https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

Assessment of humidity ground checks for GRUAN-compliant automatic radiosonde operations

Gonzague Romanens[1], Giovanni Martucci[1], Marco Rosoldi[3], Yann Poltera[4,10], Cristina Cardós[2], Natalia Prats[2], Sunji Ambe[7], Gessica Cosimato[3], Antoine Farah[5], Christian Félix[1], Ilaria Gandolfi[3], Alexander Haefele[1], Miguel Hernández[2], Masami Iwabuchi[7], Patrick Jann[6], Jason Ming Chun Lam[8], Fabrizio Marra[3], José María Rodríguez[2], Yves-Alain Roulet[1], Yuta Tanaka[7], Ruud J. Dirksen[11], and Fabio Madonna[9]

[1]Federal Office of Meteorology and Climatology, MeteoSwiss, Payerne, Switzerland
[2]Agencia Estatal de Meteorologıa (AEMET), Tenerife, Spain
[3]Consiglio Nazionale delle Ricerche-Istituto di Metodologie per l'Analisi Ambientale (CNR-IMAA), Tito Scalo (Potenza), Italy
[4]Institute of Applied Physics, University of Bern, Bern, Switzerland
[5]Meteomodem, Ury, France
[6]Météo-France, Toulouse, France
[7]Japan Meteorological Agency (JMA), Tokyo, Japan
[8]Hong Kong Observatory, Hong Kong, China
[9]Department of Physics, University of Salerno, Salerno, Italy
[10]Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland
[11]GRUAN Lead Centre, Meteorologisches Observatorium Lindenberg (DWD),Tauche/Lindenberg, Germany

Correspondence: Gonzague Romanens (gonzague.romanens@meteoswiss.ch)

Abstract. Automatic Radiosonde Launchers (ARLs) are increasingly adopted for upper-air observations as they offer major operational advantages compared to manual launches, but they introduce new challenges for maintaining reference-quality humidity measurements within the GCOS Reference Upper-Air Network (GRUAN). In particular, the applicability of the pre-launch Standard Humidity Chamber (SHC) test, routinely used for manual radiosonde launches, has not been systematically assessed for radiosondes stored inside ARLs prior to launch.

This study investigates how residence time inside ARLs affects radiosonde humidity sensor's performance and the adequacy of SHC tests for automated GRUAN operations. A coordinated experiment was conducted at six sites using three radiosonde models (Vaisala RS41, Meisei IMS-100, and MeteoModem M10) and three ARL systems (Vaisala AS15, Meisei ARS and MeteoModem Robotsonde). Radiosondes were tested in an SHC prior to storage and again after residence times of 7 and 14 days. Relative humidity biases at saturation ($\mathrm{RH} = 100\%$) were analyzed using aggregated statistics, including a paired RH bias difference, i.e. changes in bias between pre-storage and post-storage measurements, to quantify the bias temporal drift. A non-parametric bootstrap resampling analysis was first applied to assess the robustness and confidence intervals of the paired RH bias difference estimates. Statistical significance of the paired RH bias difference distributions was subsequently evaluated using the Wilcoxon signed-rank test. A normalized storage-effect metric combining the systematic drift and the variability of the paired RH bias differences was compared with the GRUAN humidity uncertainty contribution associated with the ground-preparation.

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

Results indicate that radiosonde humidity performance remains stable for residence times up to 14 days, under the tested storage conditions, particularly for the RS41–AS15 configuration, where only small negative drifts were observed ($\sim -0.1\%$ RH week^{-1}). In contrast, the IMS-100–ARS configuration exhibited a pronounced increase in drift with storage time ($\sim -0.8\%$ RH week^{-1}), amplified under high-humidity and temperature seasonal conditions, while the M10–Robotsonde configuration showed a reversed drift with partial bias reduction during storage ($\sim +0.12\%$ RH week^{-1}). These findings demonstrate that radiosonde behavior during storage strongly depends on the combined radiosonde–ARL system.

Based on the experimental and statistical evidence, an operational upper limit of 14 days for radiosonde storage inside ARLs is recommended for GRUAN-compliant operations, with residence times up to 7 days providing optimal reliability. The study confirms the feasibility of implementing SHC tests for ARL operations, while highlighting the need for manufacturer and site-specific procedures to ensure long-term traceability and consistency of reference humidity measurements. The results demonstrate that residence time inside ARLs constitutes a system-dependent source of uncertainty, warranting explicit inclusion as an additional component in the GRUAN Data Product (GDP) humidity uncertainty budget.

1 Introduction

The GCOS Reference Upper-Air Network (GRUAN) is an international observing network designed to provide high-quality reference measurements of atmospheric essential climate variables through the troposphere and stratosphere (e.g. Seidel et al., 2009; Immler et al., 2010; Dirksen et al., 2014; Bodeker et al., 2016). Its primary objective is to ensure measurement traceability to SI standards and to deliver rigorous vertically resolved uncertainty estimates for all observations. This is achieved through well-documented metadata, measurement sensors, processing algorithms, standardized procedures, calibration protocols, and dedicated pre-launch tests such as that using a Standard Humidity Chamber (SHC), which assesses and validates humidity measurements of a Radiosonde (RS) under controlled conditions by a manufacturer-independent method.

The rapid development of automatic observing technologies, including Automatic Radiosonde Launchers (ARLs) systems, has enabled more frequent and cost-effective measurements, particularly during nighttime and in remote or harsh environments (Madonna et al., 2020). These advances address logistical challenges and reduce operational costs, leading to a growing adoption of ARL systems by meteorological services worldwide.

However, this transition also introduces several challenges related to data consistency, traceability, and operational harmonization. Ensuring consistency and homogeneity between historical manual observations and new automated measurements is essential for maintaining long-term Climate Data Records (CDRs) (Global Climate Observing System (GCOS), 2010; European Space Agency, 2026). GRUAN plays a key role in safeguarding data quality by providing reference-quality observations, and ensuring that automation does not compromise the integrity of climate datasets (GCOS, 2007). This is particularly important for climate research and for the calibration and validation (cal/val) of satellite and surface-based remote sensing systems.

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

While SHC tests are routinely performed for manual launches within GRUAN, standardized procedures for automatic launches remain insufficiently established. The need for performing a SHC procedure for ARLs is highlighted in Madonna et al. (2020). To enable the integration of ARLs into the GRUAN network in compliance with its requirements, a procedure must be defined and evaluated for implementing a manufacturer-independent Ground Check (GC), specifically the SHC, for these systems. To achieve this objective, a team of scientists from five GRUAN sites and the MeteoModem headquarters, assessed the impact of the residence time t that a RS spends inside the ARL on its ability to measure $\mathrm{RH} = 100\%$ within the SHC. This measurement is compared to a reference measurement performed prior to storage inside the ARL at loading time t_0. Based on the type of RH sensor adopted by the RS participating in this study, a prolonged residence time inside the ARL is expected to progressively degrade the humidity sensor response through cumulative exposure to environmental contaminants, temperature and humidity fluctuations, and storage-related aging effects. The participating sites are: Hong Kong Observatory (HKO), China; the Federal Office of Meteorology and Climatology MeteoSwiss, Payerne (PAY), Switzerland; the Institute of Methodologies for Environmental Analysis of the National Research Council (CNR-IMAA), Potenza (POT), Italy; the Japan Meteorological Agency, Tateno (TAT), Japan; the Meteorological State Agency (AEMET), Tenerife (TEN), Spain; the MeteoModem headquarter, Ury (URY), France. Of the six sites participating in the experiment, four sites operate a Vaisala RS41 RS (Jauhiainen et al., 2014; Survo et al., 2014a, b; Sommer et al., 2022) with the AS15 ARL, one operates the MeteoModem M10 RS (Dupont et al., 2020; Diarra et al., 2026) with the Robotsonde and one the Meisei IMS-100 RS with the ARS ARL (Arkan and Nababan, 2017; Kizu et al., 2018; Hoshino et al., 2022).

The description of the sites and their role within the framework of the experiment is provided in Section 2, followed by a description of the SHC procedures in Section 3. The design of the experiment and its limitations are presented in Section 4. A dedicated study investigating the impact of Manufacturer Ground-Check (MGC) reconditioning on the capability of the humidity sensor to recover its standard performance is presented in Section 5. Section 6 presents the statistical approach used to assess the establishment of the SHC test as a GRUAN-compliant independent ground-check for ARLs. The main findings from the statistical analysis are given for each site and their statistical significance is assessed in the context of the GDP humidity uncertainties associated with the ground-preparation component. Finally, Section 7 presents the conclusions and provides a set of recommendations for the use of SHC tests to meet GRUAN requirements in ARL systems.

2 Experimental framework and sites description

For the experiment, the residence time t was selected to range from 0 (loading time, t_0) to 14 days after loading (t_{14}), with an intermediate assessment time at 7 days (t_7). The choice of residence times is motivated by prevailing operational practices, whereby ARLs are typically loaded on a weekly basis and can accommodate between 12 and 24 RSs.

Assuming a typical launch frequency of two RSs per day, the maximum residence time of a RS inside the ARL does not generally exceed two weeks. The experiment was designed to evaluate, at each site, how residence time (at t_7 and t_{14}) and storage conditions affect the ability of the RS to measure $\mathrm{RH} = 100\%$ inside the SHC. The results are compared to a reference SHC test performed at t_0, before loading the RS into the ARL.

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

Table 1 summarizes the key characteristics of the six sites participating in the experiment. The sites span a wide geographical region, with TEN (28.32°N, 16.38°W) representing the westernmost location and TAT (36.06°N, 140.13°E) the easternmost. They also cover a broad latitudinal range, from URY (48.20°N, 2.36°E) as the northernmost site to HKO (22.31°N, 114.17°E) as the southernmost. These geographical differences translate into markedly diverse climatological conditions. As a consequence, RSs stored inside the ARLs may experience substantially different environmental conditions, including varying temperature and humidity diurnal and seasonal cycles, despite whether the ARL units are climate-controlled to mitigate ambient indoor or outdoor variations (fifth column). A climate-controlled (CC) ARL can feature air-conditioned (A/C), air dryer (AD) device or heating (H). Each site follows a distinct RS loading procedure that is summarized in the seventh column of Table 1. This procedure is defined by three parameters: the frequency with which the ARL is loaded, the number of ARL slots that are routinely filled after loading, and the maximum residence time t (in days) of a RS inside the ARL. For the URY site, the loading procedure is not operationally defined, as this site was activated solely for the purpose of this experiment and is normally used as a testbed for MeteoModem RSs.

The last two columns report the number of measurements performed at each site for the 7-day and 14-day experiments, respectively. The minimum requirement for each experiment is 20 RSs tested at t_7 and 20 RSs tested at t_{14}, while the target is 30 RSs. Deviations from the target sample size primarily result from limited time and logistical resources or from RSs failing the MGC. In addition, not all sites were able to fully allocate the available ARL capacity to achieve the target number of RSs for both the t_7 and t_{14} experiments.

Table 1. List of sites used in the experiment.

Site	GPS	RS	ARL	CC	SHC	Loading procedure†	$N_{\mathrm{RS}_{t_7}}$	$N_{\mathrm{RS}_{t_{14}}}$
HKO	22.31°N, 114.17°E	RS41-SG	AS15	YES (A/C, AD, H)	SHC1	weekly, 24, 11	20	20
PAY	46.81°N, 6.95°E	RS41-SG	AS15	YES (A/C, AD, H)	SHC1	weekly, 12, 8	30	30
POT	40.60°N, 15.72°E	RS41-SG	AS15	YES (A/C, AD, H)	SHC1	weekly, 2, 14	30	30
TEN	28.32°N, 16.38°W	RS41-SGP	AS15	YES (A/C, AD, H)	SHC1	weekly, 24, 12	30	30
TAT	36.06°N, 140.13°E	IMS-100	ARS	YES (A/C, AD)	RD-08-GC	4x weekly, 12, 14	24	28
URY	48.20°N, 2.36°E	M10	Robotsonde	YES (A/C,AD,H)	SHC1	not applicable	30	29

* HKO: the same 20 RS were used for t_7 and t_{14} experiments

* TAT: one year (≈ eight t_7 and t_{14} tests per season) instead of short-duration campaign, RH. Ref. = RD-08-GC + Rotronic-HC2 instead of SHC1

† loading procedure: how often the robot is loaded, how many slots are always filled after loading, maximum residence time in days of a sonde in the ARL.

Table 1 also shows which SHC type is in use at each site (sixth column), a description of the SHC is provided in the next section.

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

3 Standard Humidity Chamber - SHC

A standard humidity chamber for radiosonde testing is a controlled environmental device designed to expose humidity sensors to well-defined, known conditions. A commonly used well-defined condition is the thermodynamic equilibrium between water vapor and liquid water, corresponding to $\mathrm{RH} = 100\%$ at thermodynamic equilibrium (Hardy, 1998). The SHC has been used in several studies conducted within the framework of GRUAN (Immler et al., 2010; Dirksen et al., 2014; Sairanen et al., 2014, 2015; Lee et al., 2019; Dirksen et al., 2020). Through the last two decades, the SHC became a standard to assess the performance of a manually launched RS and its ability to measure the reference value of $\mathrm{RH} = 100\%$. In particular, during the latest World Meteorological Organization (WMO) Upper-Air Instrument Intercomparison, the SHC was extensively employed to characterize the laboratory and pre-flight performance of the humidity sensors from ten different RS models (Dirksen et al., 2024). The SHC chamber consists of a sealed enclosure containing a water solution with which the water vapor establishes thermodynamic equilibrium. For pure water, the resulting relative humidity of the air inside the SHC corresponds exactly to $\mathrm{RH} = 100\%$. Air circulation within the chamber ensures uniform conditions, while in certain setups high-accuracy reference instruments (traceable hygrometers and thermometers) provide the baseline against which the RS measurements are compared. The RS sensors are placed inside the chamber, by means of an adapter, often mounted in a way that mimics operational airflow conditions, and are allowed to equilibrate before measurements are recorded.

The importance of this test lies in ensuring the reliability and traceability of humidity measurements in the atmosphere. RS data are widely used in weather forecasting, climate monitoring, and numerical weather prediction models. Testing at $\mathrm{RH} = 100\%$ is especially critical because this condition represents the upper limit of atmospheric humidity and is frequently encountered in clouds, fog, and precipitation regions. At such high humidity levels, humidity sensors are particularly susceptible to saturation effects, condensation, and response-time limitations (Miloshevich et al., 2009; World Meteorological Organization (WMO), 2024). A standard humidity chamber allows operators to verify that the radiosonde can accurately reach and report values close to 100% RH without significant bias or drift (i.e., evolution of the bias over time). It also helps identify issues such as sensor contamination, aging, or manufacturing inconsistencies that could compromise performance. By systematically validating RS performance under saturated conditions, the humidity chamber test contributes to maintaining data quality, inter-comparability between stations, and long-term consistency of atmospheric observations. Within GRUAN, and as an example of application to the RS41 RS data, the SHC measurements are integrated into the processing and uncertainty characterization framework of the GDP as described in section 3.2.2.4 of Sommer et al. (2023a).

4 Design of the experiment

The heterogeneity of ARLs and RS models within the GRUAN network results in significant differences in the loading procedures across sites. These differences include the sequencing of the MGC and SHC tests, the duration of storage inside the ARL, and the environmental conditions under which storage occurs. The participating ARL systems operate under diverse climate-controlled temperature and humidity conditions and may perform MGC at different time (e.g., during the loading phase or immediately prior to launch). More generally, the deployment of a ARL is often driven by limited staffing resources and/or

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

the need to operate in remote locations. Consequently, the introduction of additional manufacturer-independent tests, such as the SHC, can pose operational challenges for routine activities. Procedural differences and site dependent capabilities have to be taken into account when implementing a procedure to use the SHC. For instance, at the TAT site, the GRUAN procedure for the ARS is a complete ground-check sequence comprising the MGC, followed by SHC-0% and subsequently SHC-100% completed by a final MGC. In contrast, other GRUAN sites, such as those operating the Vaisala Vaisala Autolauncher (AS15), do not currently follow a standardized SHC procedure, nor is there an established method to integrate SHC data into the generation of the RS GRUAN Data Product (GDP). Nevertheless, the SHC serves as the reference system regardless of the specific model or implementation (see Table 1), as it defines the reference value $\mathrm{RH} = 100\%$. This study ultimately aims to define a GRUAN-compliant procedure for implementing the SHC during automatic flights, thereby compensating for the identified methodological inhomogeneities between manual and automated launches within the network. As a first step toward establishing such a procedure, the impact of the residence time t spent by a RS inside the ARL is assessed with respect to its ability to accurately measure $\mathrm{RH} = 100\%$.

The experimental procedure is summarized in the diagram shown in Figure 1. The experiment consists of a sequence of preparation, storage, and post-storage checks. First, all RSs undergo an initial MGC, after which only the qualifying units proceed to the $\mathrm{RH} = 100\%$ test in the SHC. The RSs are then loaded into the ARL at t_0, i.e. $N_{\mathrm{RS}_{t_7}}$ radiosondes are loaded at $t_{0,7}$ and $N_{\mathrm{RS}_{t_{14}}}$ at $t_{0,14}$.

Subsets of RSs are subsequently retrieved after defined residence times: the 7-day group is unloaded at t_7, and the 14-day group at t_{14}. Following removal from the ARL, all RSs undergo a final round of MGC and SHC tests to assess potential performance degradation.

The diagram represents the experimental procedure adopted at all stations except HKO. At HKO, all RSs are loaded into the ARL simultaneously at t_0 and unloaded at t_7, at which time the MGC and SHC tests are performed. Subsequently, all RSs are reloaded into the ARL and unloaded again at t_{14} to perform the final MGC and SHC tests. The experiments were conducted over a period of approximately two months at all sites, with the exception of TAT. In contrast to the other sites, TAT provided a substantially different dataset in terms of temporal coverage, extending over nearly eight months, from January to September 2025. This extended duration was specifically designed to monitor and quantify potential seasonal effects on the performance of the RS humidity sensors.

We define here the RH bias of a RS with respect to the SHC at time t, as $\Delta\mathrm{RH}_t$, with $t \in \{t_{0,7}, t_{0,14}, t_7, t_{14}\}$. As shown in Figure 1, $N_{\mathrm{RS}_{t_7}}$ radiosondes at time $t_{0,7}$ and $N_{\mathrm{RS}_{t_{14}}}$ radiosondes at $t_{0,14}$, undergo the SHC test after completion of the MGC. This yields a total of $N_{\mathrm{RS}_{t_7}} + N_{\mathrm{RS}_{t_{14}}}$ initial bias measurements $\Delta\mathrm{RH}_{t_{0,7}}$ and $\Delta\mathrm{RH}_{t_{0,14}}$ obtained prior to loading the RSs into the ARL. The combined dataset $[\Delta\mathrm{RH}_{t_{0,7}} \cup \Delta\mathrm{RH}_{t_{0,14}}]$ exhibits a statistical spread that is characterized by its standard deviation, denoted as $\Omega_{\Delta\mathrm{RH}_{t_0}}$. The RH interval $100\% \pm \Omega_{\Delta\mathrm{RH}_{t_0}}$ therefore represents the minimum inherent variability of the $\mathrm{RH} = 100\%$ measurement obtained by a RS inside the SHC at $t > t_{0,7}$ and $t > t_{0,14}$.

The change in bias between pre-storage and after-storage time is defined as $\Delta(\Delta\mathrm{RH})_{t_7} = \Delta\mathrm{RH}_{t_7} - \Delta\mathrm{RH}_{t_{0,7}}$ (for residence time of 7 days), and $\Delta(\Delta\mathrm{RH})_{t_{14}} = \Delta\mathrm{RH}_{t_{14}} - \Delta\mathrm{RH}_{t_{0,14}}$ (for residence time of 14 days). These bias differences, hereafter referred to as paired RH bias difference, quantify the impact of the residence time t (attributable to storage, handling and

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

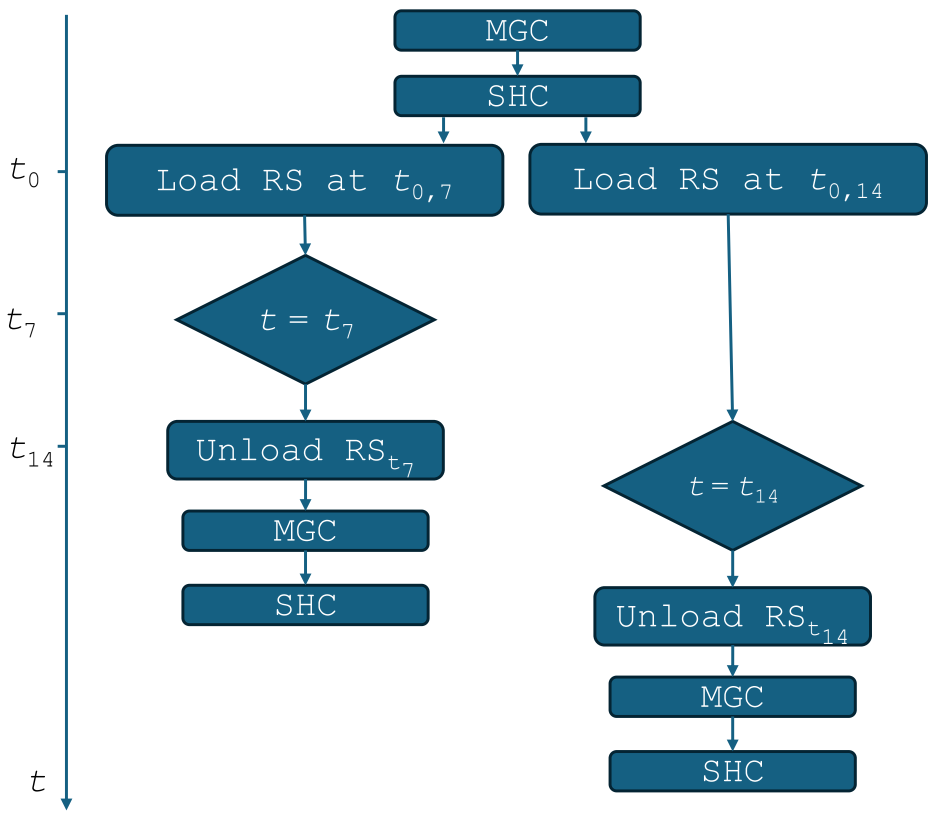

Figure 1. Experimental procedure flowchart. The left and right branches show the progression from the initial MGC and SHC tests at $t_{0,7}$ and $t_{0,14}$ to the final tests at t_7 and t_{14}.

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

environmental conditions effects) on the ability of the RS to accurately measure $\mathrm{RH} = 100\%$ at t_7 and t_{14}. The spread $\Omega_{\Delta \mathrm{RH}_{t_0}}$ corresponds then to the intrinsic repeatability of the SHC measurement system under reference conditions and provides a practical criterion to assess whether subsequent paired RH bias differences remain compatible with the expected variability of the SHC measurements.

4.1 Limitations of the experimental design

The experiment evaluates RS residence time under realistic site conditions, meaning that cross-site variations encompass not only storage duration, but also distinct environmental, configurational, and operational factors. It should be noted that the $t_{0,7}$ and $t_{0,14}$ tests were conducted on different days. Consequently, differences in environmental or operational conditions may have contributed to the observed variations, potentially affecting the reproducibility and direct comparability of the results. Also, the participating sites span a wide range of climatological regimes, and the ARLs differ in terms of internal climate control, with a different exposure to ambient indoor or outdoor conditions. Table 2 summarizes the mean values and standard deviations of the annual cycles of T and RH climatological conditions inside the ARLs at the different sites.

Based on the information provided in Table 1, all Vaisala sites are equipped with the full climate-controlled suite inside the AS15. This is reflected in the standard deviations of T and RH (shown in bold in Table 2), which are comparable across the Vaisala sites. In contrast, these values are nearly twice as large for the Meisei site despite the A/C and dry air injection inside the storage unit used to reduce the relative humidity. The strong variability of T and RH inside the ARS at TAT may therefore contribute to the larger paired RH bias differences discussed further on in this study. The MeteoModem URY site exhibits a temperature annual cycle with a similar variability as for the Vaisala sites, but with a humidity variability more aligned with the TAT site.

Table 2. Mean values and standard deviations of the annual cycles of T and RH climatological conditions inside the ARLs at the different sites.

Site	$\overline{T} \pm \sigma_T$ [K]	$\overline{\mathrm{RH}} \pm \sigma_{\mathrm{RH}}$ [%]
HKO	28.6 $\pm\mathbf{2.3}$	32.4 $\pm\mathbf{5.8}$
PAY	21.6 $\pm\mathbf{4.2}$	36.9 $\pm\mathbf{5.8}$
POT	23.8 $\pm\mathbf{4.1}$	36.9 $\pm\mathbf{9.8}$
TEN	22.1 $\pm\mathbf{1.5}$	36.4 $\pm\mathbf{7.5}$
TAT	18.4 $\pm\mathbf{7.5}$	53.0 $\pm\mathbf{17.5}$
URY	23.0 $\pm\mathbf{3.5}$	41.5 $\pm\mathbf{17.4}$

Instead of using these variations as categorization criteria, this study evaluates the impact of residence time on both RS humidity performance and SHC test representativeness under the diverse operational conditions of the GRUAN network. As a consequence, the observed paired RH bias differences shall be interpreted as system-dependent operational responses resulting

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

from the combined influence of residence time, environmental exposure, and handling procedures, rather than as isolated causal effects of residence time alone. Finally, the fact that the analysis is based on the data from four Vaisala sites and one Meisei and MeteoModem site represents an unbalanced contribution to the study. However, it is representative of the current RS presence across the GRUAN network (www.gruan.org, last access 21.05.2026).

5 Radiosondes and manufacturer ground-check

Table 3 summarizes the main characteristics of the three radiosonde models included in the experiment: Vaisala RS41, Meisei IMS-100, and MeteoModem M10. All three radiosondes employ capacitive thin-film polymer humidity sensors, but they differ in sensor integration, reconditioning capability, recalibration strategy, and Quality Controlled (QC) criteria. The Vaisala RS41 is equipped with a capacitive thin-film polymer RH sensor with heating capability combined with a platinum resistance temperature sensor (Pt1000) (Vaisala Oyj, 2017). This model includes an active reconditioning process in which the RH sensor is heated to remove contamination deposits, and recalibration is applied only to the humidity sensor. Default QC thresholds are more stringent for RH than for temperature, with radiosondes rejected if RH errors exceed $\pm 2\%$ in dry conditions or if temperature errors exceed $\pm 3\,\mathrm{K}$. The Meisei IMS-100 radiosonde combines a capacitive thin-film polymer RH sensor with a glass bead thermistor for temperature measurements (Kizu et al., 2018). Unlike the RS41, the IMS-100 does not implement sensor reconditioning, and RH sensor recalibration is applied only later when generating the GDP. The default QC rejection thresholds are $\pm 5\%$ for RH and $\pm 0.5\,\mathrm{K}$ for temperature. The MeteoModem M10 radiosonde integrates a capacitive thin-film polymer RH sensor with an integrated temperature sensor (MeteoModem, 2023; Dupont et al., 2020). No sensor reconditioning is performed, and recalibration is applied only to the temperature sensor. The QC rejection limits for the M10 are $\pm 5\%$ for RH and $\pm$2K for temperature.

5.1 Vaisala RS41

For the Vaisala RS41–AS15 configuration, there is not yet an established GRUAN procedure that includes the use of a SHC. The Vaisala RS41 Manufacturer Ground-Check (MGC) procedure (Vaisala Oyj, 2016) comprises several sequential checks (rejection criteria are reported under the column “default QC” in the first row of Table 3). For the temperature sensor GC, a short comparison (approximately 10–15 seconds) is performed between the main temperature sensor and the temperature sensor integrated on the humidity sensor chip; the temperature calibration is not updated following the MGC.

For the humidity sensor GC, the procedure begins with reconditioning by heating the sensor for approximately 180 seconds, followed by a verification of the RH value at the “physical zero-RH” while the sensor is cooled back to ambient air temperature (approximately 60 seconds), from which a correction value $\delta\mathrm{RH}_{\mathrm{corr}}$ is derived. After this, the manufacturer-provided RH calibration coefficient is updated using the derived $\delta\mathrm{RH}_{\mathrm{corr}}$ value (Vaisala Oyj, 2016; Sommer et al., 2023b). The correction $\delta\mathrm{RH}_{\mathrm{corr}}$ is applied across the full RH range (0–100%RH) using scaling factors to $\delta\mathrm{RH}_{\mathrm{corr}}$ that vary from 1 in dry conditions to 2.5 in saturated conditions, respectively (Vaisala, private communication, presented at the 16th Implementation and Coordination Meeting, Tenerife, Spain, 2025).

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

Table 3. List of radiosondes used in the experiment.

RS model	RH sensor	T sensor	reconditioning	recalibration	default QC
Vaisala RS41	capacitive thin film polymer + T (pt1000)	Platinum resistor pt1000	YES (heated RH sensor)	YES (RH sensor)	$> \pm 2\%$ RH rejected; $> \pm 3\,\mathrm{K}$ rejected, $>$ 3 hPa rejected
Meisei IMS-100	capacitive thin film polymer + T sensor	Glass bead thermistor	NO	YES (RH sensor)	$> \pm 5\%$ RH rejected; $> \pm 0.5\,\mathrm{K}$ rejected
Meteomodem M10	capacitive thin film polymer + T sensor	Glass bead thermistor	NO	YES (T sensor)	$> \pm 5\%$ RH rejected; $> \pm 2$ K rejected

Finally, the pressure sensor GC lasts about 30–60 seconds for RS41-SGP and RS41-D models; prior to launch, the pressure calibration is updated through comparison with a reference barometer (Vaisala Oyj, 2016).

5.2 Meisei iMS-100

For the Meisei IMS-100–ARS configuration, a standard GRUAN procedure that includes the use of the SHC has been published by Kizu et al. (2018). The IMS-100 RS undergoes a sequence of MGC + SHC + MGC prior to launch. An initial MGC is performed more than one day before launch under room temperature and RH conditions inside a ventilated container, where temperature and RH measurements are compared against the reference sensor Rotronic HC2-A probe. This step determines RS acceptance, with criteria of $\pm 5\%$ for RH and $\pm 0.5\,\mathrm{K}$ for temperature; no temperature recalibration is applied.

Following a successful initial MGC, the SHC test is conducted more than 24 hours before launch. The humidity sensor is sequentially exposed to approximately 0% RH (molecular sieve, $< 1\%$ RH, 7 minutes) and 100% RH (humidified sponge, 10 minutes) in ventilated containers, with RH biases derived from measurements. Afterward, the RS is stored in a low-humidity chamber for more than 24 hours.

Immediately prior to loading the IMS-100 into the ARS, to ensure RS acceptance, a second MGC is carried out under room conditions. No recalibration of the sensor is applied following this step. The RH deviations obtained from the first MGC and the two SHC reference points (0% and 100% RH) are subsequently used for humidity sensor recalibration in the GDP (Kizu et al., 2018, Section 2.4).

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

5.3 Meteomodem M10

For the Meteomodem M10–Robotsonde configuration, a standard GRUAN procedure that includes the use of the SHC is currently under review by the GRUAN Lead Center, (personal communication by Antoine Farah, MeteoModem). Prior to launching the M10 RS, a ground preparation procedure is performed to configure, verify, and validate the system. This includes setting the transmission frequency, verifying key system parameters using the preparation software, and checking battery voltage, GPS 3D acquisition, and sensor data accuracy.

The MGC relies on factory calibration for temperature and relative humidity using a Rotronic HC2-A probe, with recalibration required every three years at Meteomodem. Measurements are conducted within a ground-check device consisting of a plastic enclosure that houses the radiosonde alongside reference temperature and humidity sensors operating under ambient conditions, as well as a GNSS signal transmitter and an infrared transmitter.

During the measurement phase, the enclosure is closed and a fan ensures homogeneous air conditions. The radiosonde sensors, positioned near the reference sensors, are allowed to stabilize before comparison. The resulting deviations in temperature and relative humidity are evaluated against predefined thresholds listed in Table 3, which determine whether a correction to the thermistor is applied and whether the radiosonde is validated or rejected.

5.4 Impact of Vaisala MGC on the humidity sensor

In addition to the main experiment, at the PAY site, three dedicated experiments were conducted with the objective of investigating the detailed functioning and the impact of the reconditioning of the RH sensor performed by the Vaisala MGC on the performance of the humidity sensor itself.

The first experiment aimed at assessing the role and impact of 10 consecutive MGCs on the capability to recondition the humidity sensor. The second experiment pursued the same objective, but included an immediately subsequent SHC test after each MGC in order to assess the performance of the reconditioned humidity sensor inside the SHC. The third experiment focused on evaluating the capability of the MGC to recondition the humidity sensor after short (2 minutes) to long (150 minutes) exposure to different Volatile Organic Compounds (VOCs) contaminants.

The first experiment, referred to as "sequential MGC", consists of repeated calibration cycles on two Vaisala RS41 radiosondes (RS_1 and RS_2). Two sets of MGC devices, each connected to a dedicated computer, are used in parallel to perform 10 consecutive MGCs on each radiosonde, with the correction value $\delta\mathrm{RH}_{\mathrm{corr}}$ recorded after every cycle.

The second experiment (performed on a different day), denoted as "sequential MGC + SHC", extends this approach by alternating calibration and high-humidity testing. After preparing the MGC devices and stabilizing the ventilated SHC, the radiosondes RS_3 and RS_4 undergo 10 successive cycles combining MGC and SHC. Following each MGC, the correction value $\delta\mathrm{RH}_{\mathrm{corr}}$ is stored, and the radiosondes are alternately placed in the SHC. During each cycle, their humidity biases $\Delta\mathrm{RH}_{\mathrm{SHC}}$ relative to the $\mathrm{RH} = 100\%$ reference are recorded.

Figure 2 shows the results of the first two experiments in terms of the correction values $\delta\mathrm{RH}_{\mathrm{corr}}$ for RS_1 (blue squares) and RS_2 (yellow squares) as a function of time. In the left panel ("sequential MGC"), both RS41 radiosondes, RS_1 and RS_2,

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

exhibit a nearly linear decrease of $\delta\mathrm{RH}_{\mathrm{corr}}$ with increasing cycle number. This behavior indicates a cleansing effect of the MGC on the humidity sensor: with each successive MGC, contaminating material is progressively removed from the sensor surface, resulting in progressively smaller correction values $\delta\mathrm{RH}_{\mathrm{corr}}$, i.e. the humidity sensors require less correction at each cycle. The right panel of Figure 2 shows the results of the second experiment ("sequential MGC + SHC"). Again, both RS_3 and RS_4 display a decreasing trend in $\delta\mathrm{RH}_{\mathrm{corr}}$ over the 10 MGCs. In contrast to the first experiment, and coincidentally, the $\delta\mathrm{RH}_{\mathrm{corr}}$ values of RS_3 and RS_4 exhibit no relative offset and align along the same negative trend. The cleansing effect of the MGC is also clearly reflected in the SHC results: the corresponding $\Delta\mathrm{RH}_{\mathrm{SHC}}$ values for both radiosondes are distributed around the reference value $\mathrm{RH} = 100\%$, without any systematic trend beyond measurement noise. This demonstrates that, after each MGC, the humidity sensor is effectively reconditioned and retains its ability to accurately measure the reference value $\mathrm{RH} = 100\%$ inside the SHC, with no observable degradation in performance over successive cycles.

The third experiment, denoted as "sequential MGC + SHC with humidity sensor contamination", investigates the impact of controlled contamination on sensor performance. Two Vaisala MGCs systems connected to dedicated computers are used to record the correction value $\delta\mathrm{RH}_{\mathrm{corr}}$, and a ventilated SHC is prepared and stabilized prior to testing.

The same two radiosondes of the first experiment, RS_1 and RS_2, undergo, respectively, 2 and 3 successive cycles combining in quick succession (i) contamination of the sensor, (ii) SHC (contaminated measurement), (iii) MGC (reconditioning), and (iv) SHC (decontaminated measurement). The contamination is applied directly to the humidity sensors using three approaches: exposure to Volatile Organic Compounds (VOCs) from commercial skin care creams at close range for varying durations (100, 150, and 15 minutes; cases "cont1", "cont2", and "cont4", respectively), exposure to tobacco smoke for 2 minutes ("cont3"), and exposure to ambient air for 110 minutes ("cont5").

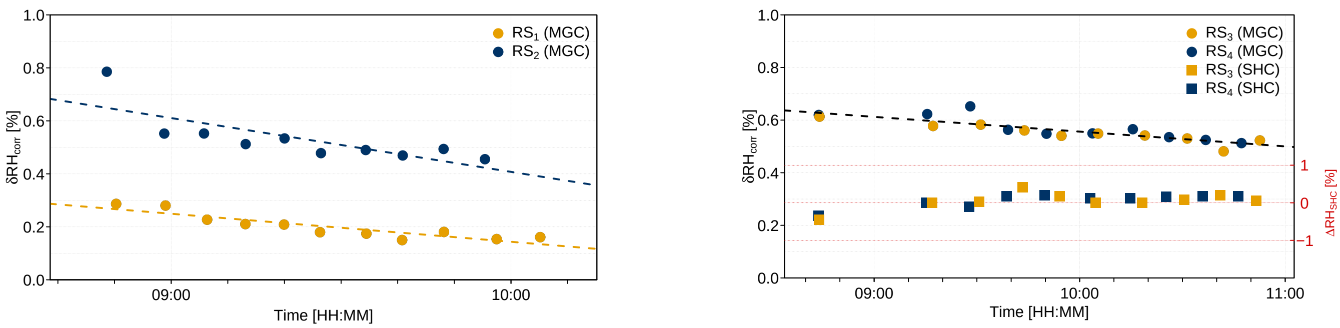

Figure 2. Left: 10 consecutive MGC cycles and corresponding $\delta\mathrm{RH}_{\mathrm{corr}}$ values for RS_1 (blue) and RS_2 (yellow) with linear fit. Right: 10 consecutive MGC + SHC cycles and corresponding $\delta\mathrm{RH}_{\mathrm{corr}}$ values for RS_3 (blue circles) and RS_4 (yellow circles) with linear fit on top part. On bottom part, the corresponding $\Delta\mathrm{RH}_{\mathrm{SHC}}$ for RS_1 (blue squares) and RS_2 (yellow squares) obtained from the SHC test. The $[-1, 1]$ $\Delta\mathrm{RH}_{\mathrm{SHC}}$ is shown in red on the right axis.

Figure 3 shows the results of the third experiment, "sequential MGC + SHC with humidity sensor contamination". The time series displays the RH values measured by the two RS41 radiosondes, RS_1 and RS_2, inside the SHC following intentional contamination of their humidity sensors by the three aforementioned contaminants.

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

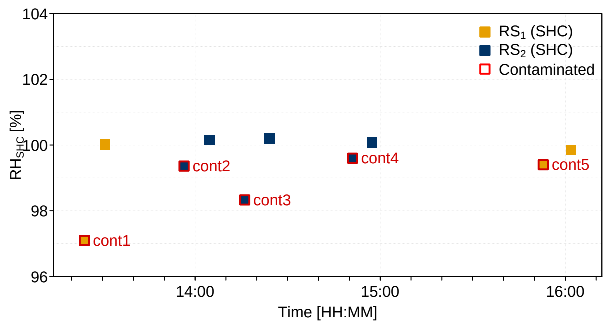

Figure 3. $\mathrm{RH_{SHC}}$ measured in the SHC as a function of time during the sequential contamination – MGC – SHC. Five contamination conditions: VOC exposure of different durations (cont1, cont2, cont4), cigarette smoke (cont3), and ambient air (cont5). After each contamination and subsequent MGC recalibration, $\mathrm{RS_1}$ and $\mathrm{RS_2}$ are alternately measured in the SHC. Blue and yellow markers indicate reference-stable measurements near $\mathrm{RH} = 100\%$, while red-contoured markers highlight deviations associated with each contamination type, and illustrate the impact of contamination on humidity measurement bias $\Delta\mathrm{RH_{SHC}}$.

The $\Delta\mathrm{RH_{SHC}}$ values measured by $\mathrm{RS_1}$ and $\mathrm{RS_2}$ immediately after contamination and prior to reconditioning are substantially larger than the uncontaminated measurements. For example, the $\Delta\mathrm{RH_{SHC}}$ value obtained by $\mathrm{RS_1}$ after contamination "cont1" by VOC exposure for 100 minutes reaches nearly $3\,\%\,\mathrm{RH}$.

Following contamination, $\mathrm{RS_1}$ undergoes reconditioning by the MGC, which exhibits a clear cleansing effect. This effect is reflected in the subsequent SHC test, which yields a $\Delta\mathrm{RH_{SHC}}$ value close to zero, indicating that the humidity sensor is restored to its nominal performance. The same pattern is observed for the other contaminants, i.e. tobacco smoke and ambient air, although with a smaller effect on the humidity sensor. Commercial VOC products can therefore induce different levels of contamination. For example, the "cont2" case, despite a longer exposure time (150 minutes) than "cont1" (100 minutes), produced a smaller effect on the RH measurements inside the SHC. This difference is attributed to the use of different creams in the two experiments. The deterioration of the RS41 humidity sensor when exposed to VOC contamination arises from the occupation of $\mathrm{H_2O}$ bonding sites on the thin-film polymer sensor by VOC molecules. This reduces the number of $\mathrm{H_2O}$ molecules that can bind to the polymer, leading to a lower capacitance and consequently a lower RH output (i.e. larger $\Delta\mathrm{RH_{SHC}}$) (Vaisala Oyj, 2020, 2012). The nominal performance of the humidity sensor is restored following the subsequent MGC.

These results provide a basis for interpreting the data from the Vaisala site HKO. This site is unique in that the same set of RSs was used throughout the entire experiment; consequently, the RSs reaching $t = t_{14}$ underwent one additional MGC+SHC iteration compared to those at the other sites. Based on the findings in Sect. 5.4, the RH bias of the RSs at HKO after t_{14} is therefore expected to be smaller than in the case of a single MGC (as applied at the other Vaisala sites), owing to the additional cleansing effect associated with the reconditioning of the humidity sensor.

The experiment evaluating different contaminants was designed to simulate a residence time of the RS inside the ARL, extending well beyond t_{14}. The results suggest that, for long residence times combined with uncontrolled environmental

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

conditions, which lead to sensor contamination, the RH measurements obtained in the free atmosphere after launch may no longer comply with GRUAN measurements quality requirements. To provide a realistic example, the RS sensor boom is exposed during the time t to talcum powder released by the balloon envelop inside the ARL. As the third experiment has proven, any exposure to contaminants without reconditioning reduces the capability of the RH sensor to accurately measure $\mathbf{RH} = 100\%$.

6 Analysis and Results

The results of the dedicated tests presented in Section 5 provide useful context for interpreting the evolution of the SHC measurements discussed hereafter. Figure 4 presents the statistical results for all sites. The left panel shows box plots of the bias distributions $\Delta\mathrm{RH}_t$ for $t \in \{t_{0,7}, t_{0,14}, t_7, t_{14}\}$. The median values illustrate the central tendency of the biases at each experimental stage. The boxes have vertical width corresponding to the inter-quartile range ($25^{\mathrm{th}} - 75^{\mathrm{th}}$ percentile), and whiskers extend to the most extreme values within 1.5 times the inter-quartile range.

The two right panels show the corresponding distributions of paired RH bias difference, $\Delta(\Delta\mathrm{RH})_{t_7}$ and $\Delta(\Delta\mathrm{RH})_{t_{14}}$. These quantities represent the statistical drift of $\Delta\mathrm{RH}_t$ from $t_{0,7}$ to t_7 and from $t_{0,14}$ to t_{14}, respectively.

The vertical zero line in the paired RH bias difference plots indicates the absence of systematic drift in the bias during the residence time inside the ARL. The gray shaded areas represent the repeatability interval $\pm\Omega_{\Delta\mathrm{RH}_{t_0}}$ around the zero-line and indicate the expected variability of the measurement.

The error bars associated with the mean values of $\Delta(\Delta\mathrm{RH})_{t_7}$ and $\Delta(\Delta\mathrm{RH})_{t_{14}}$ correspond to the standard deviation of the paired RH bias difference distributions at t_7 and t_{14}, respectively. Individual paired RH bias difference values are shown as black crosses.

6.1 Sites results

6.1.1 Hong Kong Observatory (HKO)

The first row of Figure 4 shows the results of the experiment performed at the HKO site using the Vaisala AS15 ARL. As described in Section 4, the RSs at HKO were loaded into the ARL at t_0 and subsequently unloaded at t_7 and t_{14} to perform the MGC and SHC tests.

The narrow inter-quartile range of the distribution at t_0 indicates that the repeatability of the SHC test for the $N_{\mathrm{RS}_{t_0}}$ RSs is high and characterized by a very small spread. A negative trend in the $\Delta\mathrm{RH}_t$ values is observed over $t \in \{t_0, t_7, t_{14}\}$, with the mean (and median) bias at t_{14} being the most negative among the three measurement times. The tendency for the bias to increase in magnitude with residence time inside the ARL is further shown in the paired RH bias difference plots. In particular, the mean value of $\Delta(\Delta\mathrm{RH})_{t_{14}}$ falls outside the repeatability interval (gray shaded region), i.e. $\Delta(\Delta\mathrm{RH})_{t_{14}} < -\Omega_{\Delta\mathrm{RH}_{t_0}}$, indicating a systematic negative drift after 14 days of residence time exceeding the repeatability interval $\pm\Omega_{\Delta\mathrm{RH}_{t_0}}$. This value is accentuated by the fact that the RS at HKO underwent one additional MGC+SHC iteration at t_7 compared to the other sites.

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

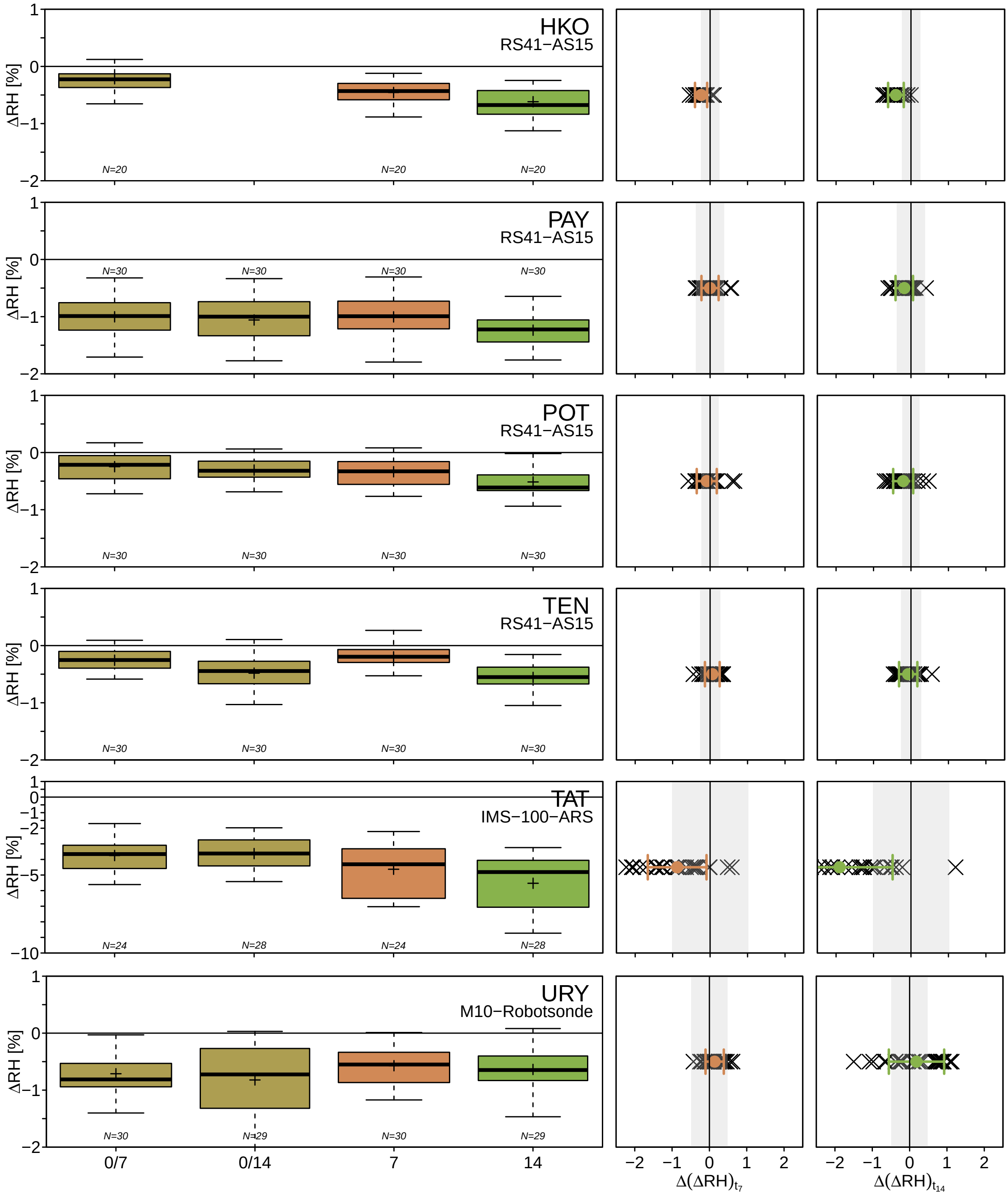

Figure 4. For the six sites, each row shows box plots of RH bias distributions $\Delta\mathrm{RH}_t$ at the four stages $t_{0,7}$, $t_{0,14}$, t_7, and t_{14} (left). Boxes indicate the inter-quartile range and horizontal lines the median. The two right panels show paired RH bias differences $\Delta(\Delta\mathrm{RH})_{t_7}$ and $\Delta(\Delta\mathrm{RH})_{t_{14}}$, representing changes after 7 (orange) and 14 days (green) in the ARL. Crosses denote individual measurements and filled circles the mean. The vertical zero line marks no drift; shaded gray regions indicate $\pm\Omega_{\Delta\mathrm{RH}_{t_0}}$ around the zero-slope reference.

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

Table 4. Summary statistics of paired RH bias differences.

Site	$\overline{\Delta}(\Delta\mathrm{RH})_{t_7}$ [%]	$\overline{\Delta}(\Delta\mathrm{RH})_{t_{14}}$ [%]	$\Omega_{\Delta\mathrm{RH}_{t_0}}$ [%]
HKO*	-0.24 ±0.16	-0.40 ±0.21	±0.24
PAY	0.00 ±0.23	-0.18 ±0.23	±0.37
POT	-0.09 ±0.27	-0.21 ±0.27	±0.23
TEN	0.06 ±0.20	-0.07 ±0.25	±0.27
TAT	-0.88 ±0.78	-1.90 ±1.41	±1.02
URY	0.14 ±0.24	0.19 ±0.74	±0.48

* HKO, $\overline{\Delta}(\Delta\mathrm{RH})_{t_7,t_{14}}$ = -0.16 ±0.26%

345 ### 6.1.2 Meteoswiss, Payerne (PAY)

The second row of Figure 4 shows the results of the experiment performed at the PAY site using the Vaisala AS15 ARL. Overall, the mean values of the $\Delta\mathrm{RH}_t$ biases at $t_{0,7}$, $t_{0,14}$, t_7, and t_{14} range between approximately -1.0% and -1.3%. This systematic underestimation is likely related to the orientation of the RS sensor relative to the airflow direction inside the SHC, which can introduce a bias in relative humidity (Sommer et al., 2024).

350 The box plots indicate that the $\Delta\mathrm{RH}_t$ distributions at $t_{0,7}$ and $t_{0,14}$ exhibit very similar mean values and inter-quartile ranges, suggesting high repeatability of the SHC measurement performed by the $N_{\mathrm{RS}_{t_7}}$ RSs at $t_{0,7}$ and the $N_{\mathrm{RS}_{t_{14}}}$ RSs at $t_{0,14}$. Although the spread is larger than that observed at the HKO site, the repeatability ensures high confidence.

The mean value of $\Delta(\Delta\mathrm{RH})_{t_7}$ is very close to zero $(0.00\% \pm 0.23\%)$, as also reported in Table 4, indicating that no measurable degradation occurs after 7 days of residence time inside the ARL. In contrast, the mean value at t_{14}, $\overline{\Delta}(\Delta\mathrm{RH})_{t_{14}} =$
355 $-0.18\ \% \pm 0.23\ \%$, deviates further from zero reference, but remains within the repeatability interval $\pm\Omega_{\Delta\mathrm{RH}_{t_0}} = \pm 0.37\ \%$. This suggests good stability of the RH measurements over the investigated residence times, despite the presence of a persistent systematic bias.

6.1.3 CNR-IMAA, Potenza (POT)

The third row of Figure 4 shows the results of the experiment performed at the POT site using the Vaisala AS15 ARL. Overall,
360 the results of the statistical analysis are consistent with those obtained at the other Vaisala sites. In the box plots shown in the left panel, the mean values of the $\Delta\mathrm{RH}_t$ biases at $t_{0,7}$, $t_{0,14}$, t_7, and t_{14} range between approximately -0.2% and -0.7%, indicating a smaller systematic underestimation of the reference value $\mathrm{RH} = 100\%$ inside the SHC compared to PAY, likely due to its opposite positioning of the RS inside the SHC.

The inter-quartile ranges of the $\Delta\mathrm{RH}_t$ distributions are narrow and remain very similar across all experimental stages at $t_{0,7}$,
365 $t_{0,14}$, t_7, and t_{14}. In particular, the small spread observed at $t_{0,7}$ and at $t_{0,14}$ suggests high repeatability of the RSs performance inside the SHC.

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

The paired RH bias differences, shown in the right panels of Figure 4 and summarized in Table 4, indicate negative trends both at t_7 and at t_{14} even though they remain within the repeatability interval $\pm\Omega_{\Delta\mathrm{RH}_{t_0}} = \pm 0.23\,\%$. Specifically, after 7 days, the mean value is $\overline{\Delta}(\Delta\mathrm{RH})_{t_7} = -0.09\,\% \pm 0.27\,\%$, while after 14 days it increases in magnitude to $\overline{\Delta}(\Delta\mathrm{RH})_{t_{14}} = -0.21\% \pm 0.27\%$.

6.1.4 AEMET-Tenerife (TEN)

The fourth row of Figure 4 shows the results of the experiment performed at the TEN site using the Vaisala AS15 ARL. Overall, the results of the statistical analysis are consistent with those obtained at the other Vaisala sites. The mean values of the $\Delta\mathrm{RH}_t$ biases at $t_{0,7}$, $t_{0,14}$, t_7, and t_{14} range between approximately -0.1% and -0.5%, indicating a smaller systematic underestimation of the reference value $\mathrm{RH} = 100\%$ inside the SHC compared to the other Vaisala sites, likely due to most appropriate positioning of the RS inside a new update version of the SHC at TEN. However, the mean of $\Delta\mathrm{RH}_t$ biases at $t_{0,7}$ and $t_{0,14}$ show a larger difference compared to the other Vaisala sites. As mentioned in section 4.1, The difference of the mean at $t_{0,7}$ and $t_{0,14}$ can be explained by different environmental and operational conditions. Despite this apparent inconsistency in the experiment condition at the loading times, the paired RH bias differences, shown in the right panels of Figure 4 and summarized in Table 4, indicate a virtually zero trend both at t_7 and at t_{14}. Specifically, after 7 days, the mean value is slightly positive $\overline{\Delta}(\Delta\mathrm{RH})_{t_7} = 0.06\% \pm 0.20\%$, while after 14 days it becomes slightly negative $\overline{\Delta}(\Delta\mathrm{RH})_{t_{14}} = -0.07\% \pm 0.25\%$ also within the repeatability interval $\pm\Omega_{\Delta\mathrm{RH}_{t_0}} = \pm 0.27\,\%$.

6.1.5 Japan Meteorological Agency, Tateno (TAT)

The fifth row of Figure 4 presents the results of the experiment performed at the TAT site using the Meisei ARS ARL. As already mentioned in Section 4, the TAT dataset extended from January to September 2025 to monitor and quantify the seasonal effects on humidity measurements.

In the context of the experiment objectives, the Japan Meteorological Agency (JMA) team at TAT has reported that the IMS-100 RS humidity sensor exhibits a pronounced dry bias under high-humidity and temperature conditions, such as those inside the SHC. In particular, during the summer season, the magnitude of the dry bias increases with longer residence times of the RS inside the ARL. As the experiment spans multiple seasons, the resulting datasets show a considerably larger spread in the $\Delta\mathrm{RH}_t$ distributions across all experimental stages at $t_{0,7}$, $t_{0,14}$, t_7, and t_{14}. Moreover, the mean bias values at all stages range between approximately -3.5% and -4.5%, indicating a substantially larger systematic underestimation of the reference value $\mathrm{RH} = 100\%$ inside the SHC compared to the other sites.

Figure 5 further illustrates the dependence of the biases $\Delta\mathrm{RH}_t$ on both residence time and seasonal environmental conditions. The largest paired RH bias differences are observed during Summer and for the longer residence time t_{14}. These results suggest that prolonged storage under high-temperature and high-humidity environmental conditions, particularly longer than 7 days, may negatively affect the humidity sensor performance. This climate-driven response of the humidity sensors is the subject of ongoing investigation by the TAT team in collaboration with the Meisei manufacturer.

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

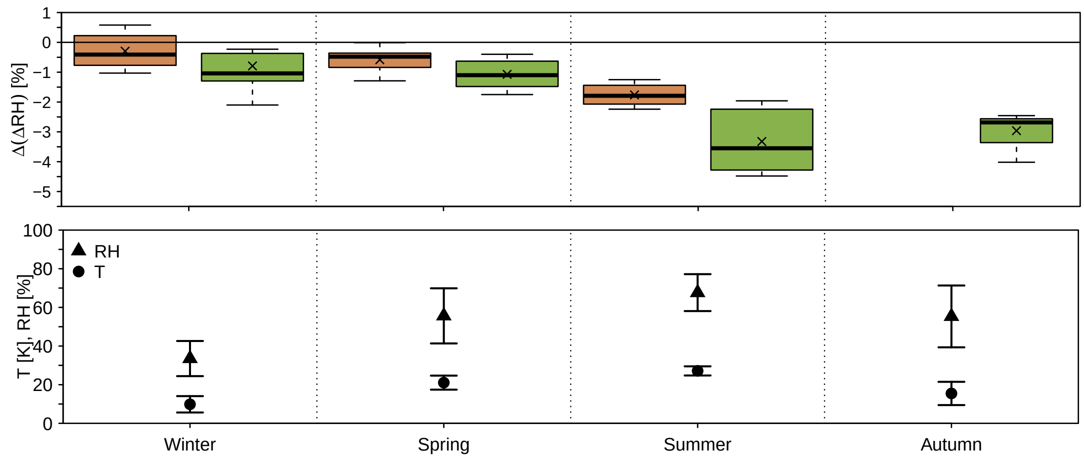

Figure 5. Top panel: seasonality of $\Delta(\Delta\mathrm{RH})$ at t_7 (orange) and t_{14} (green). Bottom panel: seasonality of T (circles) and RH (triangles). Highest T and RH mean values correspond to a larger inter-quartile range of the Summer paired RH bias difference at t_{14}

The paired RH bias differences shown in the right panel of Figure 4 confirm, over the full duration of the experiment, the strong impact of the combination of environmental conditions and residence time on reducing the ability of the RS to accurately measure $\mathrm{RH} = 100\%$ inside the SHC. Consistently with the results obtained at the Vaisala-operated sites, Table 4 indicates negative trends for both residence times t_7 and t_{14}. Specifically, after 7 days, the mean value is $\overline{\Delta}(\Delta\mathrm{RH})_{t_7} = -0.88\% \pm 0.78\%$, while after 14 days it increases markedly in magnitude to $\overline{\Delta}(\Delta\mathrm{RH})_{t_{14}} = -1.90\% \pm 1.41\%$ with a repeatability interval $\pm\Omega_{\Delta\mathrm{RH}_{t_0}} = \pm 1.02\ \%$.

6.1.6 MeteoModem (URY)

The last row of Figure 4 shows the results obtained at the URY site, which operates the MeteoModem M10 RS in combination with the Robotsonde ARL. The box plots of the $\Delta\mathrm{RH}_t$ distributions indicate moderately negative mean biases at all experimental stages, with a spread comparable to that observed at PAY and TEN operating the Vaisala configuration. The inter-quartile range at $t_{0,14}$ reveal a substantially larger spread compared with the bias distribution at $t_{0,7}$. This disparity suggests that the repeatability of the SHC measurements prior to storage is low, likely due to changes in environmental and operational conditions between $t_{0,7}$ and $t_{0,14}$. In contrast to all other sites, the paired RH bias differences for URY exhibit a reversed trend. Both show predominantly positive mean values, $\Delta(\Delta\mathrm{RH})_{t_7} = 0.14\% \pm 0.24\%$ and $\Delta(\Delta\mathrm{RH})_{t_{14}} = 0.19\% \pm 0.74\%$, indicating that the RH bias tends to decrease in magnitude with increasing residence time inside the ARL. This behavior suggests an apparent reduction in the initial RH bias magnitude after storage, rather than the deterioration observed at the other sites. Such a reversed tendency is not consistent with the expected gradual deterioration associated with prolonged storage inside the ARL. Although the mean values remain largely within the interval $\pm\Omega_{\Delta\mathrm{RH}_{t_0}} = \pm 0.48\ \%$, the consistent sign reversal relative to the other GRUAN sites points to site- or system-specific effects, potentially related to the interaction between the M10 sensor

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

characteristics and the Robotsonde ARL storage conditions. This distinct response warrants further investigation to clarify the underlying mechanisms.

6.2 Statistical significance of the results

Figure 6 summarizes the paired RH bias differences by grouped sites according to RS and ARL type.

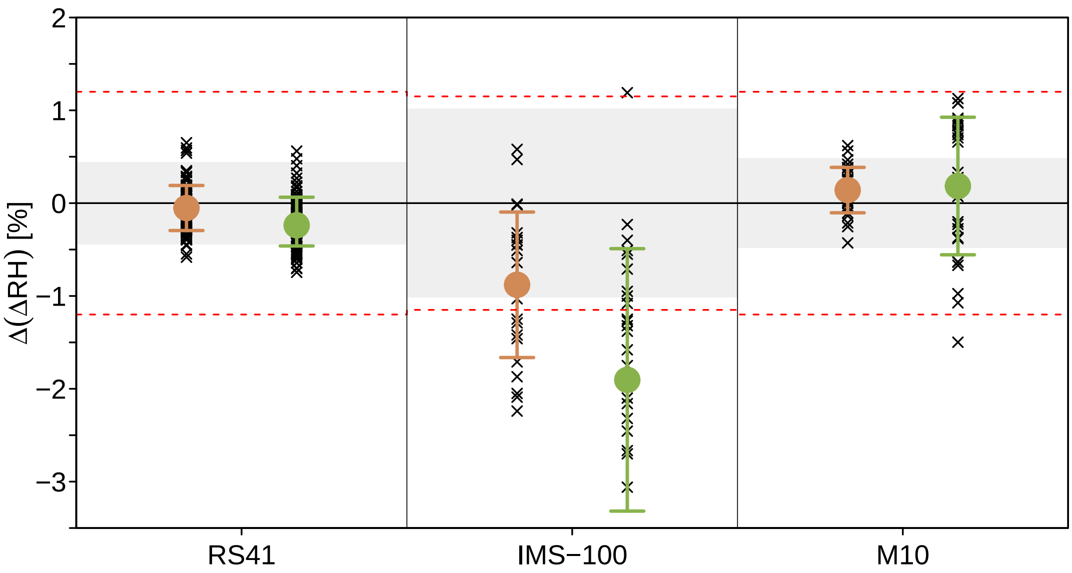

Figure 6. Summary of paired RH bias differences grouped by radiosonde and ARL configuration. The left column shows combined results from the four Vaisala sites (HKO, PAY, POT, TEN), the center column shows the results from the Meisei site (TAT), and the right column shows the results from the MeteoModem site (URY). "X" markers represent individual $\Delta(\Delta\mathrm{RH})$ values after 7 (mean in orange circle marker) and 14 (mean in green circle marker) days of residence time. The zero line indicates no systematic drift, the gray shading denotes $\pm\Omega_{\Delta\mathrm{RH}_{t_0}}$ and the dashed red lines indicate the uncertainty component $\pm\mathrm{U}_{\mathrm{G}}^{\mathrm{RS}}$ of the GDP after ground-preparation.

The left panel shows the aggregated paired RH bias difference distributions for the Vaisala RS41 RS operated with the AS15. The four Vaisala sites have overall mean values $\Delta(\Delta\mathrm{RH})_{t_7} = -0.05 \pm 0.24$ and $\Delta(\Delta\mathrm{RH})_{t_{14}} = -0.20 \pm 0.26$. The narrow clustering of the means reflects the overall robustness and consistency of the RS41–AS15 system observed across the four Vaisala sites. The majority of the individual values remain within the repeatability interval $\pm\Omega_{\Delta\mathrm{RH}_{t_0}}$ (gray-shaded around zero-line) and within the uncertainty interval $\pm\mathrm{U}_{\mathrm{G}}^{\mathrm{RS}}$ (uncertainty component of the GDP after ground-preparation), indicating that, on average, residence times up to 14 days inside the AS15 ARL do not lead to a large degradation of the RS41 humidity measurements (red dashed lines for $\pm\mathrm{U}_{\mathrm{G}}^{\mathrm{RS}}$).

The central panel presents the results from the TAT site. In contrast to the Vaisala group, a substantial fraction of individual measurements lies outside the interval $\pm\Omega_{\Delta\mathrm{RH}_{t_0}}$ and $\pm\mathrm{U}_{\mathrm{G}}^{\mathrm{RS}}$, highlighting the sensor response to both residence time and

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

seasonal high-humidity and temperature conditions, and confirms that prolonged storage inside the ARL significantly reduces the sensor's capability to accurately measure $\mathrm{RH} = 100\%$ already after t_7.

The right column shows, although most values remain within $\pm\Omega_{\Delta\mathrm{RH}_{t_0}}$ and $\pm\mathrm{U}^{\mathrm{RS}}_{\mathrm{G}}$, the reversed drift at the URY site compared with the other sites and indicates a distinct sensor–ARL interaction. On average, residence times up to 7 days inside the Robotsonde ARL do not lead to a large degradation of the M10 humidity measurements

As an additional robustness assessment, Table 5 summarizes the results obtained by applying a non-parametric bootstrap (BS) resampling ($10,000$ iterations) to the paired RH bias differences $\Delta(\Delta\mathrm{RH})^{\mathrm{BS}}_{t_7}$ and $\Delta(\Delta\mathrm{RH})^{\mathrm{BS}}_{t_{14}}$. The bootstrap-generated distributions reproduced the observed statistical behavior of the different radiosonde–ARL configurations, supporting the robustness and stability of the inferred residence-time effects.

Table 5. Non-parametric bootstrap (BS) resampling of the paired RH bias differences. $10,000$ bootstrap iterations were performed for each system configuration.

Configuration	$\overline{\Delta}(\Delta\mathrm{RH})^{\mathrm{BS}}_{t_7}$ [%]	95% CI [%]	$\overline{\Delta}(\Delta\mathrm{RH})^{\mathrm{BS}}_{t_{14}}$ [%]	95% CI [%]
RS41–AS15	-0.05	$[-0.10; 0.00]$	-0.20	$[-0.25; -0.15]$
IMS-100–ARS	-0.88	$[-1.18; -0.57]$	-1.90	$[-2.42; -1.40]$
M10–Robotsonde	0.14	$[0.05; 0.23]$	0.18	$[-0.09; 0.44]$

The mean values $\overline{\Delta}(\Delta\mathrm{RH})_{t_7}$ and $\overline{\Delta}(\Delta\mathrm{RH})_{t_{14}}$, shown in Figure 4, together with the bootstrap estimates $\overline{\Delta}(\Delta\mathrm{RH})^{\mathrm{BS}}_{t_7}$ and $\overline{\Delta}(\Delta\mathrm{RH})^{\mathrm{BS}}_{t_{14}}$ obtained from non-parametric resampling, are characterized by their statistical confidence intervals (Tables 4 and 5). When these values overlap with the repeatability interval $\pm\Omega_{\Delta\mathrm{RH}_{t_0}}$, no conclusion can be drawn as to whether the paired RH bias differences are statistically significant. To further analyze the results for the different RS–ARL configurations and to assess whether the paired RH bias differences significantly depart from zero, a Wilcoxon signed-rank test was applied to the paired RH bias difference distributions. The results are summarized in Table 6. The Wilcoxon signed-rank test is a non-parametric paired test that does not assume normality and accounts for the dependence between repeated measurements (Wilcoxon, 1945). The measurements are tested based on the null-hypothesis that the $\Delta(\Delta\mathrm{RH})_{t_7}$ and $\Delta(\Delta\mathrm{RH})_{t_{14}}$ are equal to zero. While the test gives a p-value (significance), the effect size r tells how strong the difference from zero is (see Table 6 notes for r). For RS41 (HKO, PAY, POT, TEN), a statistically significant difference was observed at t_7 ($p = 0.011$), associated with a small effect size ($r = 0.24$). A highly significant difference was found at t_{14} ($p < 0.001$), with a large effect size ($r = 0.62$), indicating an increase in the magnitude of the differences over the residence times t_7 and t_{14}. The Meisei RS IMS-100 (TAT) exhibited statistically significant differences at both t_7 and t_{14} ($p < 0.001$), with large effect sizes ($r = 0.77$ and $r = 0.83$, respectively), indicating a strong and consistent shift between time points. For M10 (URY), a statistically significant difference was detected at t_7 ($p = 0.0037$) with a large effect size ($r = -0.51$), indicating a substantial shift in the opposite direction compared to RS41 and IMS-100. No significant difference was observed at t_{14} ($p = 0.14$), with a small effect size ($r = -0.28$), most likely due to the large scatter of biases observed at $t_{0,14}$. These results show that the magnitude and direction of the temporal evolution of $\Delta(\Delta\mathrm{RH})$ depend also on the instrument type. In addition, the results highlight that statistical

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

significance does not necessarily imply a large effect size, as illustrated by RS41 at t_7, where a significant p-value is associated with a small effect.

Table 6. Wilcoxon signed-rank test results comparing initial conditions (t_0) to t_7 and t_{14}. Effect sizes (r) quantify the magnitude of the paired differences.

Configuration	p_{t_0,t_7}	r_{t_0,t_7}	Interpretation	$p_{t_0,t_{14}}$	$r_{t_0,t_{14}}$	Interpretation
RS41–AS15	0.011*	0.24	small	< 0.001***	0.62	large
IMS-100–ARS	< 0.001***	0.77	large	< 0.001***	0.83	large
M10–Robotsonde	0.004**	-0.51	large	0.14 (NS)	-0.28	small (NS)

Wilcoxon signed-rank tests were used to assess paired differences between time points.
Effect sizes were computed as $r = Z/\sqrt{N}$, where $Z = \frac{x-\bar{x}}{\sigma_x}$ is the standardized test statistic derived from the Wilcoxon signed-rank test and N is the number of paired observations.
Interpretation of r: small ($\sim$0.1), moderate ($\sim$0.3), large ($\geq$0.5).
Significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$; NS: not significant.

To further assess the practical relevance of the storage-related effects, the paired RH bias differences were evaluated relative to the GRUAN GDP humidity uncertainty associated with the ground-preparation component, U_G^{RS}, for each radiosonde type (Table 7). The corresponding uncertainty estimates after ground-preparation ($k = 1$) are approximately 1.2 %RH for RS41 (Sommer et al., 2023b), 1.15 %RH for IMS-100 (Kizu et al., 2018), and approximately 1.2 %RH for M10 based on the current GRUAN processing assessment under review by the GRUAN Lead Centre, (personal communication by Antoine Farah, MeteoModem).

Following (Dirksen et al., 2024, Appendix J, Eq. J.2), we show here the equivalence between the root-mean square value of the paired RH bias differences and the sum of the squared mean and the squared standard deviation. For a series of N paired RH bias differences, let $z_e = \Delta(\Delta\mathrm{RH})_e$ with $e \in [1;\ldots;N]$, with arithmetic mean $\bar{z}$ and standard deviation σ_z, their mean square z_{rms}^2 can be expressed as:

$$\begin{aligned}
z_{rms}^2 &= \frac{1}{N}\sum_e {z_e}^2 = \bar{z}^2 + z_{rms}^2 - 2\bar{z}^2 + \bar{z}^2 \\
&= \bar{z}^2 + \frac{1}{N}\sum_e {z_e}^2 - \frac{2}{N}\bar{z}\sum_e z_e + \frac{1}{N}\sum_e \bar{z}^2 \\
&= \bar{z}^2 + \frac{1}{N}\sum_e (z_e - \bar{z})^2 \\
&= \bar{z}^2 + {\sigma_z}^2. \qquad (1)
\end{aligned}$$

Following from Equation 1, in order to account for both the magnitude and variability of the storage-related effects, the normalized metric is defined as,

$$\mathrm{M_U} = \frac{\sqrt{\bar{z}^2 + \sigma_z^2}}{\mathrm{U_G^{RS}}}. \qquad (2)$$

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

The metric $\mathrm{M_U}$ therefore represents the quadratic combination of the systematic storage-related drift and its random variability normalized by the corresponding GRUAN humidity uncertainty after ground-preparation. Values of $\mathrm{M_U}$ much smaller than unity indicate that the storage-related effects remain limited compared to the expected GDP uncertainty range. In contrast, values approaching or exceeding unity indicate that storage-related effects become comparable to, or larger than, the uncertainty contribution associated with the GDP ground-preparation component.

For the RS41–AS15 configuration, $\mathrm{M_U}$ remained well below unity at both t_7 and t_{14}, indicating limited impact of storage on the SHC measurements. Although statistically significant negative drifts were identified, particularly at t_{14}, their magnitudes remain small relative to the GRUAN humidity uncertainty budget.

In contrast, the IMS-100–ARS configuration exhibited substantially larger $\mathrm{M_U}$ values, approaching or exceeding unity depending on the residence time considered. This behavior is consistent with the pronounced dry-bias increase and the strong statistical significance (Table 6), indicating that storage-related effects become relevant for residence times extending toward 14 days, particularly under high-humidity and temperature environmental conditions.

For the M10–Robotsonde configuration, $\mathrm{M_U}$ remained below unity but increased substantially at t_{14}, mainly due to the larger variability of the paired RH bias differences rather than to a pronounced systematic drift. Although the mean drift remained moderate, the increased spread suggests a less stable humidity response after prolonged storage.

Table 7. Normalized metric $\mathrm{M_U}$ results. The statistical significance was quantitatively assessed from the magnitude of $\mathrm{M_U}$. Values of $\mathrm{M_U}$ much smaller than unity indicate that the storage-related effects remain limited compared to the expected GDP uncertainty range. In contrast, values approaching or exceeding unity indicate that storage-related effects become comparable to, or larger than, the uncertainty contribution associated with the GDP ground-preparation component.

Configuration	$\mathrm{M_{U,t_7}}$	$\mathrm{M_{U,t_{14}}}$	Significance
RS41–AS15	0.21	0.27	Limited impact
IMS-100–ARS	1.02	2.06	Increasing impact with residence time
M10–Robotsonde	0.23	0.64	Limited impact despite elevated variability

In summary, the paired RH bias difference quantifies systematic storage-related bias changes, $\Omega_{\Delta \mathrm{RH}_{t_0}}$ characterizes the intrinsic variability of the SHC measurements, bootstrap resampling assesses the robustness and confidence intervals of the estimated effects, Wilcoxon signed-rank tests and associated effect sizes evaluate the statistical significance and magnitude of the observed changes, while the normalized $\mathrm{M_U}$ metric evaluates their relevance relative to the ground-preparation component of the GDP humidity uncertainty.

7 Conclusions

This study investigates the effect of radiosonde residence time inside ARLs on the accuracy and stability of humidity measurements within the SHC. It also aims to evaluate the suitability of combining SHC measurements with automatic observations from ARLs in the context of GRUAN operations. The analysis, conducted across six sites and three RS–ARL configurations,

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

demonstrates that the response to storage is strongly system-dependent and influenced by environmental conditions and sensor handling procedures. For the Vaisala RS41–AS15 configuration, humidity measurements remain generally stable for residence times up to 14 days. Although small negative drifts are observed, they are typically within the expected measurement variability and well below the GRUAN uncertainty budget, indicating limited impact. These results highlight the robustness of the RS41 system, which is likely reinforced by the sensor reconditioning carried out during the Manufacturer Ground-Check (MGC).

In contrast, the Meisei IMS-100–ARS configuration exhibits a pronounced increase in dry bias with increasing residence time, particularly under high-humidity and temperature conditions. Storage-related effects increase markedly with residence time and become strongly statistically significant after 14 days, highlighting the sensitivity of this system to environmental exposure and the lack of an active reconditioning step. The MeteoModem M10–Robotsonde configuration shows a distinct behavior, with a partial reduction of the initial bias during storage, although the underlying mechanisms require further investigation.

Across all configurations, the optimal residence time for minimizing humidity bias drifts is found to be within 7 days, while storage up to 14 days represents an upper limit for maintaining GRUAN-compliant performance, i.e. after ground-preparation uncertainty estimation ($k = 1$). Extended residence times, particularly in uncontrolled environments, may lead to sensor contamination and reduced effectiveness of sensor reconditioning procedures. Overall, the results confirm the feasibility of implementing SHC tests for ARL-based operations and suggest that the use of the SHC may support compliance with the high-level measurement requirements of reference observing networks such as GRUAN.

The analysis clearly shows that the residence time inside ARLs can be explicitly considered as an additional source of uncertainty in the GRUAN GDP humidity uncertainty budget. The observed paired RH bias differences demonstrate that the effect of storage is system-dependent and may become significant under specific environmental and operational conditions, particularly for prolonged residence times and non-climate-controlled ARLs. Therefore, a dedicated residence-time uncertainty component should be developed and propagated into the final GDP uncertainty budget. Such an approach would ensure that the long-term traceability and reference-quality requirements of GRUAN are met under automated conditions.

Code and data availability. The analysis scripts developed for this study will be archived and made publicly available upon publication of the manuscript. The radiosonde and SHC datasets used in this analysis originate from GRUAN-related operational activities and are distributed according to the GRUAN data policy. The data are available through the GRUAN data portal and can be obtained from the corresponding author or the respective contributing institutions upon reasonable request.

Author contributions. G.R., G.M., and M.R. conceived the study design and coordinated the GRUAN-Task Team 4 (TT4) experimental activities. Measurements, radiosonde preparation, and SHC tests were carried out at the participating sites by G.R., N.P., C.C., M.R., Y.P., S.A., G.C., A.F., I.G., M.H., M.I., P.J., J.M.C.L., J.M.R. and Y.T.

G.R. performed the statistical analysis and prepared the figures with contributions from G.M. and Y.P.. The interpretation of the results was discussed collectively within the GRUAN-TT4 team.

G.R and G.M. led the manuscript writing with major contributions from C.C., N. P., M. R. and Y.P..

Fabio Madonna, C.F., A.H., and Y.-A.R. contributed through scientific coordination, institutional support, and critical review of the manuscript.

All authors contributed to the scientific discussion, reviewed the manuscript, and approved the final version.

Competing interests. The authors declare no competing interests.

Acknowledgements. The authors would like to thank all GRUAN-TT4 participants and site operators for their contribution to the experimental activities and for the valuable scientific discussions conducted throughout the project.

We gratefully acknowledge the operational support provided by the participating GRUAN sites: HKO, PAY, POT, TAT, TEN, and URY. Special thanks are extended to the technical staff involved in radiosonde preparation, SHC operations, and ARL maintenance during the measurement campaigns.

The authors also acknowledge the support and collaboration of Vaisala, Meisei, and MeteoModem for technical exchanges related to the radiosonde and ARL systems investigated in this study.

This work was conducted within the framework of the GCOS Reference Upper-Air Network (GRUAN) Task Team 4 (TT4) activities dedicated to the implementation of Standard Humidity Chamber (SHC) procedures for Automatic Radiosonde Launchers (ARLs). Y.P. is member of the Oeschger Centre for Climate Change Research and contributed to this work with support from the "Swiss H2O Hub for High-quality water vapor measurements from ground to space" project financed by GCOS Switzerland (https://www.meteoswiss.admin.ch/about-us/research-and-cooperation/programmes-gaw-ch-and-gcos-ch/gaw-ch-and-gcos-ch-supported-activities.html, last access: 2026-05-19). The authors further acknowledge the institutional support provided by MeteoSwiss, CNR-IMAA, AEMET and the University of Bern.

The authors thank the anonymous reviewers and colleagues involved in the internal review process for their constructive comments and suggestions, which helped improve the manuscript.

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

The authors used an AI-based language tool to improve the grammar and readability of the manuscript and reviewed the output. The authors take full responsibility for the content of the paper.

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

List of acronyms

ARL	Automatic Radiosonde Launcher
ARLs	Automatic Radiosonde Launchers
AS15	Vaisala Autolauncher
GC	Ground Check
GCOS	Global Climate Observing System
GDP	GRUAN Data Product
GRUAN	GCOS Reference Upper-Air Network
JMA	Japan Meteorological Agency
MGC	Manufacturer Ground-Check
QC	Quality Controlled
RS	Radiosonde
SHC	Standard Humidity Chamber
VOC	Volatile Organic Compound
TT4	Task Team 4
WMO	World Meteorological Organization

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

References

Arkan, F. and Nababan, D.: Radiosonde Transmitter Meisei iMS-100 Study as Observation on Air at Meteorology Station Klas I Depati Amir Pangkalpinang, Journal of Electrical Technology UMY, 1, 1–8, https://doi.org/10.18196/jet.1212, 2017.

Bodeker, G. E., Bojinski, S., Cimini, D., Dirksen, R. J., Haeffelin, M., Hannigan, J. W., Hurst, D. F., Leblanc, T., Madonna, F., Maturilli, M., Mikalsen, A. C., Philipona, R., Reale, T., Seidel, D. J., Tan, D. G. H., Thorne, P. W., Vömel, H., and Wang, J.: Reference Upper-Air Observations for Climate: From Concept to Reality, Bulletin of the American Meteorological Society, 97, 123–135, https://doi.org/10.1175/BAMS-D-14-00072.1, 2016.

Diarra, S., Montoux, N., Baray, J.-L., Keckhut, P., Dupont, J.-C., Farah, A., and Alraddawi, D.: Comparison of M10 and M20 Meteomodem radiosondes relative humidity measurements with ECMWF ERA5 above France: focus on the upper troposphere, EGUsphere, 2026, 1–28, https://doi.org/10.5194/egusphere-2026-482, 2026.

Dirksen, R. et al.: Report of WMO's 2022 Upper-Air Instrument Intercomparison Campaign, IOM-143, World Meteorological Organization, https://library.wmo.int/idurl/4/68808, 2024.

Dirksen, R. J., Sommer, M., Immler, F. J., Hurst, D. F., Kivi, R., and Vömel, H.: Reference quality upper-air measurements: GRUAN data processing for the Vaisala RS92 radiosonde, Atmospheric Measurement Techniques, 7, 4463–4490, https://doi.org/10.5194/amt-7-4463-2014, 2014.

Dirksen, R. J., Bodeker, G. E., Thorne, P. W., Merlone, A., Reale, T., Wang, J., Hurst, D. F., Demoz, B. B., Gardiner, T. D., Ingleby, B., Sommer, M., Von Rohden, C., and Leblanc, T.: Managing the transition from Vaisala RS92 to RS41 radiosondes within the Global Climate Observing System Reference Upper-Air Network (GRUAN): A progress report, Geoscientific Instrumentation, Methods and Data Systems, 9, 337–355, https://doi.org/10.5194/gi-9-337-2020, 2020.

Dupont, J.-C., Haeffelin, M., Badosa, J., Clain, G., Raux, C., and Vignelles, D.: Characterization and Corrections of Relative Humidity Measurement from Meteomodem M10 Radiosondes at Midlatitude Stations, Journal of Atmospheric and Oceanic Technology, 37, 857 – 871, https://doi.org/10.1175/JTECH-D-18-0205.1, 2020.

European Space Agency: Climate Data Records at a Glance, https://climate.esa.int/en/climate-data-records-at-a-glance/, accessed: 2026-05-29, 2026.

GCOS: GCOS Reference Upper-Air Network (GRUAN): Justification, Requirements, Siting and Instrumentation Options, Tech. rep., World Meteorological Organization (WMO), https://library.wmo.int/idurl/4/48112, GCOS-112, Version 1, April 2007, 2007.

Global Climate Observing System (GCOS): Guideline for the Generation of Datasets and Products Meeting GCOS Requirements, Tech. Rep. 143, World Meteorological Organization, Geneva, Switzerland, https://library.wmo.int/records/item/39563-guideline-for-the-generation-of-datasets-and-products-meeting-gcos-requirements, gCOS-143, 2010.

Hardy, B.: ITS-90 Formulations for Vapor Pressure, Frostpoint Temperature, Dewpoint Temperature, and Enhancement Factors in the Range -100 to $+100$ °C, in: Proceedings of the Third International Symposium on Humidity & Moisture, Teddington, London, England, April 1998, https://www.thunderscientific.com/wp-content/uploads/2022/12/its90formulas.pdf, accessed: 2026-05-19, 1998.

Hoshino, S., Sugidachi, T., Shimizu, K., Kobayashi, E., Fujiwara, M., and Iwabuchi, M.: Comparison of GRUAN data products for Meisei iMS-100 and Vaisala RS92 radiosondes at Tateno, Japan, Atmospheric Measurement Techniques, 15, 5917–5948, https://doi.org/10.5194/amt-15-5917-2022, 2022.

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

Immler, F. J., Dykema, J., Gardiner, T., Whiteman, D. N., Thorne, P. W., and Vömel, H.: Reference Quality Upper-Air Measurements: guidance for developing GRUAN data products, Atmospheric Measurement Techniques, 3, 1217–1231, https://doi.org/10.5194/amt-3-1217-2010, 2010.

Jauhiainen, H., Lentonen, J., Survo, P., Lehtinen, R., and Pietari, T.: The Implications of Vaisala's new Radiosonde RS41 on Improved in-situ Observations for Meteorological Applications, in: 94th American Meteorological Society Annual Meeting, February 2–6, 2014, Atlanta, GA, USA, pp. 1–15, American Meteorological Society, https://ams.confex.com/ams/94Annual/webprogram/Manuscript/Paper230538/6_2ImplicationsofVaisalaRadiosondeRS41.pdf, 2014.

Kizu, N., Sugidachi, T., Kobayashi, E., Hoshino, S., Shimizu, K., Maeda, R., and Fujiwara, M.: Technical characteristics and GRUAN data processing for the Meisei RS-11G and iMS-100 radiosondes, GRUAN Technical Document 5 (GRUAN-TD-5), GRUAN, https://www.gruan.org/documentation/gruan/td/gruan-td-5/, 2018.

Lee, S. W., Choi, B. I., Woo, S. B., Kim, J. C., and Kim, Y. G.: Development of a low-temperature low-pressure humidity chamber for calibration of radiosonde humidity sensors, Metrologia, 56, 1—16, https://doi.org/10.1088/1681-7575/ab0cc0, 2019.

Madonna, F., Kivi, R., Dupont, J.-C., Ingleby, B., Fujiwara, M., Romanens, G., Hernandez, M., Calbet, X., Rosoldi, M., Giunta, A., Karppinen, T., Iwabuchi, M., Hoshino, S., von Rohden, C., and Thorne, P. W.: Use of automatic radiosonde launchers to measure temperature and humidity profiles from the GRUAN perspective, Atmos. Meas. Tech., 13, 3621–3649, https://doi.org/10.5194/amt-13-3621-2020, 2020.

MeteoModem: M10 Radiosonde Technical Specifications, Tech. rep., MeteoModem, Ury, France, https://simtech.com.br/wp-content/uploads/2023/07/M10-EN.pdf, accessed: 2026-05-19, 2023.

Miloshevich, L. M., Vömel, H., Whiteman, D. N., and Leblanc, T.: Accuracy Assessment and Correction of Vaisala RS92 Radiosonde Water Vapor Measurements, Journal of Geophysical Research, 114, D11 305, https://doi.org/10.1029/2008JD011565, 2009.

Sairanen, H., Heinonen, M., Högström, R., Lakka, A., and Kajastie, H.: A Calibration System for Reference Radiosondes that Meets GRUAN Uncertainty Requirements, NCSLI Measure, 9, 56–60, https://doi.org/10.1080/19315775.2014.11721696, 2014.

Sairanen, H., Heinonen, M., and Högström, R.: Validation of a calibration set-up for radiosondes to fulfil GRUAN requirements, Measurement Science and Technology, 26, 105 901, https://doi.org/10.1088/0957-0233/26/10/105901, 2015.

Seidel, D. J., Berger, F. H., Diamond, H. J., Dykema, J., Goodrich, D., Immler, F., Murray, W., Peterson, T., Sisterson, D., Sommer, M., Thorne, P., Vömel, H., and Wang, J.: Reference Upper-Air Observations for Climate: Rationale, Progress, and Plans, Bulletin of the American Meteorological Society, 90, 361–369, https://doi.org/10.1175/2008BAMS2540.1, 2009.

Sommer, M., von Rohden, C., and Simeonov, T.: User Guide of the RS41 GRUAN Data Product Version 1 (RS41-GDP.1), GRUAN Technical Note 13 (GRUAN-TN-13), GRUAN Lead Centre, https://www.gruan.org/documentation/gruan/tn/gruan-tn-13, v1.0 (2022-11-21), 2022.

Sommer, M., von Rohden, C., Simeonov, T., Oelsner, P., Naebert, T., Romanens, G., Jauhiainen, H., Survo, P., and Dirksen, R.: GRUAN characterisation and data processing of the Vaisala RS41 radiosonde, GRUAN Technical Document 8 (GRUAN-TD-8), GRUAN, https://www.gruan.org/documentation/gruan/td/gruan-td-8, v1.0.0 (2023-06-28), 2023a.

Sommer, M., von Rohden, C., Simeonov, T., Oelsner, P., Naebert, T., Romanens, G., Jauhiainen, H., Survo, P., and Dirksen, R.: GRUAN characterisation and data processing of the Vaisala RS41 radiosonde (GRUAN-TD-8), GRUAN Technical Document, 8, –, https://www.gruan.org/documentation/gruan/td/gruan-td-8, v1.0.0, published 2023-06-28, 2023b.

Sommer, M., von Rohden, C., Oelsner, P., and Dirksen, R.: Orientation effect of RS41 radiosonde in the Standard Humidity Chamber (SHC), GRUAN Technical Note x (GRUAN-TN-x), GRUAN Lead Centre, DRAFT - Manuscript in preparation, 2024.

https://doi.org/10.5194/egusphere-2026-3215
Preprint. Discussion started: 12 August 2026

Survo, P., Hiltunen, E., Jauhiainen, H., Lentonen, J., Leppänen, J., Salo, T., and Turunen, M.: Atmospheric temperature and humidity measurements of Vaisala radiosonde RS41, in: 17th Symposium on Meteorological Observations and Instrumentation, 9–13 June 2014, Westminster, CO, USA, pp. 1–14, American Meteorological Society, https://ams.confex.com/ams/21Applied17SMOI/webprogram/Manuscript/Paper247141/Extended%20Abstract_AMS%20SMOI%202014_%20Atmospheric%20Temperature%20and%20Humidity%20measurements%20of%20RS41.pdf, 2014a.

Survo, P., Lehtinen, R., and Kauranen, J.: SI traceability of Vaisala radiosonde RS41 sounding data – calibration and uncertainty analysis, in: Proceedings of the WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation (TECO 2014), Saint Petersburg, Russian Federation, 07-09 July 2014, pp. 1–8, https://library.wmo.int/idurl/4/53114, 2014b.

Vaisala Oyj: Measuring Humidity and Dew Point in Environments with Chemical Contaminants (Ref. B211241EN-A), Tech. rep., Helsinki, Finland, https://www.kenelec.com.au/wp-content/uploads/2018/04/Vaisala-B21-1241EN-A-Measuring-Humidity-and-Dew-Point-in-Environments-With-Chemical-Contaminants-AppNote.pdf, accessed: 2026-05-19, 2012.

Vaisala Oyj: White Paper: Ground Check Device RI41. Ref. B211539EN-A, Tech. rep., Vaisala Oyj, Helsinki, Finland, https://www.vaisala.com/sites/default/files/documents/WhitepaperGroundCheckofRS41B211539EN-A.pdf, ©Vaisala, 2016.

Vaisala Oyj: White Paper: Vaisala Radiosonde RS41 Measurement Performance. Ref. B211356EN-B, Tech. rep., Vaisala Oyj, Helsinki, Finland, https://www.vaisala.com/sites/default/files/documents/WEA-MET-RS41-Performance-White-paper-B211356EN-B-LOW-v3.pdf, ©Vaisala, 2017.

Vaisala Oyj: Vaisala HUMICAP® Sensor for Measuring Relative Humidity (Ref. B210781EN), Vaisala Oyj, Helsinki, Finland, https://www.vaisala.com/sites/default/files/documents/HUMICAP-Technology-description-B210781EN.pdf, technology description brochure, 2020.

Wilcoxon, F.: Individual Comparisons by Ranking Methods, Biometrics Bulletin, 1, 80–83, https://doi.org/10.2307/3001968, 1945.

World Meteorological Organization (WMO): Guide to Instruments and Methods of Observation (WMO-No. 8), Volume I, World Meteorological Organization, Geneva, https://doi.org/10.59327/WMO/CIMO/1, 2024.