma_{\bar{\chi}_i}^2 = \frac{n}{n-1} \left[\sum_j (c_{i,j} \chi_j - \bar{c}_i \bar{\chi}_i)^2 - 2 \bar{\chi}_i \sum_j (c_{i,j} - \bar{c}_i) (c_{i,j} \chi_j - \bar{c}_i \bar{\chi}_i) + \bar{\chi}_i^2 \sum_j (c_{i,j} - \bar{c}_i)^2 \right] \quad (4)$$

Due to the action of the PID controller and the thermal lag of the mirror and condensate dynamics frostpoint observations cannot be considered randomly distributed and show some degree of autocorrelation. This autocorrelation leads to an underestimate of the uncertainty using Equation 4. For simplicity we assume a first order auto-regressive model, for which the underestimation due to the autoregression can be expressed as

25

$$\widehat{\sigma}_{\bar{\chi}_i}^2 = \frac{\sigma_{\bar{\chi}_i}^2}{1 - \rho^2} \quad (5)$$

30 The autoregression parameter ρ can be estimated by the time lag λ of the instrument as $\rho = e^{-\frac{1}{\lambda}}$ and depends on the detector characteristics, the PID settings as well as the frost coverage. Comparisons with the Fast In situ Stratospheric Hygrometer



(FISH) during the AquaVIT-1 laboratory intercomparison (Fahey et al., 2014) indicate that the time response of the CFH at very low frost point temperatures is faster than 20 s. Here we use soundings with pronounced stratospheric features to study time lag in the CFH by comparing ascent and descent profiles. A profile at Lindenberg launched on 20 May 2014 showing two prominent features in the stratosphere at 22 km and 23.5 km was sampled both on ascent and descent (Figure 2). The time difference between the sampling of these features on ascent and descent is small (about 10 min), since the balloon burst at 24 km altitude. Furthermore, the wind shear in this altitude region was small, so that nearly the same air mass was sampled both on ascent and descent. The parachute opened immediately after balloon burst and slowed the instrument to less than 25 m/s. Time lag impacts measurements both on ascent and descent; however, due to the faster velocity this time lag is more noticeable on parachute descent. Applying a time lag correction following Miloshevich et al. (2004) to the entire profile data using a stratospheric time constant of $\lambda = 10$ s brings the features during ascent and descent into agreement. The same result was obtained using data from another sounding launched on 22 June 2011 at Lindenberg (not shown). Therefore, we may use this stratospheric time lag to estimate the autocorrelation coefficient and the frostpoint uncertainty given in Equation 5. Note that in the troposphere the time lag is estimated to be less, although we currently don't have a good way to quantify it. We estimate it as $\lambda \leq 1$ s near the surface, which is less than the temporal resolution of the measurements.

Since the uncertainty estimate expressed in Equation 5 is based on the variance of the data over a particular time period, it represents the uncertainty of \bar{x}_i over this time period and not of individual one-second data points. Therefore, uncertainty due to controller response and vertical resolution are directly related. The statistical uncertainty can be further reduced by widening the kernel, however at the expense of a reduced vertical resolution. If further vertical averaging is required it is essential to know over which time period the uncertainty should be applied and how the uncertainties of neighboring layers can be averaged.

The width of the averaging kernel is chosen manually for a best compromise between uncertainty and resolution and typically varies from 3 s in the lower troposphere to 30 s in the middle stratosphere. For well-behaved instruments the uncertainty due to controller oscillations may be significantly less than the estimate by Vömel et al. (2007a). However, for fast-changing features in the troposphere and or poorly behaving instruments the uncertainty due to controller stability may be larger. Providing this vertically resolved uncertainty estimate together with the data will help the user in the interpretation of the data.

Figure 3 shows the profiles already shown in Figure 1 but smoothed using the Gaussian filter. The left panel show the smoothed data of both profiles, and the center panel shows the uncertainty estimates using the kernel widths shown in the right hand panel. Since the raw data for instrument 2L2506 (Figure 1, right panel) are significantly noisier, a wider filter has been applied to these data (Figure 3, right panel). The resulting uncertainty is also larger than for instrument 2L2501 (Figure 3, center panel).



The mixing ratio difference between the smoothed profiles is shown in Figure 4 together with the combined uncertainties which have been added in quadrature. Throughout most of the stratosphere the smoothed profiles agree to within 0.25 ppmv. The combined uncertainty is slightly larger than this, and is dominated by the controller oscillations of instrument 2L2506. This plot shows that the mean difference between these two profiles is significantly better than 5% for most parts of the profile and that the total uncertainty is roughly of that magnitude. Only at 15.5 km is the difference between the profiles significantly larger than 5%, a result of the poorer behavior of instrument 2L2506. However, since the uncertainty of the measurements of this instrument at that altitude is quite large, both instruments are considered in agreement within their combined uncertainty.

Several of the instruments during the 2010 WMO radiosonde intercomparison campaign at Yangjiang, China, (Nash et al., 2011) suffered from poor controller stability. The data processing of that campaign provided for the first time a simple estimate of the controller stability, which has been refined here. Using the experiences of this campaign the tuning of the PID controller and the noise characteristics of the detector have been significantly improved leading to much better performance throughout the entire troposphere and the stratosphere in almost all conditions. Current instruments achieve controller uncertainty as low as 0.1 K (e.g. Figure 1, left panel).

For simplicity the profiles shown here have been limited to the stratospheric part of the sounding, though all arguments given here apply to the tropospheric part of the profile as well. Data files generated with this processing provide in addition to the actual measurement, the vertically resolved kernel width (i.e. vertical resolution) and the vertically resolved uncertainty estimate. Therefore, researchers using CFH measurements should also pay attention to the uncertainty estimate and the vertical resolution.

2.2 Calibration

Frostpoint and dewpoint hygrometers rely on the accurate calibration of the mirror temperature measurement. They are not calibrated against water vapor standards and are considered water vapor standards in themselves. The uncertainty of the frostpoint temperature due to calibration and other instrument specific factors is considered to be small with a total of about 0.1 K (Vömel et al., 2007a). Here we present an update on the long term calibration stability. Five thermistors have been recalibrated in more than 40 different calibration runs, which allows identification of possible drifts in the reference thermometer and calibration setup. Figure 5 shows the calibration deviation of the mean of these thermistors at four select temperatures. Over the entire record, the variation is less than 0.02 K, which for stratospheric water vapor measurements is equivalent to a mixing ratio variation of about 0.4%. Most importantly, there is no significant drift in this calibration record. In addition to the accuracy of the thermistor calibration, the accuracy of the resistance measurement circuitry has been tested for 508 CFH instruments by measuring a high precision 10 k Ω reference resistor with the circuitry of each instrument prior



to installation into the instrument. This 10 k Ω resistance is equivalent to a temperature of -35.84°C using average instrument calibration coefficients. At this equivalent temperature the instruments showed a negligible bias of -0.006 K +/- 0.006 K. The standard deviation shows the repeatability of the resistance measurement among all instruments. The repeatability of the resistance measurements at 100 k Ω (corresponding to a temperature of -77.56°C) is less than 0.005 K and at 1k Ω (22.7°C) it is 0.08 K. At any of these standard resistances that were used to test the CFH measurement electronics, no drift could be detected.

Seven dual CFH soundings conducted at Lindenberg, one of which is shown in Figure 1, are investigated in greater detail to verify the total calibration uncertainty. The main goal of these soundings was to test new developments or modifications while at the same time maintaining the consistency of the observations. For each dual sounding the differences of the two profiles have been averaged over the entire troposphere, over the entire stratosphere and over the entire profile. Using large layer averages removes all random effects due to controller stability and other random processes in any particular profile. In the first of these dual soundings a manufacturing defect was found in one instrument, which led to a very large difference between the two instruments. This sounding has therefore been removed from the statistics. Within the remaining six dual soundings the standard deviation of the difference is 0.11 K in the troposphere and 0.09 K in the stratosphere. The standard deviation of the difference over the entire profile is 0.08 K.

These six dual soundings imply that on average any two randomly-picked instruments are expected to agree with each other to within better than 0.11 K. This difference is only slightly larger than the expected agreement of about 0.1 K and is in part due to the fact that the second instrument in each pair was used to test modifications rather than to test a second identical instrument.

2.3 Laboratory tests

In 2007 an intensive laboratory campaign took place at the Aerosols Interaction and Dynamics in the Atmosphere (AIDA) test chamber at the Karlsruhe Institute of Technology (Fahey et al., 2014). During this campaign stratospheric and upper tropospheric water vapor concentrations were used to test a large number of water vapor instruments. The experiments conducted during the first week of the campaign were averaged to remove any random effect and short term deviations, which is comparable to the stratospheric and tropospheric averages of the dual soundings. The mixing ratios measured by the CFH were within 10% of the campaign reference (approximately within 0.5 K frostpoint temperature) for all static levels during that experiment and within 4% (approximately within 0.2 K frostpoint temperature) for most levels. Although this campaign was not intended as a comparison against an absolute reference, this result may be taken as indication that a calibration uncertainty estimate of better than 0.2 K is a reasonable estimate.

The calibration uncertainty term is assumed to be a fixed value over the entire profile. It therefore needs to be added in square to the vertically resolved controller stability. However, this must be done after any vertical or temporal averaging of the controller stability has been done, since this term cannot be reduced by increasing the averaging. The calibration uncertainty is the lower limit for the overall measurement uncertainty that current production technology can achieve.



2.4 Measurement bias

Miloshevich et al. (2009) indicated that the CFH may have some unexplained bias in the lower troposphere. Furthermore Nash et al. (2011) indicated poor control and possible bias in some instruments. These biases and poor control impacted some instruments and have motivated further instrument improvements. The most important of these is a modification in the production procedure related to installation of the thermistor into the mirror. The location of thermistor in the mirror and its installation has been largely the same since the original design of the NOAA frostpoint hygrometer by Mastenbrook. Although the location of the mirror temperature measurement has previously been found not to be critical (Vömel et al., 2007a), we found that small variations in the manufacturing process may have had a noticeable influence on the accuracy of the temperature measurement. To eliminate the possibility of this bias the assembly procedure of the thermistor has been changed. This change was introduced in the production starting with 2L2901.

The impact of this change can be illustrated in the frequency of supersaturation over liquid water observed in the lower troposphere. Supersaturation over liquid water is expected to be small in the real atmosphere and measurements showing significant supersaturation over liquid water are most likely related to high biases in the measurement of water vapor or cold biases in the measurement of air temperature. From a set of 1022 available CFH soundings, we extracted the maximum values of relative humidity over liquid (RH) in the lower troposphere measured in each profile. Figure 7 shows the distribution of this peak value grouped by CFH version. Here we used all available data, including profiles that had been flagged as suspicious in the original data processing and ignored in previous scientific analyses. The lower panel shows the distribution of peak values of RH for all instruments up to serial number 2L28xx. This includes all instruments interfaced to the Vaisala RS80 radiosonde, the InterMet iMet-1 radiosonde and the Meisei RS06G radiosonde. The upper panel shows the distribution for all available data from instruments starting with serial number 2L2901. Fifteen percent of instruments from prior to the production change show supersaturation over liquid water larger than 5%, which is a direct indication of a possible bias in these instruments. This bias is the same that had been noted already by Miloshevich et al. (2009). In the top panel, which shows the observations from all instruments after the manufacturing change, only 1.6 % of the soundings show a significant bias. Therefore, we can conclude that the bias that had impacted some instruments in the past has been effectively eliminated in all instruments starting with serial number 2L2901.

Suspected bias in the older instruments can be corrected using an empirical model, which depends on measured frostpoint temperature and pressure:

$$Fp' = Fp - k(Fp - 0.964 \ln(P)^{2.2} - 141.328) \quad (6)$$

where Fp' is the corrected frostpoint temperature, Fp is the originally measured frostpoint temperature, P is the ambient pressure, and k is an empirical correction constant for the entire profile, which has to be estimated based on the level of supersaturation observed in the lower troposphere. The change in frostpoint temperature impacts lower tropospheric observations more strongly than upper tropospheric measurements. Figure 8 shows the distribution of the correction as a function of altitude for all soundings, where a bias is suspected. The most common correction in the lower troposphere is



around 0.4 K, dropping to about 0.2 K in the stratosphere. This correction is within the previously published total uncertainty estimate of 0.51 K; however, since this is a systematic error, it will impact large data sets, trend estimates, or comparisons with satellite instruments. Although most of the impacted instruments show a bias within the previously published uncertainty estimates, a few instruments show significantly larger biases. All of these were noted in the initial data
5 processing and were either rejected or empirically corrected. The right hand panel of Figure 8 shows the weighted average correction profile, which is strongly influenced by large outliers. The weighted mean correction for these soundings varies between 0.45 K in the lower troposphere and 0.3 K in the stratosphere. This change is equivalent to a change in mixing ratio of about 3% in the lower troposphere and about 5% in the stratosphere. Since only about 15 % of all instruments are suspected to be impacted by this bias, the impact on studies using a large number of soundings such as satellite comparisons
10 is considered small.

2.5 Frost layer morphology

It has been pointed out in the literature (e.g. Fujiwara et al., 2003; Vömel et al., 2003), that it is essential to know the phase of the condensate on the mirror of a frostpoint hygrometer to be able to convert the mirror temperature to a water vapor partial pressure. Above 0°C it is safe to assume that the condensate is liquid and below a condensate temperature of about -
15 35°C it is safe to assume that it is ice. In the mirror temperature range between 0°C and -35°C the condensate on the mirror may be liquid or ice or mixture of liquid drops and ice crystals. If the phase is not well known, the instrument cannot produce reliable reference data. The CFH addresses this issue by forcing the condensate to freeze at a condensate temperature of -15°C (Vömel et al., 2007a). Above this temperature the condensate is almost always supercooled water. Visual inspection of each profile and the associated engineering data is required to guarantee a unique identification of the
20 condensate phase.

Different condensate layer morphologies may introduce unexpected behavior depending on the details of the PID implementation. Figure 10 shows four images of the condensate layer on the CFH mirror which has a diameter of about 7 mm. These images were obtained during the AquaVIT-1 laboratory intercomparison with a specially-built laboratory version of the CFH (see Fahey et al., 2014) that allowed visual observation of the mirror in parallel with the detector. The top left
25 image shows a layer of supercooled liquid water droplets at $T = -28.0^{\circ}\text{C}$ that had been generated purposely. The instrument is able to maintain this condensate layer with good control and therefore measures the dewpoint temperature. This image is typical for dewpoint measurements. The top right image shows a fine frost condensate layer at $T = -87.9^{\circ}\text{C}$. This is a typical frost layer image for most properly prepared instruments. The morphology of this condensate layer is reasonably stable for extended periods of time. That is, once it has been formed, its structure does not significantly change throughout the time of
30 a typical sounding. The image on the bottom left shows a coarse frost layer with patches of small liquid droplets at $T = -26.8^{\circ}\text{C}$. Although the PID controller is able to maintain a stable reflectivity, the mirror temperature is not a measurement of the amount of water vapor, since both liquid droplets and small ice crystals are present on the mirror at the same time. All liquid droplets must freeze before the mirror temperature represents the frostpoint temperature instead of the dewpoint



temperature. The force-freezing algorithm avoids this ambiguity in a sounding instrument. The image on the bottom right shows a coarse frost layer at $T = -88.9^{\circ}\text{C}$. Here the PID controller is unstable and not able to properly control the amount of condensate. This condition may be caused by a number of factors, one of which is the cleanliness of the mirror. Instrument operators normally clean the mirror prior to launch to avoid this issue and to guarantee that the frost layer morphology is consistent between all soundings. Improperly cleaned mirrors may lead to very different condensate layer morphologies and thereby to unexpected instrument behaviors. These conditions give rise to substantially increased uncertainties in the measurements, which however are difficult to quantify.

5
These images highlight the fact that the proper instrument preparation and operation may strongly impact the instrumental uncertainty. Therefore, the instrument preparation and set-up needs to be captured in the metadata to evaluate the quality of the sounding data after the observations have been completed.

10

2.6 Contamination and descent observations

Water vapor observations on balloon ascent always carry the risk that the measurements are contaminated by the outgassing of water vapor from the balloon envelope, the parachute, the load line or the intake tubes of the instrument. This risk is particularly elevated when the sounding passes through liquid water clouds where supercooled water may freeze on any surface. Most soundings show some degree of contamination in the highest parts of the stratospheric profile. With the design of the CFH contamination is almost never encountered in the troposphere, but in extreme cases may impact the data soon after passing the tropopause. Therefore, all soundings have to be visually screened and contamination has to be flagged manually.

15

Outgassing by the balloon envelope and the parachute can be minimized by using unwinders of at least 50-m length. Self-contamination by icing inside the inlet tubes cannot be avoided unless heated inlet tubes are being used, which so far has not yet been done on small sounding balloons. However, it can be minimized by using short polished stainless-steel inlet tubes such as those used by the CFH.

20

CFH soundings rely largely on ascent measurements, allowing the use of ascent pressure, temperature and humidity data from the parallel radiosonde, which are considered more reliable during ascent.

25 2.7 Trace gas interferences

The frostpoint principle is based on the two-phase system of condensed and gas phase water. There have been speculations that the presence of other trace gases may interfere with this basic assumption and change the equilibrium temperature. One of the atmospheric trace gases known to change the condensation temperature of ice in polar stratospheric clouds is nitric acid (Crutzen and Arnold, 1986). Nitric acid may co-condense together with water vapor to form solid nitric acid trihydrate, for which the frostpoint is a few kelvin above the frostpoint of pure water ice. Therefore the possibility exists that nitric acid co-condenses in the condensate layer on the mirror of a frostpoint hygrometer. Laboratory studies with different concentrations of nitric acid have shown that at atmospheric concentrations small amounts of nitric acid are incorporated into

30



the ice layer of a frost point hygrometer; however, these amounts are insufficient to change the equilibrium temperature between ice and gas phase water vapor (Thornberry et al., 2012). Laboratory tests using high concentrations of carbon dioxide also have not shown indication of any significant interference with the ice frostpoint temperature. Therefore, trace gas interferences are currently excluded as systematic error sources.

5 2.8 Discussion of uncertainties

The uncertainty due to controller stability is considered random and correlated over the time of the smoothing within a single profile. For long time series, this term can be considered uncorrelated as well as random and can be treated accordingly. Here it is assumed that the oscillations are symmetric around the true frost point temperature. However, for strongly oscillating instruments, this may not be the case due to the difference in frost formation and evaporation, in particular at cold temperatures. Therefore, for strongly oscillating instruments the uncertainty due to controller stability may mask a systematic error, which cannot be quantified at this point. However, the statistical uncertainty of these soundings is usually so large that the measurements become less significant and should not be used in time series analyses.

Likewise, for instruments with negligible uncertainty due to controller stability, a systematic error due to controller lag may be present which cannot be quantified. For tropospheric and stratospheric observations the CFH is a fast-responding instrument and lag issues are not suspected.

For stratospheric observations the agreement of measurements between ascent and descent can be taken as indication that possible systematic errors are less than the difference between the ascent and descent. Observations where large differences exist between ascent and descent data that are not due to a fast descent and not due to contamination may be due to frost-layer morphology issues. These soundings require an increase in the uncertainty estimate, which is at least as large as the difference between the ascent and descent measurements over the affected altitude regions.

The uncertainty due to calibration is that of the thermistor calibration and the measurement electronics. Both are considered small. Instrumental and production variability have caused larger deviations in the past and should therefore always be checked prior to launch. A pre-launch check (see following section) is recommended, which should consist of at least a one-point measurement compared to a known reference after the cryogen has been added to the instrument.

25 3 Ground check

Upper air observations using disposable instrumentation rely heavily on the stability of the manufacturing process, which is usually not in the control of the personnel using the equipment. Even though manufacturers of high-quality equipment pay great attention to the stability of their product, factors outside their control may impact the quality of their sounding equipment. The purpose of a ground check is to provide evidence that each instrument behaves as expected prior to launch and that instrumental issues that may be detectable prior to launch are indeed detected. The ground check procedure is most important for long term observations, where an independent verification of the instrument calibration is essential, even if it



cannot be done over the full range of parameters to be measured. It assures that at least a one-point traceability to a reference can be established to verify the long-term stability of the instrumentation.

The ground check for the CFH consists of a measurement of relative humidity inside the inlet tube prior to launch using an integrated temperature and polymer humidity sensor (Ahlborn FHAD462). A small propeller generates highly turbulent air motion inside the sensor housing to assure proper ventilation across the mirror. The polymer humidity sensor is regularly recalibrated using a NaCl salt solution at 75.3% and a MgCl salt solution at 32.8% RH which avoids any long term drifts of this sensor. The relative humidity measurement is converted to dewpoint temperature using the air temperature measurement inside the inlet tube which serves as the reference to the CFH dewpoint measurement. This ground check was introduced at Lindenberg in 2014 and has been used in a total of 31 soundings so far.

Figure 9 shows the time series of the differences between the CFH dewpoint measurements and the reference measurement inside the tube just prior to launch. The mean dewpoint difference for all 31 prelaunch checks is $\overline{DP_{CFH} - DP_{ref}} = 0.02 \pm 0.10$ K. The scatter of this comparison is due to the uncertainty in the CFH as well as in the reference measurements. Over this time period there is no apparent drift and the CFH agrees with this reference to better than 0.1 K. Under typical launch conditions an uncertainty of 0.1 K in dewpoint temperature corresponds to an uncertainty of about 0.7% in relative humidity, which is comparable to the uncertainty of the recalibration of the RH sensor used in these tests. Therefore, the uncertainty of the CFH water vapor measurements at the surface is likely better than expressed by this standard deviation.

Although the data record of ground checks is short, it demonstrates the stability of this test over this time period and the ability to identify potential issues, had any arisen. For long-term climate observations using frequently changing instruments this stability test prior to launch is essential to gain confidence in the data set (Immler et al., 2010). Although we did not start this test when the older instruments described in section 0 were still used, we are confident that the issues discussed there could have been identified much sooner using this test. Other issues may emerge in the future as technology develops and instrumentation changes. Having a continuous set of reference measurements such as this short set will be essential in defending the stability of these measurements.

4 Summary

The uncertainties of observations are one of the limiting factors in determining changes in atmospheric composition. Random errors are less likely to impact long-term trends; however, changes in systematic errors may impact long-term trends to the extent of the change. The oscillations around the mean of the frostpoint temperature, which are due to the action of the PID controller, provide a direct measure of the random uncertainty of the frostpoint temperature. It is important to note that the uncertainty refers to the uncertainty of the mean frostpoint temperature and not to the instantaneous measurement. It represents the uncertainty of a mean value in a vertical layer of the atmosphere. The extent of this vertical layer is provided by the time window that was used to smooth the data and is provided in the data files.



The systematic errors of the frostpoint measurements are largely given by the uncertainty of the calibration of the thermistor. This calibration is currently stable within 0.02 K. Parallel observations of two CFH instruments on the same balloon have so far verified an agreement to within 0.11 K throughout the entire profile, which is in part due to the fact that the second instrument has been used to test new modifications.

- 5 A manufacturer-independent ground check was introduced in 2014. This ground check demonstrates that the instruments behave prior to launch as expected and that the systematic error is less than 0.1 K. This systematic error limit of 0.1 K is currently the limiting uncertainty term for long term data series.

A systematic bias, which had been noted in previous studies in lower tropospheric water vapor measurements has been identified in some instruments and was traced back to manufacturing variability. A change in the production process has
10 eliminated this variability and thereby effectively removed this bias from the current production starting with serial number 2901. Even though the bias is smaller than the previously stated uncertainty, a correction algorithm can correct all affected profiles. Since this is a systematic error, it should be removed from instruments suspected to be impacted.

This bias is another example of the need for a manufacturer-independent ground check in reference observations using disposable sounding instrumentation. This experience highlights the need for such checks in all other sounding observations
15 if these are to be used for long-term climate observations.

Appendix: Data processing, filtering, smoothing

Raw data are transmitted once per second. The raw data consist of the mirror temperature, the reflectivity signal from the optical detector, the battery voltage and the detector temperature. Other housekeeping data may be transmitted as well, but are not considered essential. The CFH is programmed to perform one clearing pulse and one freezing pulse, during which the
20 condensate of the mirror freezes. These pulses last several seconds and their disturbance on the PID controller may impact the frost point measurement for a few tens of seconds. These data are flagged as bad data and should not be used in further processing.

The forced freezing is then examined more closely and a comparison with the parallel radiosonde verifies that the condensate changes phase during this freezing cycle. A condensate phase flag is set to indicate whether the condensate on the mirror is
25 liquid or ice. This distinction is essential since the partial pressure of water is calculated using the measured mirror temperature and the vapor pressure equation corresponding to the condensate phase. During winter or in conditions, where the dewpoint temperature at the surface is significantly below 0°C, the condensate may already be frozen prior to launch, in which case the freezing cycle does not lead to a change in the condensate phase. These situations are visually identified and properly flagged.

30 The stratospheric water vapor profile is examined and the region is identified, where contamination from outgassing is suspected. These data are flagged and must be ignored in future processing. If descent data are being used, then data during which the PID controller recovers after the beginning of the descent have to be flagged manually as well. For both



contamination and PID recovery experience is used as guide. Generally data are flagged conservatively to avoid misinterpretation of low quality data.

The data are then screened for unexpected issues. Although rare, data problems may arise from unusual balloon behavior, interferences from other instruments or unknown instrumental issues and any suspect data are flagged. All data are smoothed
5 using the low pass filtering described in Section 2.1. The smoothing algorithm also provides the uncertainty estimate for the controller stability as well as the time resolution for the uncertainty estimate. Layer averages can be calculated using only ascent data. Data from the parallel radiosonde and possibly ozone sonde are valid only on ascent and uncertainty estimates for these measurements can only be provided for ascent measurements.

Acknowledgements

10 The authors would like to thank Dr. Henry Selkirk and Dr. William Cooper for very helpful comments on this manuscript. The authors would also like to recognize the dedicated work by the observer staff of the Meteorological Observatory Lindenberg. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

References

- 15 Crutzen, P. J. and F. Arnold: Nitric acid cloud formation in the cold Antarctic stratosphere: a major cause for the springtime 'ozone hole', *Nature* 324, 651 – 655, doi:10.1038/324651a0, 1986.
- Dirksen, R. J., M. Sommer, F. J. Immler, D. F. Hurst, R. Kivi, and H. Vömel: Reference quality upper-air measurements: GRUAN data processing for the Vaisala RS92 radiosonde, *Atmos. Meas. Tech.*, 7, 4463-4490, doi:10.5194/amt-7-4463-2014, 2014.
- Endlich, R.M. et al.: Statistical analysis of precipitation chemistry measurements over the eastern United States, Utility Acid
20 Precipitation Study Program and Electric Power Research Institute, UAPSP Report 112, 3-9, 1986.
- Fahey, D. W., Gao, R.-S., Möhler, O., Saathoff, H., Schiller, C., Ebert, V., Krämer, M., Peter, T., Amarouche, N., Avallone, L. M., Bauer, R., Bozóki, Z., Christensen, L. E., Davis, S. M., Durr, G., Dyroff, C., Herman, R. L., Hunsmann, S., Khaykin, S. M., Mackrodt, P., Meyer, J., Smith, J. B., Spelten, N., Troy, R. F., Vömel, H., Wagner, S., and Wienhold, F. G.: The AquaVIT-1 intercomparison of atmospheric water vapor measurement techniques, *Atmos. Meas. Tech.*, 7, 3177-3213,
25 doi:10.5194/amt-7-3177-2014, 2014.
- Fujiwara, M., M. Shiotani, F. Hasebe, H. Vömel, S. J. Oltmans, P. W. Ruppert, T. Horinouchi, and T. Tsuda: Performance of the Meteorolabor "SnowWhite" chilled mirror hygrometer in the tropical troposphere: Comparisons with the Vaisala RS-80 A/H humicap sensors, *J. Atmos. Oceanic Technol.*, 20, 1534-1542, 2003.
- Fujiwara, M., H. Vömel, F. Hasebe, M. Shiotani, S.-Y. Ogino, S. Iwasaki, N. Nishi, T. Shibata, K. Shimizu, E. Nishimoto, J.
30 Valverde-Canossa, H. B. Selkirk and S. J. Oltmans: Seasonal to decadal variations of water vapor in the tropical lower



- stratosphere observed with balloon-borne cryogenic frost point hygrometers, *J. Geophys. Res.*, 115, D18304, doi:10.1029/2010JD014179, 2010.
- Gatz, D.F. and L. Smith: The standard error of a weighted mean concentration-I. Bootstrapping vs other methods, *Atmospheric Environment*, 29, 1185–1193, doi:10.1016/1352-2310(94)00210-C, 1995.
- 5 Hasebe, F., M. Fujiwara, N. Nishi, M. Shiotani, H. Vömel, S. Oltmans, H. Takashima, S. Saraspriya, N. Komala, and Y. Inai: In situ observations of dehydrated air parcels advected horizontally in the Tropical Tropopause Layer of the western Pacific, *Atmos. Chem. Phys.*, 7, 803-813, doi:10.5194/acp-7-803-2007, 2007.
- Immler, F. J., J. Dykema, T. Gardiner, D. N. Whiteman, P. W. Thorne, and H. Vömel: Reference Quality Upper-Air Measurements: guidance for developing GRUAN data products, *Atmos. Meas. Tech.*, 3(5), 1217–1231, doi:10.5194/amt-3-10 1217-2010, 2010.
- JCGM/WG 1, Evaluation of measurement data Guide to the expression of uncertainty in measurement, Bureau International des Poids et Mesures, www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf, Working Group 1 of the Joint Committee for Guides in Metrology, JCGM 100:2008.
- Miloshevich, L. M., A. Paukkunen, H. Vömel, and S. J. Oltmans: Development and validation of a time-lag correction for 15 Vaisala radiosonde humidity measurements, *J. Atmos. Ocean. Technol.*, 21, 1305-1327, 2004.
- Miloshevich, L. M., H. Vömel, D. N. Whiteman, T. Leblanc: Accuracy assessment and correction of Vaisala RS92 radiosonde water vapor measurements, *J. Geophys. Res.*, 114, doi:10.1029/2008JD011565, 2009.
- Nash J., T. Oakley, H. Vömel, and W. Li: WMO Intercomparison of High Quality Radiosonde Systems, Yangjiang, China, 12 July – 3 August 2010, Instruments and Observing Methods Report No. 107, World Meteorological Organization, 248 p, 20 2011.
- Selkirk, H. B., H. Vömel, J. M. Valverde Canossa, L. Pfister, J. A. Diaz, W. Fernández, J. Amador, W. Stolz, and G. S. Peng: Detailed structure of the tropical upper troposphere and lower stratosphere as revealed by balloon sonde observations of water vapor, ozone, temperature, and winds during the NASA TCSP and TC4 campaigns, *J. Geophys. Res.*, 115, D00J19, doi:10.1029/2009JD013209, 2010.
- 25 Shibata, T., H. Vömel, S. Hamdi, S. Kaloka, F. Hasebe, M. Fujiwara, and M. Shiotani: Tropical cirrus clouds near cold point tropopause under ice supersaturated conditions observed by lidar and balloon-borne cryogenic frost point hygrometer, *J. Geophys. Res.*, 112, D03210, doi:10.1029/2006JD007361, 2007.
- Thornberry, T., T. Gierczak, R.-S. Gao, H. Vömel, L. A. Watts, J. B. Burkholder, and D. W. Fahey: Laboratory evaluation of the effect of nitric acid uptake on frost point hygrometer performance, *Atmos. Meas. Tech.*, 4, 289-296, doi:10.5194/amt-4-30 289-2011, 2011.
- Vömel, H., S. J. Oltmans, D. J. Hofmann, T. Deshler, and J. M. Rosen: The evolution of the dehydration in the Antarctic stratospheric vortex, *J. Geophys. Res.*, 100, 13919-13926, 1995.
- Vömel, H., M. Fujiwara, M. Shiotani, F. Hasebe, S. J. Oltmans, and J. E. Barnes: The behavior of the Snow White chilled-mirror hygrometer in extremely dry conditions, *J. Atmos. Oceanic Technol.*, 20, 1560-1567, 2003.



Vömel, H., D. E. David, and K. Smith: Accuracy of tropospheric and stratospheric water vapor measurements by the cryogenic frost point hygrometer: Instrumental details and observations, *J. Geophys. Res.*, 112, D08305, doi:10.1029/2006JD007224, 2007a.

Vömel, H., H. Selkirk, L. Miloshevich, J. Valverde, J. Valdés, E. Kyrö, R. Kivi, W. Stolz, G. Peng, and J. A. Diaz: Radiation
5 dry bias of the Vaisala RS92 humidity sensor, *J. Atmos. Oceanic Technol.*, 24, 953–963, 2007b.

Vömel, H and P. Jeannot: Balloon-borne frostpoint-hygrometry, in *Monitoring atmospheric water vapour, ground-based remote sensing and in-situ methods*, ISSI scientific report series vol. no. 10, edited by Niklaus Kämpfer, pp. 39-54, Springer, New York, 10.1007/978-1-4614-3909-7, 2013.

Figures

- 10 **Figure 1: Simultaneous water vapor profiles using two CFH instruments, showing different controller response.**
- Figure 2: Water vapor profile launched at Lindenberg, showing a pronounced vertical feature, which is again detected on descent. Left panel: Raw data at 1 s resolution. Right panel: Ascent and descent data time lag corrected using a 10 s time lag constant.**
- Figure 3: Profiles shown in Figure 1 now smoothed using the Gaussian filter. The left panel shows the filtered profiles, the center panel their respective uncertainties, and the right hand panel their respective kernel widths.**
- 15 **Figure 4: Absolute mixing ratio difference of the two simultaneous profiles shown in Figure 3. Also shown is the combined uncertainty. To provide a guide for the relative difference, a 5% difference in mixing ratio is indicated by the thin dotted lines.**
- Figure 5: Deviation of the mean of five reference thermistors from the reference calibration temperature for each calibration run.**
- Figure 6: Comparisons frostpoint measurements in dual soundings. Blue dots indicate the tropospheric difference, red dots indicate the stratospheric difference and large orange dots indicate the mean difference for the entire profile.**
- 20 **Figure 7: Frequency distribution of measured peak RH values in the lower troposphere. The bottom panel shows the distribution using the CFH instrument through serial number 2L28xx. The upper panel shows the current instrument series, which started with 2L29xx. Peak values above 100% are an indication for bias either in the temperature measurement or in the water vapor measurement. While some significant measurements of supersaturation were taken using instruments through 2L28xx, no significant supersaturation has been observed with instruments starting with 2L29xx.**
- 25 **Figure 8: Deviation of the mean of five reference thermistors from the reference calibration temperature for each calibration run.**
- Figure 9: Difference of CFH dewpoint measurements prior to launch and a reference measurement inside the inlet tube. Dots indicate the individual measurements; the solid line indicates a linear fit through all data.**
- Figure 10: Images of the condensate layer at different conditions. The diameter of the mirror is 7 mm. Top left: supercooled liquid water at $T = -28.0^{\circ}\text{C}$. Top right: fine frost condensate at $T = -87.9^{\circ}\text{C}$. Bottom left: Coarse frost layer with patches of small liquid
30 drops at $T = -26.8^{\circ}\text{C}$. Bottom right: Coarse frost layer at $T = -88.9^{\circ}\text{C}$**

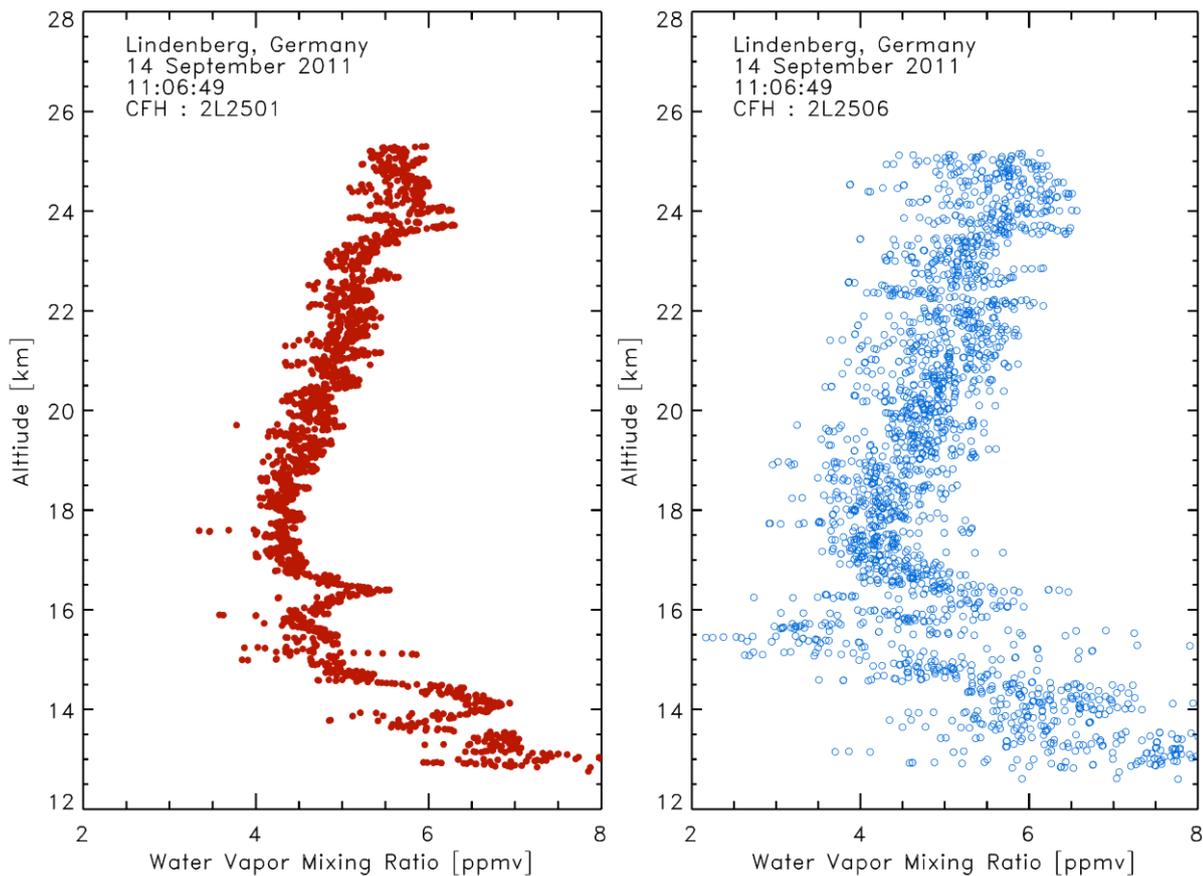


Figure 1: Simultaneous water vapor profiles using two CFH instruments, showing different controller response.

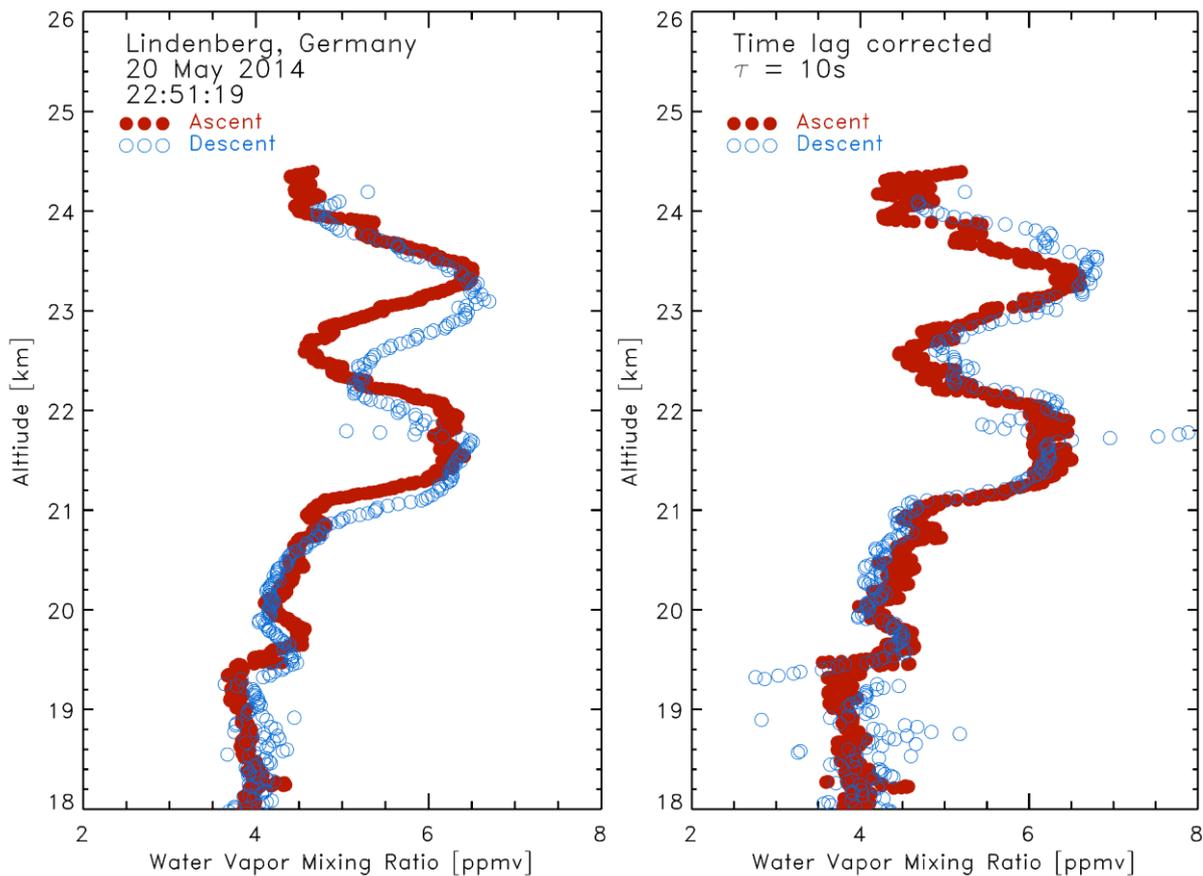


Figure 2: Water vapor profile launched at Lindenberg, showing a pronounced vertical feature, which is again detected on descent. Left panel: Raw data at 1 s resolution. Right panel: Ascent and descent data time lag corrected using a 10 s time lag constant.

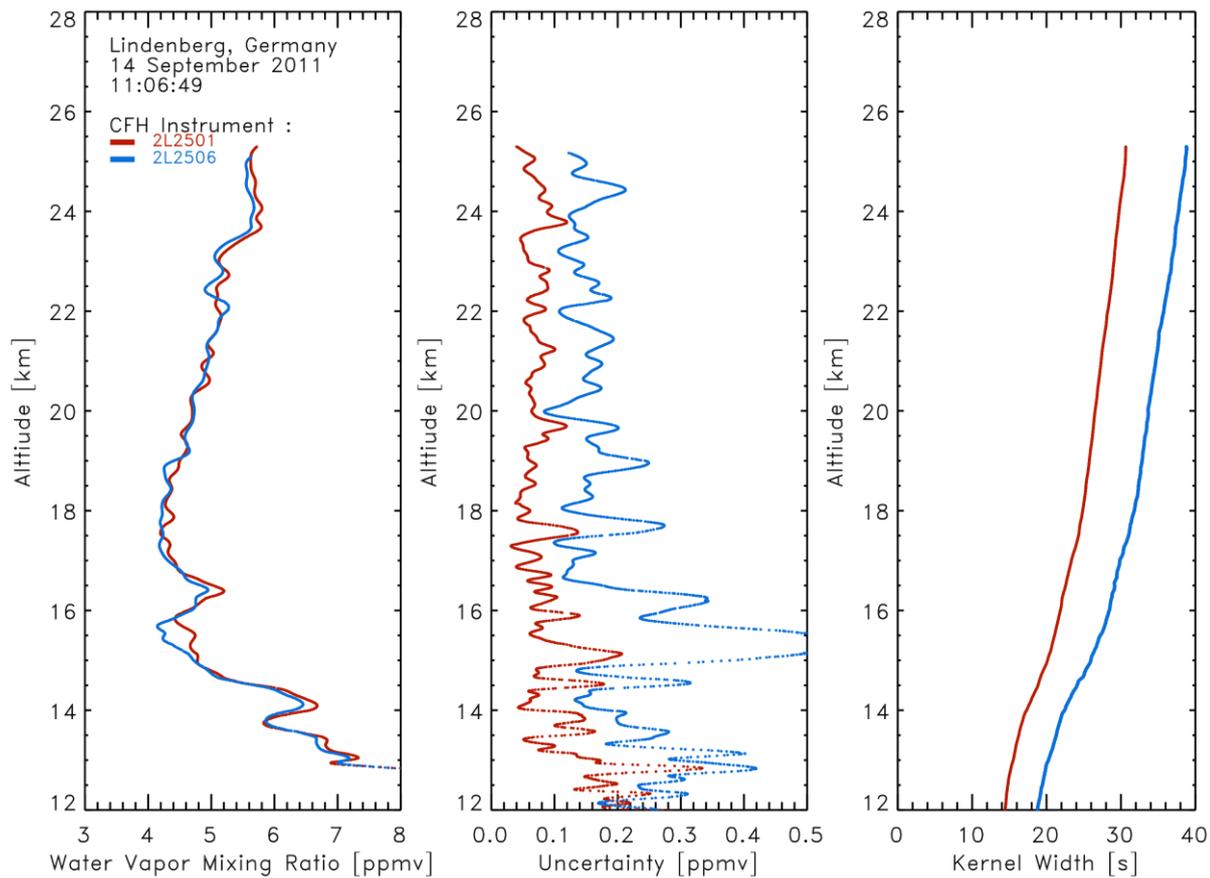


Figure 3: Profiles shown in Figure 1 now smoothed using the Gaussian filter. The left panel shows the filtered profiles, the center panel their respective uncertainties, and the right hand panel their respective kernel widths.

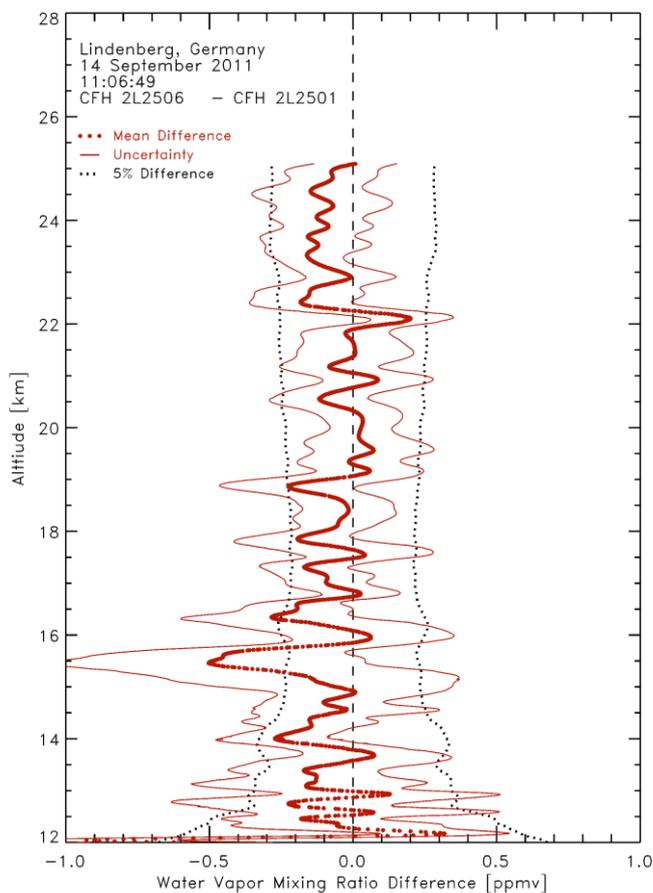


Figure 4: Absolute mixing ratio difference of the two simultaneous profiles shown in Figure 3. Also shown is the combined uncertainty. To provide a guide for the relative difference, a 5% relative difference in mixing ratio is indicated by the thin dotted lines.

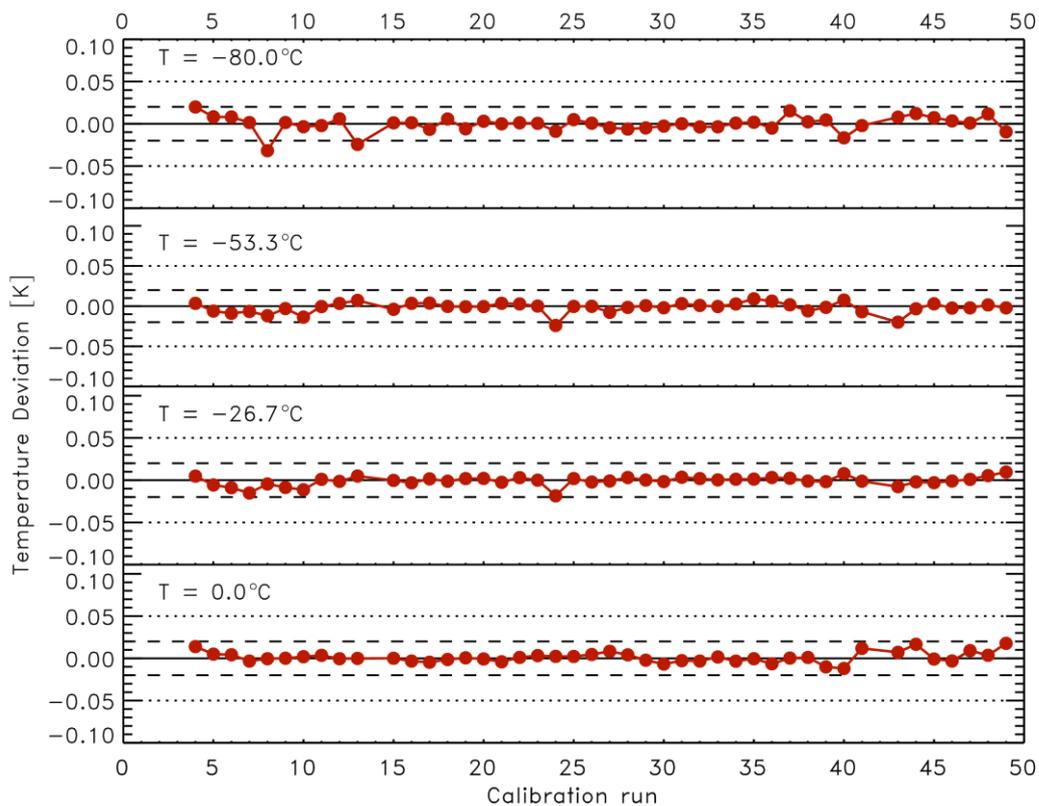


Figure 5: Deviation of the mean of five reference thermistors from the reference calibration temperature for each calibration run.

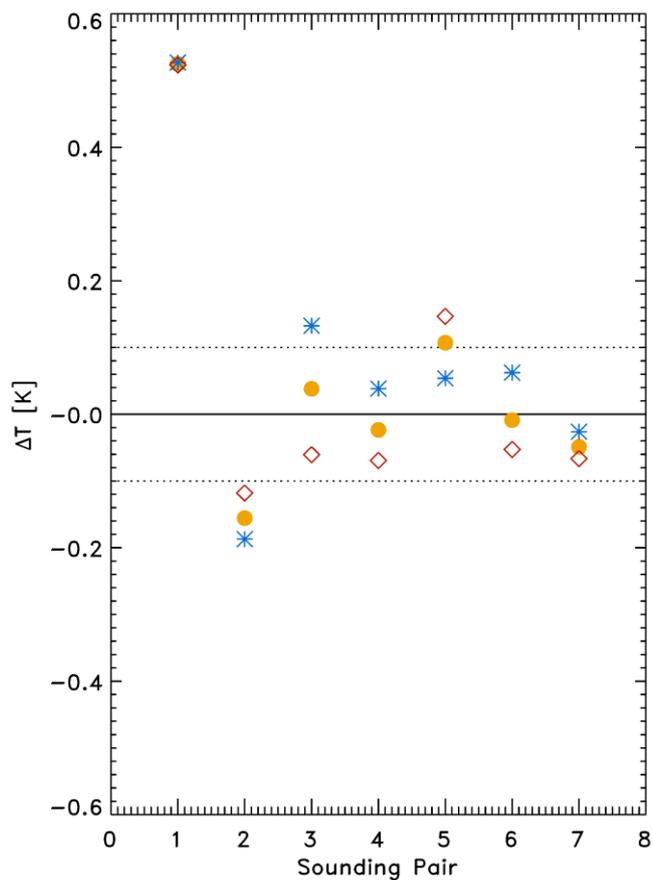
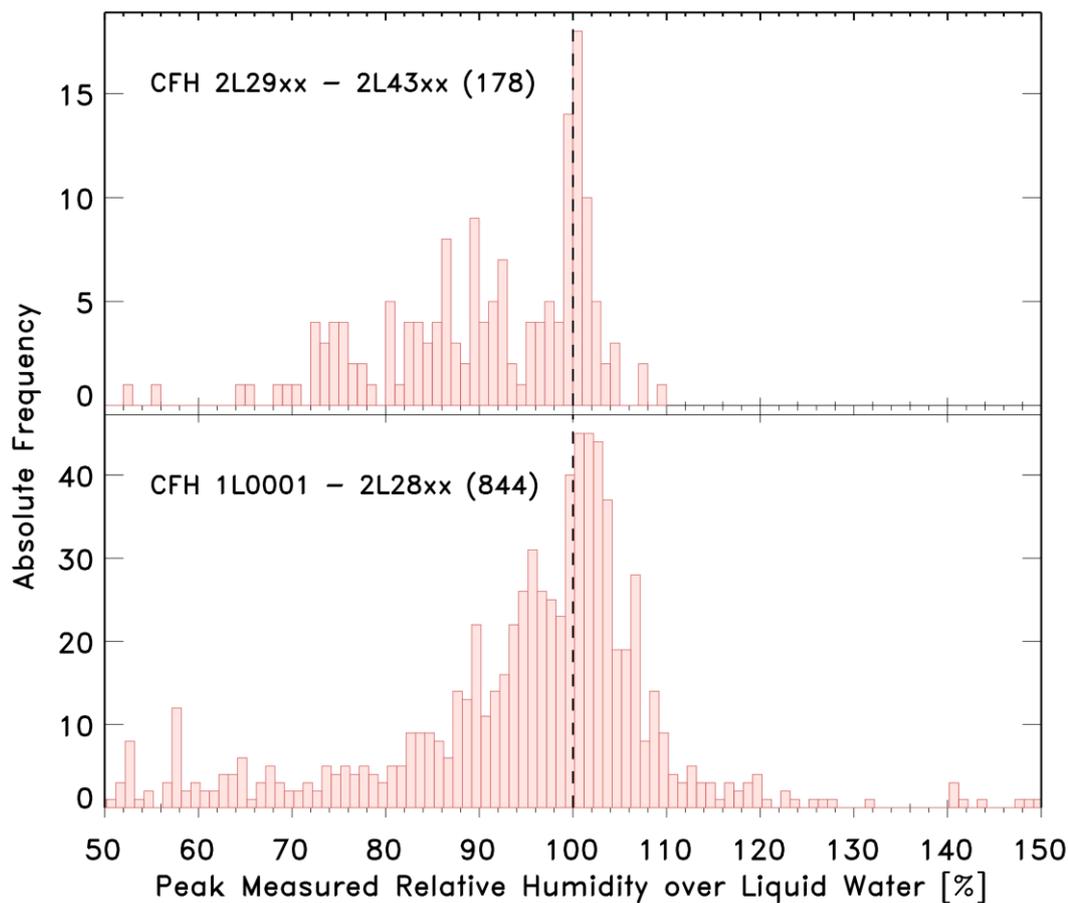


Figure 6: Comparisons frostpoint measurements in dual soundings. Stars indicate the tropospheric difference, diamonds indicate the stratospheric difference and dots indicate the mean difference for the entire profile.



5 **Figure 7: Frequency distribution of measured peak RH values in the lower troposphere. The bottom panel shows the distribution using the CFH instrument through serial number 2L28xx. The upper panel shows the current instrument series, which started with 2L29xx. Peak values above 100% are an indication for bias either in the temperature measurement or in the water vapor measurement. While some significant measurements of supersaturation were taken using instruments through 2L28xx, no significant supersaturation has been observed with instruments starting with 2L29xx.**

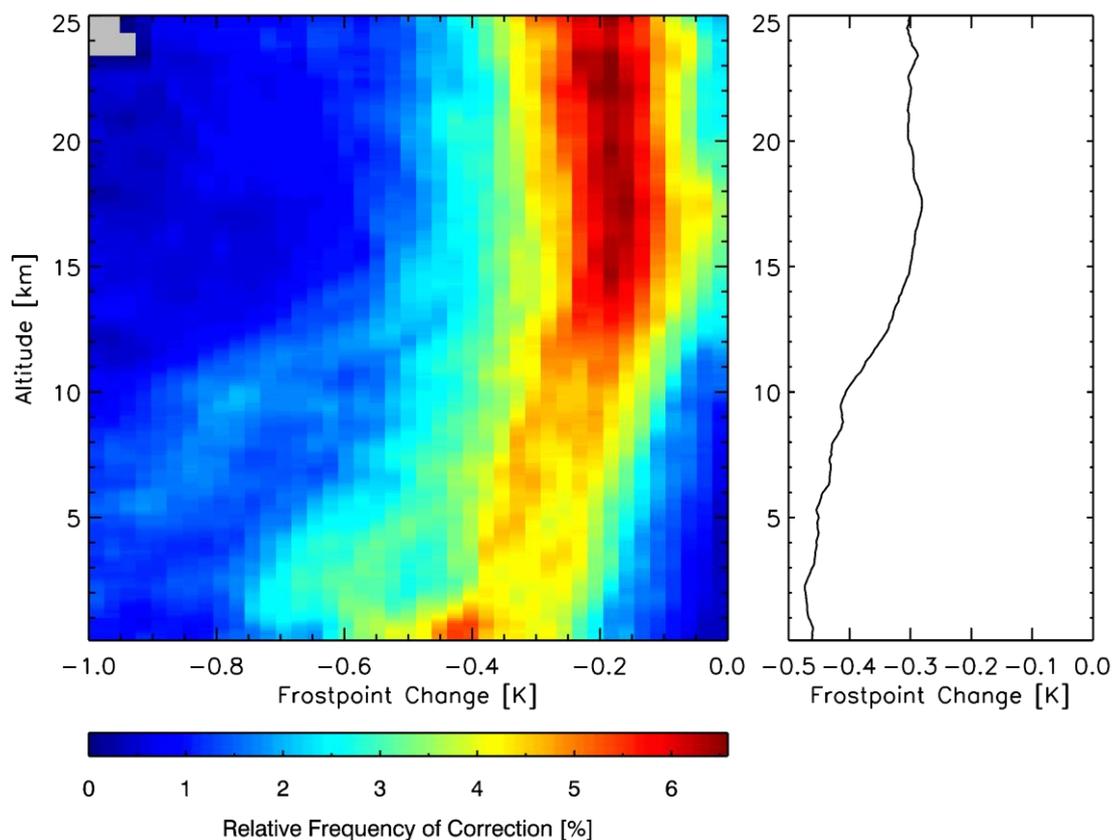


Figure 8: Deviation of the mean of five reference thermistors from the reference calibration temperature for each calibration run.

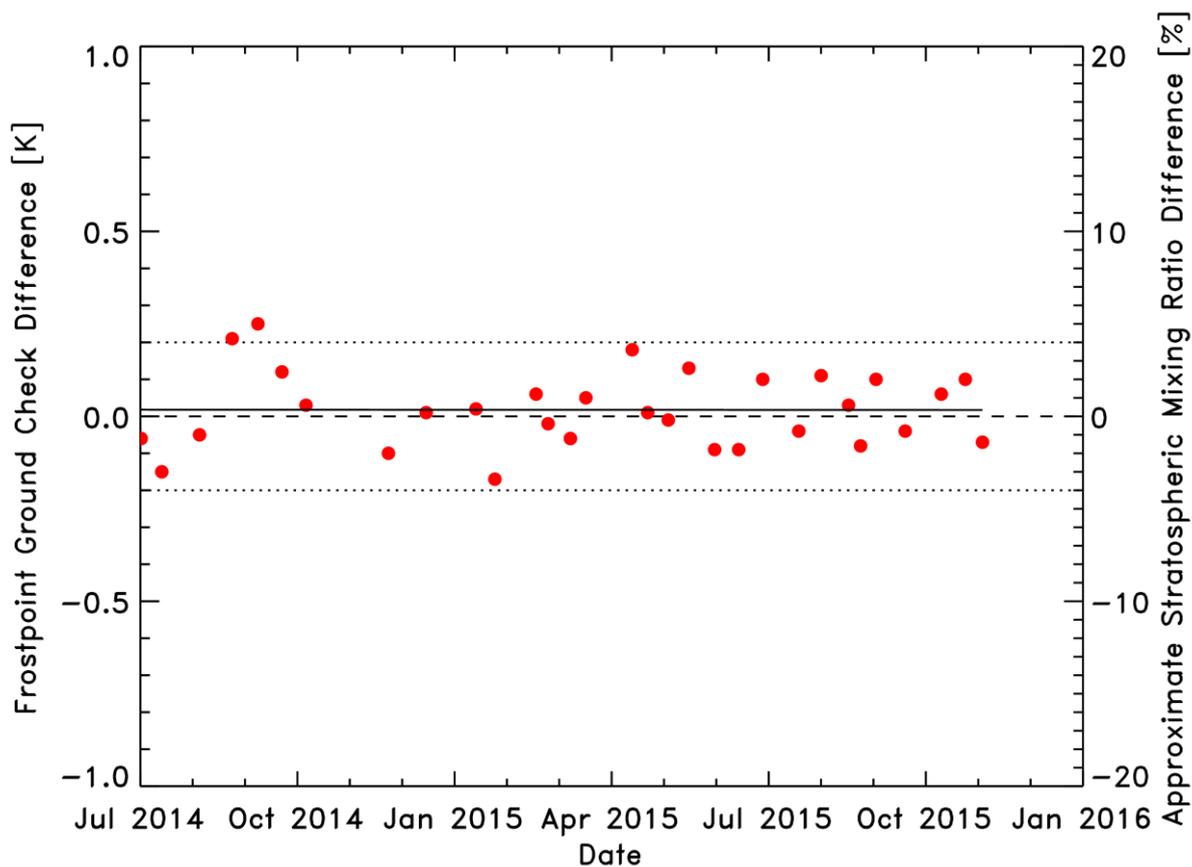


Figure 9: Difference of CFH dewpoint measurements prior to launch and a reference measurement inside the inlet tube. Dots indicate the individual measurements; the solid line indicates a linear fit through all data.

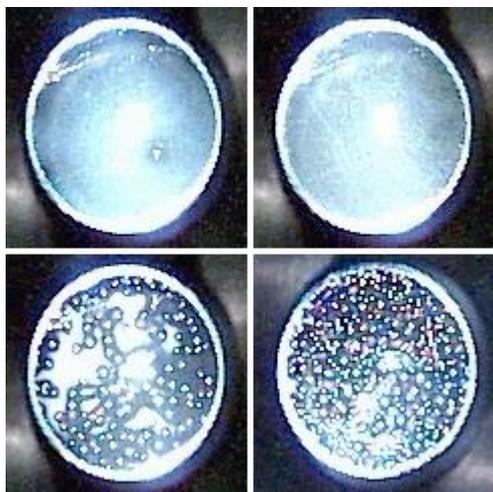


Figure 10: Images of the condensate layer at different conditions. The diameter of the mirror is 7 mm. Top left: supercooled liquid water at $T = -28.0^{\circ}\text{C}$. Top right: fine frost condensate at $T = -87.9^{\circ}\text{C}$. Bottom left: Coarse frost layer with patches of small liquid drops at $T = -26.8^{\circ}\text{C}$. Bottom right: Coarse frost layer at $T = -88.9^{\circ}\text{C}$