Atmos. Meas. Tech. Discuss., 7, 6327–6357, 2014
www.atmos-meas-tech-discuss.net/7/6327/2014/
doi:10.5194/amtd-7-6327-2014

Atmospheric Measurement Techniques Discussions
Open Access

Quantifying the value of redundant measurements at GRUAN sites

F. Madonna[1], M. Rosoldi[1], J. Güldner[2], A. Haefele[3], R. Kivi[4], M. P. Cadeddu[5], D. Sisterson[5], and G. Pappalardo[1]

[1]Consiglio Nazionale delle Ricerche, Istituto di Metodologie per l'Analisi Ambientale (CNR-IMAA), C. da S. Loja, Tito Scalo, Potenza, 85050
[2]Deutscher Wetterdienst, Meteorologisches Observatorium Lindenberg Richard Assmann Observatorium, Am Observatorium 12, 15848 Tauche/Lindenberg, Germany
[3]Federal Office of Meteorology and Climatology, MeteoSwiss, Chemin de l'Aérologie, 1530 Payerne, Switzerland
[4]Finnish Meteorological Institute, Arctic Research, Tähteläntie 62, 99600 Sodankylä, Finland
[5]Argonne National Laboratory, 9700 South Cass Avenue, Argonne, IL 60439-4801, USA

Received: 28 April 2014 – Accepted: 3 June 2014 – Published: 23 June 2014

Correspondence to: F. Madonna (fabio.madonna@imaa.cnr.it)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

The potential for measurement redundancy to reduce uncertainty in atmospheric variables has not been investigated comprehensively for climate observations. We evaluated the usefulness of entropy and mutual correlation concepts, as defined in information theory, for quantifying random uncertainty and redundancy in time series of atmospheric water vapor provided by five highly instrumented GRUAN (GCOS [Global Climate Observing System] Reference Upper-Air Network) Stations in 2010–2012. Results show that the random uncertainties for radiosonde, frost-point hygrometer, Global Positioning System, microwave and infrared radiometers, and Raman lidar measurements differed by less than 8 %. Comparisons of time series of the Integrated Water Vapor (IWV) content from ground-based remote sensing instruments with in situ soundings showed that microwave radiometers have the highest redundancy and therefore the highest potential to reduce random uncertainty of IWV time series estimated by radiosondes. Moreover, the random uncertainty of a time series from one instrument should be reduced of ∼ 60 % by constraining the measurements with those from another instrument. The best reduction of random uncertainty resulted from conditioning of Raman lidar measurements with microwave radiometer measurements. Specific instruments are recommended for atmospheric water vapor measurements at GRUAN sites. This approach can be applied to the study of redundant measurements for other climate variables.

1 Introduction

The use of redundant measurements is considered the best approach for improving the knowledge of atmospheric processes derived from in situ and remote sensing measurements and also the most efficient way to reduce uncertainty regarding an atmospheric variable. For this reason, several atmospheric observatories have extended their observing capabilities and have acquired multiple instruments to measure the

AMTD
7, 6327–6357, 2014

Quantifying the value of redundant measurements at GRUAN sites

F. Madonna et al.

same atmospheric variables with different techniques and and to estimate retrieval algorithms.

Without doubt, redundant measurements provide added value. The advantages are related to

- filling gaps and improving measurement continuity over time and vertical range;
- increasing the sampling rate by merging measurements from different instruments;
- addressing instrument noise and identifying possible biases or retrieval problems by comparing different techniques and instruments;
- providing advanced products and exploiting instrument synergy and data integration.

However, comprehensive studies to quantify the effective value of redundant measurements and their ability to reduce uncertainty in essential climate variables (ECVs), as retrieved by multiple ground-based techniques and in situ active and passive remote sensing, are missing. To this end, GRUAN (GCOS [Global Climate Observing System] Reference Upper-Air Network) aims at providing long-term, highly accurate measurements of atmospheric profiles, complemented by surface-based state-of-the-art instrumentation, for full characterization of ECVs and their changes in the complete atmospheric column (Thorne et al., 2013). GRUAN, which is now being implemented, will soon support a network of 30–40 high-quality, long-term upper-air observing stations, building on existing observational networks.

Cross-checking of redundant measurements for consistency is an essential part of the GRUAN quality assurance procedures. A fully equipped GRUAN site should make at least three redundant measurements of all GCOS ECVs (http://www.wmo.int/pages/prog/gcos/Publications/GCOS_brochure2010.pdf). As a consequence, the GRUAN community has fostered GATNDOR (GRUAN Analysis Team for Network Design and Operations Research), a scientific team charged to address key scientific

AMTD
7, 6327–6357, 2014

questions of major interest to GRUAN and identify reliable metrics for quantifying the value of redundant measurements.

The present study used observations of the vertical-profile of water vapor mixing ratio and the integrated water vapor (IWV) content from a few GRUAN sites equipped with radiosondes, global positioning system (GPS), lidars, radiometers, spectrometers, and radars. Studies of redundant measurements should be based on the preliminary identification of a reliable metric. Linear correlation (Pearson's or Spearman's) has typically been used to study redundant measurements and their reliability. More recently, Fassò et al. (2013) presented a new approach for statistical modeling of the relationships between uncertainty of atmospheric measurement collocation and a set of environmental factors. The approach, which can decompose the total uncertainty, could be adapted to evaluate the measurement redundancy. In this paper, we present the results of the GATNDOR study of redundant measurements at GRUAN sites. This study identified mutual correlation (MC), which is related to the concept of entropy, as a suitable metric for quantifying the value of measurement redundancy. In information theory, entropy is a measure of the probabilistic uncertainty associated with a random variable. The approach presented here represents a fast, efficient way to quantify the value of redundant measurements and to correlate the value with factors such as number of instruments, type of measurement techniques, and retrieval algorithms.

The aims of the paper are:

- to show the potential of entropy and MC as metrics for quantifying uncertainty (in a probabilistic sense) and the value of redundancy in climate time series;

- to study, according to GRUAN standards, the uncertainty and the value of redundancy of in situ and ground-based remote sensing techniques for estimating ECVs;

- to provide the GRUAN community and others interested in the observation of atmospheric thermodynamics with recommendations for the establishment of an

AMTD
7, 6327–6357, 2014

observation protocol to reduce the uncertainty of a measurement time series through measurement redundancy.

Section 2 outlines information theory concepts used for the study of redundancy and presents the data sets considered in this work. The data sets were provided by five GRUAN sites: the Atmospheric Radiation Measurement (ARM) Program Southern Great Plains in Oklahoma (Miller et al., 2003), USA; CIAO (Consiglio Nazionale delle Ricerche, Istituto di Metodologie per l'Analisi Ambientale [CNR-IMAA] atmospheric observatory) in Potenza (Madonna et al., 2011), Italy; Lindenberg in Germany (Adam et al., 2005); Payerne in Switzerland (Calpini et al., 2011); and Sodankyla in Finland (Hirsikko et al., 2014). Section 3 provides results and preliminary remarks on the value of redundant measurements in reducing uncertainty and introduces a possible criterion for addressing redundancy in the frame of GRUAN. Section 4 summarizes conclusions.

2 Methodology

2.1 Comparison methods

Comparisons among time series of in situ and ground-based remote sensing measurements have been performed mostly by using the concept of variance and root-mean-square difference, less frequently in terms of "information" content (e.g., Majda and Gershgorin, 2010). In information theory, as in thermodynamics, entropy is a measure of the number of specific ways a system can be arranged. Entropy is often considered a measure of disorder, of the freedom in selection of an event, or of uncertainty in the outcome or the prediction of an event. Commonly used in time series analysis is the Shannon–Wiener entropy measure (Cover and Thomas, 1991). Given x events in the population X occurring with probabilities $p(x)$, the Shannon entropy is defined as

$$H(x) = -\sum_{x\in X} p(x)\log(p(x)). \tag{1}$$

AMTD
7, 6327–6357, 2014

Therefore, H is a measure of probabilistic uncertainty or dispersion of the probabilities of events. The entropy is calculated from a histogram of probabilities; it has a maximum value if all measurements have equal probability of occurrence or a minimum value of 0 if the probability of one measurement is 1 and the probability of all the others is 0. H is not equivalent to variance (σ), though for particular classes of distributions (e.g., Gaussian) H is simply some function of σ, and these can be considered almost equivalent. Entropy generalizes the concept of measurement uncertainty for calculations of MC. Normalized H is used here to quantify the uncertainty of a time series, and H is normalized by dividing H by the logarithm of the number of states (i.e., the number of possible entries in the related histogram).

In information theory, MC is a measure of the statistical dependence between two random variables or, equivalently, the amount of information that one variable contains about the other (Cover and Thomas, 1991). The MC value can be considered a qualitative indication of how well one measurement explains the other. This means that MC quantifies the reduction of uncertainty in a variable Y after one observes another variable X. The advantage in using MC with respect to Pearson's or Spearman's correlation coefficient (ρ) is that MC is applied to linear, non-linear, and non-monotonic correlations.

The MC of two discrete random variables X and Y can be defined as (Cover and Thomas, 1991)

$$\mathrm{MC}(X,Y) = \sum_{y \in Y} \sum_{x \in X} p(x,y) \log\left(\frac{p(x,y)}{p(x)p(y)}\right), \tag{2}$$

where $p(x,y)$ is the joint probability distribution function of X and Y, and $p(x)$ and $p(y)$ are the marginal probability distribution functions of X and Y, respectively. For continuous random variables, the summation is implemented with a definite double integral. Redundancy concept is a generalization of mutual information to N variables $(X_1, X_2, \ldots, X_N)$. Given as marginal entropies $H(X)$ and $H(Y)$, MC can be also defined

AMTD
7, 6327–6357, 2014

as:

$$\mathrm{MC}(X,Y) = H(X) + H(Y) - H(X,Y). \tag{3}$$

The joint entropy $H(X,Y)$ is the total amount of information for two time series and is calculated by using the joint histogram of the two series. If the two measurements are totally unrelated, then the joint entropy will be the sum of the entropies of the individual measurements. In general, $H(X,Y) \le H(X) + H(Y)$.

The MC can be also linearized. Differences between non-linear and linear redundancy provide a qualitative test for the non-linearity of the investigated problem. The linear MC is defined as (Cover and Thomas, 1991)

$$\mathrm{LMC} = \frac{1}{2}\sum_{i=1}^{m} C_{ii} - \frac{1}{2}\sum_{i=1}^{m} \lambda_i^C, \tag{4}$$

where the C_{ii} values are the diagonal elements of the covariance matrix **C** of the m time series investigated, and the λ values are the eigenvalues of **C**. A comparison between linear and non-linear MC is in Sect. 3.4.

Many applications require a metric – a distance measure not only between points but also between data clusters (or time series of data). Different distances are defined in the literature. Here, D is defined as

$$D(X,Y) = 1 - \mathrm{MC}(X,Y)/\max(H(X),H(Y)), \tag{5}$$

where D is a metric that satisfies the triangle inequality. Calculation of MC is an effective way to compare clustering and study relationships between time series (Correa and Lindstrom, 2012). Entropy and mutual information are both rather insensitive to outliers, but even a single outlier can arbitrarily impact both the variance and correlation between two distributions, obscuring the similarity of two closely related variables. The entropy gained from a member of a mixture of distributions is the difference between the entropy of the average distribution and the average of the entropies of the

AMTD
7, 6327–6357, 2014

individual distributions. $H(X,Y)$ can be calculated by using the joint histogram of X and Y.

Finally, in agreement with the axiom of information theory, we define the conditional entropy as $H(X|Y) = H(Y) - \mathrm{MC}(X,Y)$. This definition can be generalized for two or more conditioning variables through the chain rule for joint entropy (Cover and Thomas, 1991).

2.2 Data sets and instruments

This study focuses on a dataset including radiosonde, Raman lidar, infrared, and microwave radiometry (MWR) observations from the GRUAN sites (Lindenberg [LIN], Payerne [PAY], Potenza [POT], Sodankyla [SOD], and ARM Southern Great Plains [SGP]). More information about the selected sites can be found at www.gruan.org. This study focused on the investigation of atmospheric water vapor measurements, both along the vertical porfile and integrated over the complete column, from these sites for the period 2010–2012. The instruments considered at the five selected sites are identified in Table 1. GRUAN is establishing a database of ECV measurements from the different techniques and instruments, including full characterization of the uncertainty budget (random and bias contributions). The added value of GRUAN products is related to the implementation of data processing including several corrections for spurious effects on the radiosonde measurements and therefore the fidelity of the long-term records of radiosondes used for climate applications (Immler et al., 2010; Immler and Sommer, 2010). At present, only quality-assured measurements obtained by RS92-SGP sondes is flowing into the GRUAN data archive. Unfortunately, the approach presented in this paper cannot be used to show the advantages of using GRUAN sonde products, mainly because the bias component of the total uncertainty budget cannot be quantified through the entropy analysis presented here.

Water vapor measurements from sensors not considered in this study are also available for the considered sites (as noted in Table 1); they are a subject for future study. The current water vapor measurements were selected according to data availability for

AMTD
7, 6327–6357, 2014

each site. A similar investigation could be performed for other ECVs. For coherency, we used sonde data processed at each site rather than GRUAN products, which are still not available at all sites and for all radiosondes types. Moreover, retrieval algorithms for passive instruments usually take advantage of historical radiosonde data sets as a statistical constraint.

Simultaneous data from all available instruments were selected according to the conditions of clear sky (per lidar measurements or radiosonde humidity), nighttime, and, if lidar data are available, a relative error of lidar water vapor mixing ratio at 7 km a.g.l. $< 25\,\%$. This error is considered a good compromise having an adequate lidar signal-to-noise ratio and also covering the part of the troposphere where most of the water vapor can be observed. Raman lidar measurements are integrated over 10 min around the sonde synoptic launch time to keep a good signal-to-noise ratio in the investigated region, and MWR and microwave profiler (MWP) measurements are provided every 10 min. GPS data are provided only every 15 min, because of constraints on data processing at the considered sites. The measurements closest to the sonde launch time (within 10 min) are considered for the comparison. The use of MWR to calibrate the ARM Raman lidar measurements affects the independence of the IWV comparison for lidar at the SGP; in contrast, at PAY and POT the Raman lidar is calibrated by using radiosonde humidity profiles in the lower troposphere (Madonna et al., 2011; Brocard et al., 2013).

Data from different sites are currently processed with different algorithms; this could affect the comparison. However, the study of entropy is also a good check for the effect of retrieval inconsistencies. A linear regression of the entire time series (3 years) of IWV data and vertical profiles of water vapor mixing ratio at the altitude levels removed natural or artificial trends (e.g., calibration drifts). This was done to suppress the bias component of the time series uncertainty. Therefore, the reported entropies will be related only to the random uncertainty. MC quantifies the part of the random uncertainty that can be reduced by using ancillary information.

AMTD
7, 6327–6357, 2014

2.3 Optimal binning choice and minimally sufficient data

The two crucial issues need to be considered for entropy calculation using the histogram of a variable are the minimal quantity of data required to reduce inaccuracies in the calculation and the choice of the optimal binning to represent the actual Probability Density Functions (PDFs) of the variable.

To make our histogram representative of the real underlying PDF of the variable and to calculate the related entropy, a minimal number of data points is needed. The data sets considered here include >140 cases per station (LIN = 296, PAY = 174, ARM = 144). The cases were selected by the stations, according to their quality assurance criteria. For POT and SOD, more restricted data sets (40 and 22 cases, respectively) were used, because of the unique sampling strategy at POT (one radiosonde launch per week, only in clear sky) and the limited number of cryogenic frost point hygrometer (CFH) launches made available by SOD for this study. Knuth (2013) reported that at least 100 cases should be considered to avoid underestimation of entropy, though the number might depend on the underlying distribution. Nevertheless, values of the entropy calculated for POT and SOD are quite similar to those reported for other sites. This is encouraging, though a margin of inaccuracy affecting the values can be quantified only if larger data sets become available for the specific instruments at both stations.

To determine the optimal binning, several statistical methods have been proposed (Knuth, 2013). In Fig. 1, entropy is shown as a function of the number of bins used to build the histogram for the PAY radiosonde data sets. The value of entropy increases up to 0.81 for a histogram with 100 bins. Starting at 25 bins, entropy tends to assume asymptotic behavior. In this work, in view of the behavior shown in Fig. 1 and the number of data points available, 50 bins per histogram are used.

AMTD
7, 6327–6357, 2014

3 Results and discussion

In this section, normalized entropy, MC, and conditional entropy are presented for the data sets (and instruments) identified in Table 1. Both quantities were calculated to quantify uncertainty and redundancy in the IWV time series, as well as in the times series of the vertical profile of water vapor mixing ratio. In this investigation of time series of atmospheric water vapor measurements, entropy includes all contributions affecting the uncertainty of a measurements time series – sampling uncertainty, uncertainty due to the time and vertical average, atmospheric variability, and all other relevant environmental factors (Kitchen, 1989; Fassò et al., 2013), such as solar radiation affecting daytime in situ soundings.

Figure 2 (left) is an example of the time series of the IWV for the LIN instruments (Table 1), while Fig. 2 (right) shows the corresponding histograms of the time series. After linear detrending of the time series described above, the histograms were used to calculate entropy and MC. The shape of the histograms in Fig. 2 clarifies both how outliers can occur by chance in any distribution, often indicating either measurement errors or a heavy-tailed distribution in the population, and also the absence of any guarantee that the distribution will be a normal one. The discrepancies between the time series reported in Fig. 2 (left) do translate into a sort of bi-modal distribution characterized by a high kurtosis (Fig. 2, right). Thus, caution is needed in assuming a normal distribution; statistics, like entropy, that are robust to outliers and independent on the underlying distribution are more reliable for characterizing the uncertainty of a time series.

3.1 Normalized entropy for integrated water vapour and vertical profiles

Figure 3 compares the normalized entropies $H/\log n$, where n is the number of states (histogram entries) retrieved for all instruments measuring IWV at the LIN, PAY, POT, and SGP sites. For LIN, PAY, and SGP at least four instruments are available; for POT, GPS IWV is available only from June 2011 and thus is not included in this study.

AMTD
7, 6327–6357, 2014

For the numbers of available measurements and the atmospheric variability over the different stations (because entropy decreases as the number of measurements increases), there are no large differences in the uncertainties in IWV measurements with the techniques considered. Except for PAY, lidar entropy is the closest to radiosonde entropy, whether calibrated by using the sonde itself or the MWR. Moreover, at PAY the lidar offers the lowest entropy of the instrument ensemble. At SGP, GPS has the lowest entropy, though the values for all considered instruments are quite close. Similarly, at LIN, where the MWR has the lowest entropy, all values are close. At POT, the lowest entropy value is for the MWP. As a whole, the differences in entropy of the time series among the different instruments are within 8 %. Obviously, the different atmospheric variability of each site can also result in large deviations between entropy values. This deviation could be smoothed if a longer temporal data set was investigated. Moreover, differences in the observation techniques and their experimental implementation (e.g., different measurement angles and fields of view) might also contribute to differences in the calculated entropies and to non-linear calibration drifts.

3.2 Mutual correlation and distance for integrated water vapor and vertical profiles

The statistical distance D, as defined in Eq. (5), is a dimensionless measure of the similarity of pairs of data points, data clusters or time series. Figure 4 compares the distances between the IWV time series from all instruments with the radiosonde series at the LIN, PAY, POT, and SGP sites. The plot reveals very different results at different sites. In terms of the best performance of each instrument at the different sites, MWR and MWP have the highest redundancy and therefore the highest potential to reduce the uncertainty of the radiosonde IWV time series, with lidar and GPS following. The distances from the radiosonde series are > 0.18 for lidar, > 0.32 for GPS, > 0.14 for MWR, and > 0.28 for MWP. At PAY and POT, all the techniques show good redundancy, though GPS IWV at POT is not included in the statistics, because the number of measurements is small for the considered period. However, criteria are needed to de-

AMTD
7, 6327–6357, 2014

termine the acceptable levels of uncertainty and redundancy for a climate observation network. Section 3.5 deals with this aspect in more detail.

Normalized entropy and MC are compared for the available measurements of the water vapor or relative humidity (RH) vertical profiles in Fig. 5, which compares vertical profiles of distance for the Raman lidar (RL) and Atmospheric Emitted Radiance Interferometer (AERI) water vapor profiles with respect to radiosonde (RS92) profiles at the SGP site. Lidar profiles were retrieved by integrating signals over 10 min around the sonde launch time. The AERI profiles were averaged in the same time window. To improve the comparison among in situ, active and passive remote sensing measurements, the profiles from the three instruments were averaged over a vertical range of 1 km. This should strongly reduce the differences related to instrument signal-to-noise ratio and to the effective vertical resolution, which differs for the different techniques. Moreover, for the AERI, the statistical retrieval provided by the ARM Archive was considered; this retrieval is based on the radiosonde profile as a first guess, which affects the calculation of distance. Nevertheless, the comparison is provided to test the approach for passive profiling retrievals. Figure 5 shows that, though the difference is small, AERI has lower values of distance along the entire profile, probably because of the use of collocated radiosonde data as first guesses in the retrieval algorithm.

Figure 6 (left panel) compares the entropies retrieved for the RH profiles provided by RS92 radiosondes, Intermet radiosondes (I-Met), and CFH measuring in situ water vapor vertical profiles at SOD. Figure 6 (right panel) shows the profiles for one case on 15 March 2010. Entropies have been smoothed to obtain an effective vertical resolution of 60 m. Because 20 simultaneous profiles are included in this comparison, the calculated entropy values might underestimate the real uncertainty for the sensors. A larger data set should be considered for a full assessment of the differences in entropy for the various in situ measurements; this will be considered in future work, taking advantage of the data set available in the GRUAN network at the Boulder and LIN sites. The comparison in the left panel of Fig. 6 reveals that the entropy values for all sensors along the entire vertical profiles from the ground to 10 km are similar. This observation

AMTD
7, 6327–6357, 2014

indicates that RS92 and I-Met in situ measurements of atmospheric water vapor are affected by the same probabilistic uncertainty as the CFH, considered the reference in situ profiling instrument (Suortti et al., 2008). The comparison of RH profiles in Fig. 6 (right) shows good agreement between RS92 and CFH, with a small bias affecting the RS92 in the free troposphere (Wang et al., 2013). On the other hand, the I-Met sondes are able to reproduce the vertical variability of the RH profile, but they are affected by a negative bias. Moreover, above 8 km the I-Met behavior suddenly changes, overestimating RH. The comparison results are in agreement with literature values (WMO, 2010). The differences reported for the two sonde types are related to systematic effects on the RH profiles that contribute to the total uncertainty budget. This contribution can, in principle, be modeled and removed, but because of its systematic nature it cannot be evaluated with the entropy analysis discussed here. The presented analysis allows us only to state that the RH time series measured by the RS92, I-Met, and CFH show similar random uncertainty at all altitude levels below 10 km.

3.3 Conditional entropy

The conditional entropy quantifies the amount of information needed to describe the outcome of a random variable Y, given the value of another random variable X. The conditioning usually reduces entropy. That is, given two time series X and Y, the conditional entropy $H(X|Y) \leq H(X)$. Equality occurs only if X and Y are fully independent. Figure 7 shows the values of conditional entropy retrieved for all possible combinations of instruments measuring IWV at the SGP (left panel) and POT (right panel) sites, for the data sets described above. In both plots, the values of the normalized entropies calculated for each single instrument are also reported as a comparison term to quantify the residual uncertainty affecting each instrument when one or more other instruments are assumed as good constraints. Figure 7 shows that the residual entropy obtained by conditioning one instrument with a second instrument is 30–40 % lower than the entropy obtained for a single instrument. If two instruments are used for the conditioning, the residual entropy ranges between 5 % and 20 %. This finding indicates that with

AMTD
7, 6327–6357, 2014

reliable constraints, the entropy can be reduced by about 60–65 % with respect to the use of a single instrument. The minimum residual uncertainties are obtained when the GPS is conditioned with the RL and the MWR at the SGP site, and when the RL is conditioned with the MWP at the POT site.

These results also show that the residual uncertainties obtained with two conditioning constraints (two instruments) can be better than or similar to the value with only one instrument as a constraint. This is relevant when synergetic products must be defined and retrieved by using algorithms that can integrate information from ground-based or satellite sensors. This is the case for all optimal estimation algorithms based on the Bayes' Theorem, which is frequently adopted to improve atmospheric profiling. To quantify the effective advantages of integration, the presented analysis can be performed in advance of the elaboration of algorithms integrating measurements from different sensors. Moreover, conditional entropy can be applied similarly for directly measured quantities, like radiances, as well as for data products such as water vapor ground-based remote sensing. This is the case for algorithms making use of satellite measurements from polar and geostationary satellites to improve the resolution or reduce the uncertainty affecting the estimation of ECVs, but it is also true for algorithms merging satellite and ground-based passive sensor data to improve atmospheric profiling.

3.4 Linear mutual correlation

A comparison between MC and linear mutual correlation (LMC) provides a qualitative test for the non-linearity of the investigated problem. The plot in Fig. 8 shows a comparison between MC and LMC for the lidar and radiosonde at POT. In this case, both MC and LMC are normalized over the maximum entropy (the total number of entries in the joint histogram). The LMC underestimates the correlation between the two variables by about 10 %. Above 5.5 km, the LMC overestimates the correlation by about 8 %, most likely because of the presence of outliers in the PDF. This example supports the use of MC as more accurate than the LMC for quantifying the value of redundant

AMTD
7, 6327–6357, 2014

measurements. This is due to MC's capability to account for the higher-order terms in the PDF. The result is in agreement with outcomes from previous studies that analyzed data sets including different types of data and compared Taylor's diagrams built by using standard deviation vs. correlation and entropy vs. MC (e.g., Correa and Lindstrom, 2012).

3.5 Redundancy criteria

The analysis above shows how to approach the problem of quantifying measurement redundancy by using the concepts of information theory. However, the usefulness of this approach can be clarified only if some criteria are identified to classify when two data sets are redundant. This obviously depends on the investigated variable and on the uncertainty limits assumed to be minimum requirements for studying a certain atmospheric process or climate trend.

Here, we present an example showing the relationship between distance values and the random uncertainty affecting IWV measurements. The aim is to clarify the use of MC and the related distance for quantifying redundant IWV measurements at GRUAN sites. The plot of Fig. 9 shows the distance between the radiosonde IWV time series at LIN and the corresponding time series obtained by adding variable random noise to the radiosonde time series. The random noise is added to reproduce the effect of an additional random uncertainty, with relative values of 0–100 % affecting an IWV time series with respect to the reference series. For example, a distance value lower than about 0.2 corresponds to a random uncertainty 20 % larger than that of the original time series assumed as the reference. This example indicates a very simple way to approach data sets from different instruments or techniques, fixing a threshold consistent with the desired redundancy requirements. According to the GCOS requirements for the state-of-art capability, also reported in the GRUAN manual (http://www.wmo.int/pages/prog/gcos/publications/gcos-171.pdf), atmospheric water vapor must be measured with a random error $< 5\,\%$ in the entire troposphere and stratosphere. This corresponds to a maximum random error $< 5\,\%$ affecting an

AMTD
7, 6327–6357, 2014

IWV time series. If the random uncertainty is quantified by using entropy and the radiosonde IWV time series (the reference) is affected by random errors < 5 %, an IWV time series affected by a random error < 5 % is consistent with the true series if the corresponding distance value is lower than about 0.2 (total random error < 10 %). The plot in Fig. 4 shows that the distance values for different instrument and different sites do not always meet this standard. The values > 0.2 should be classified as not redundant in terms of the threshold of 5 % random error affecting the two compared time series.

4 Conclusions

The ultimate aim of this study is to recommend the best combination of instruments for monitoring atmospheric water vapor. Though entropy and MC are robust concepts provided in information theory, representing appropriate metrics to quantify the uncertainty and redundancy of atmospheric measurements, they have never been applied extensively to climate data. In this paper, we show how entropy and MC can be used to evaluate the random probabilistic uncertainty in the ECV by analyzing measurement redundancy.

The following conclusions can be drawn from the results of this study of data sets of water vapor from five GRUAN observation stations in 2010–2012:

1. The random uncertainty in the IWV time series obtained with the different instruments considered in this study (Raman lidar, GPS, MWR, MWP, sondes) differs by < 8 %.

2. In terms of the best performances for each instrument at the different sites, the comparison of IWV time series showed that MWR and MWP have the highest redundancy and therefore the highest potential to reduce the random uncertainty of IWV time series as measured by radiosondes.

3. The distance between the time series of water vapor profiles at each altitude level has been also performed to show how to evaluate the redundancy of collocated

AMTD
7, 6327–6357, 2014

in-situ, active and passive profiling instruments, though for passive instruments this also depends on the retrieval algorithms and on which first-guess prior covariance is used.

4. Both RS92 and I-Met radiosondes can measure in situ atmospheric water vapor with the same random uncertainty as the CFH, though the sondes are affected by a bias error that cannot be evaluated with the present approach.

5. A conditional entropy analysis showed that conditioning of the time series with more than one instrument, assumed as constraints, can decrease the residual entropy by at least 60 % vs. the use of one conditioning instrument. Moreover, the use of two conditioning instruments vs. one results in similar or slightly better residual uncertainty.

6. An analysis of the relationship between distance and the random uncertainty showed that a maximum random error $< 5\,\%$ affecting the IWV estimated by two different techniques corresponds to a distance value less than about 0.2. That is, an IWV time series whose distance from a reference time series (i.e., IWV measured by radiosondes) is > 0.2 exceeds the redundancy limits identified according to the GCOS criteria.

Final recommendations can be provided only if criteria, to support a certain network, are clearly defined according to the uncertainty thresholds assumed in the study of an ECV; however, the presented approach is versatile enough to be used with different data sets, stations, and instruments to provide the required feedback in terms of uncertainty and use of redundant measurements to reduce uncertainty in ECV values. Moreover, the entropy and MC analysis can be used for a preliminary feasibility study of the effective advantages in using retrieval algorithms integrating measurements provided by different observation platforms, for both ground-based and satellite measurements and products, including both direct measurements (e.g., radiances) and retrieved products (e.g., temperature, water vapor content, aerosol optical depth).

AMTD
7, 6327–6357, 2014

Acknowledgements. The financial support of ACTRIS Research Infrastructure Project funded by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n. 262254. Data sources were as follows: ARM SGP data through the US Department of Energy (www.arm.gov); CIAO Potenza data through CNR-IMAA (http://www.ciao.imaa.cnr.it); LIN data through the Lindenberg Meteorological Observatory, Richard Assmann Observatory, Deutscher Wetterdienst (http://www.dwd.de); PAY data through the Federal Office of Meteorology and Climatology MeteoSwiss (www.meteoswiss.ch); SOD data through the Finnish Meteorological Institute (http://www.fmi.fi). The authors also gratefully acknowledge the useful comments of A. Fassò from University of Bergamo. This work was supported by the US Department of Energy, Office of Science, Office of Biological and Environmental Research, under contract DE-AC02-06CH11357.

References

Adam, W., Dier, H., and Leiterer, U.: 100 years aerology in Lindenberg and first long-time observations in the free atmosphere, Meteorol. Z., 14, 597–607, 2005.

Brocard, E., Philipona, R., Haefele, A., Romanens, G., Mueller, A., Ruffieux, D., Simeonov, V., and Calpini, B.: Raman Lidar for Meteorological Observations, RALMO – Part 2: Validation of water vapor measurements, Atmos. Meas. Tech., 6, 1347–1358, doi:10.5194/amt-6-1347-2013, 2013.

Calpini, B., Ruffieux, D., Bettems, J.-M., Hug, C., Huguenin, P., Isaak, H.-P., Kaufmann, P., Maier, O., and Steiner, P.: Ground-based remote sensing profiling and numerical weather prediction model to manage nuclear power plants meteorological surveillance in Switzerland, Atmos. Meas. Tech., 4, 1617–1625, doi:10.5194/amt-4-1617-2011, 2011.

Correa, C. and Lindstrom, P.: The mutual information diagram for uncertainty visualization, International Journal for Uncertainty Quantification, 3, 187–201, 2012.

Cover, T. M. and Thomas, J. A.: Elements of Information Theory, John Wiley & Sons, Inc., 1991.

Fassò, A., Ignaccolo, R., Madonna, F., and Demoz, B. B.: Statistical modelling of collocation uncertainty in atmospheric thermodynamic profiles, Atmos. Meas. Tech. Discuss., 6, 7505–7533, doi:10.5194/amtd-6-7505-2013, 2013.

Hirsikko, A., O'Connor, E. J., Komppula, M., Korhonen, K., Pfüller, A., Giannakaki, E., Wood, C. R., Bauer-Pfundstein, M., Poikonen, A., Karppinen, T., Lonka, H., Kurri, M., Heinonen, J., Moisseev, D., Asmi, E., Aaltonen, V., Nordbo, A., Rodriguez, E., Lihavainen, H., Laaksonen, A., Lehtinen, K. E. J., Laurila, T., Petäjä, T., Kulmala, M., and Viisanen, Y.: Observing wind, aerosol particles, cloud and precipitation: Finland's new ground-based remote-sensing network, Atmos. Meas. Tech., 7, 1351–1375, doi:10.5194/amt-7-1351-2014, 2014.

Immler, F. J. and Sommer, M.: GRUAN-TD-4 Brief Description of the RS92 GRUAN Data Product (RS92.GDP), available at: http://www.gruan.org (last access: 20 June 2014), 2010.

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, Atmos. Meas. Tech., 3, 1217–1231, doi:10.5194/amt-3-1217-2010, 2010.

Kitchen, M.: Representativeness errors for radiosonde observations, Q. J. Roy. Meteor. Soc., 115, 673–700, 1989.

Knuth, K. H.: Optimal data-based binning for histograms, arXiv preprint physics/0605197, 2013.

Madonna, F., Amodeo, A., Boselli, A., Cornacchia, C., Cuomo, V., D'Amico, G., Giunta, A., Mona, L., and Pappalardo, G.: CIAO: the CNR-IMAA advanced observatory for atmospheric research, Atmos. Meas. Tech., 4, 1191–1208, doi:10.5194/amt-4-1191-2011, 2011.

Majda, A. and Gershgorin, J.: Quantifying uncertainty in climate change science through empirical information theory, P. Natl. Acad. Sci. USA, 107, 14958–14963, 2010.

Miller, M. A., Johnson, K. L., Troyan, D. T., Clothiaux, E. E., Mlawer, E. J., and Mace, G. G.: ARM value added cloud products: description and status, in: Proc. of the 13th ARM Science Team Meeting, Broomfield, CO, ARM, available at: http://www.arm.gov/publications/proceedings/conf08/extended_abs/miller_ma.pdf (last access: 20 June 2014), 2003.

Suortti, T. M., Kivi, R., Kats, A., Yushkov, V., Kämpfer, N., Leiterer, U., Miloshevich, L. M., Neuber, R., Paukkunen, A., Ruppert, P., and Vömel, H.: Tropospheric comparisons of Vaisala radiosondes and balloon-borne frost point and Lyman-α hygrometers during the LAUTLOS-WAVVAP Experiment, J. Atmos. Ocean. Tech., 25, 149–166, doi:10.1175/2007JTECHA887.1, 2008.

Thorne, P. W., Vömel, H., Bodeker, G., Sommer, M., Apituley, A., Berger, F., Bojinski, S., Braathen, G., Calpini, B., Demoz, B., Diamond, H. J., Dykema, J., Fassò, A., Fujiwara, M., Gardiner, T., Hurst, D., Leblanc, T., Madonna, F., Merlone, A., Mikalsen, A., Miller, C. D., Reale, T., Rannat, K., Richter, C., Seidel, D. J., Shiotani, M., Sisterson, D., Tan, D. G. H., Vose, R. S., Voyles, J., Wang, J., Whiteman, D. N., and Williams, S.: GCOS reference upper air network (GRUAN): steps towards assuring future climate records, AIP Conf. Proc., 1552, 1042–1047, doi:10.1063/1.4821421, 2013.

Wang, J., Zhang, L., Dai, A., Immler, F., Sommer, M., and Vömel, H.: Radiation dry bias correction of Vaisala RS92 humidity data and its impacts on historical radiosonde data, J. Atmos. Ocean. Tech., 30, 197–214, doi:10.1175/JTECH-D-12-00113.1, 2013.
WMO/TD – No. 1580: Instruments and Observing Methods Report No. 107, edited by: Nash, J., Oakley, T., Vömel, H., and Li, W.; WMO Intercomparison of High Quality Radiosonde Systems, Yangjiang/China, 12 July–3 August 2010, 2010.

AMTD
7, 6327–6357, 2014

Quantifying the value of redundant measurements at GRUAN sites

F. Madonna et al.

AMTD
7, 6327–6357, 2014

Table 1. Instruments available (and model when applicable) at the GRUAN sites generating data sets considered in this study of uncertainty and redundancy. Simbol X indicates that the instrument is available and used for this study; while symbol ! indicates that the instrument is available at the site, but the data were not used in the study. Abbreviations: CFH, Cryogenic Frost-point Hygrometer; MWR, Microwave Radiometer; MWP, Microwave Profiler; GPS, Global Positioning System; FTIR, Fourier Transform Infrared Radiometer; AERI, Atmospheric Emitted Radiance Interferometer.

GRUAN Site/Instrument	Sonde	CFH	Lidar	MWR	MWP	GPS	FTIR
Lindenberg	RS-92 (4× day)	!	X	Radiometrics	Radiometrics	GFZ	
Payerne	SRS 400 (2× day)		X	HATPRO		GFZ	
Potenza	RS-92 (1× week)		X		Radiometrics	!	
Sodankyla	RS-92, I-Met1 RSB	X	!			!	Bruker
Southern Great Plains	RS-92 (4× day)		X	Radiometrics		Suominet	AERI

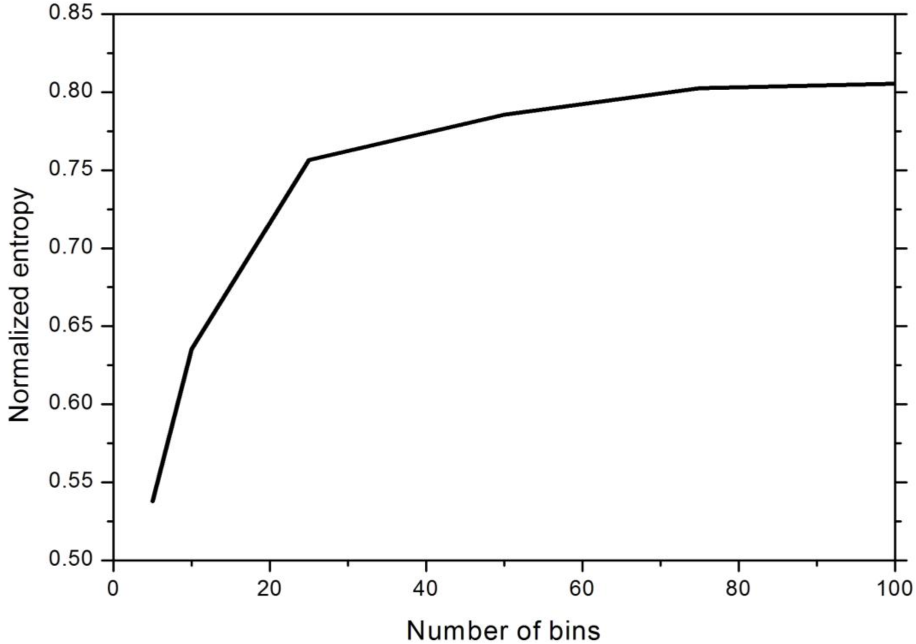

Figure 1. Entropy as a function of number of bins used to build the histogram for one of the datasets used in this study.

AMTD
7, 6327–6357, 2014

AMTD
7, 6327–6357, 2014

Quantifying the value of redundant measurements at GRUAN sites

F. Madonna et al.

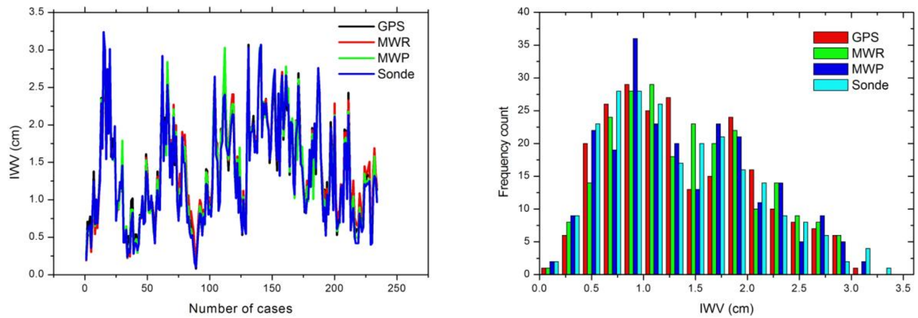

Figure 2. Example of the time series (left) of integrated water vapor obtained with the instruments available at the Lindenberg site (reported in Table 1) and histograms (right) of the frequency count of the time series shown in the left panel. After detrending of the time series, the histograms were used to calculate entropy and mutual correlation.

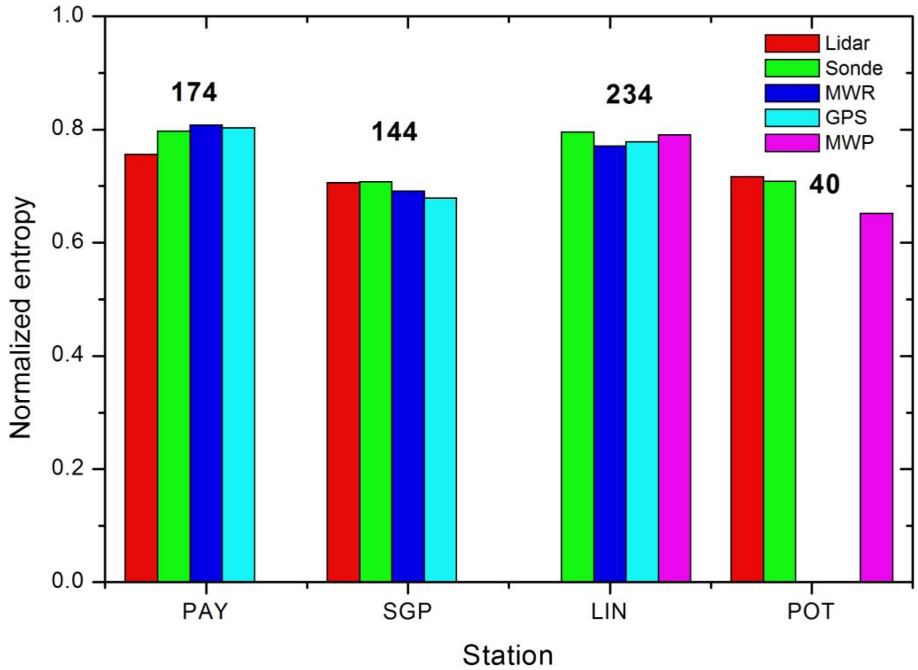

Figure 3. Comparison of the normalized entropy retrieved for the instruments measuring integrated water vapor at the Lindenberg (LIN), Payerne (PAY), Potenza (POT), and Southern Great Plains (SGP) sites. The data set considered includes all available measurements in 2010–2012. The numbers above the bars represent the number of cases selected, according to the quality assurance criteria for each station.

AMTD
7, 6327–6357, 2014

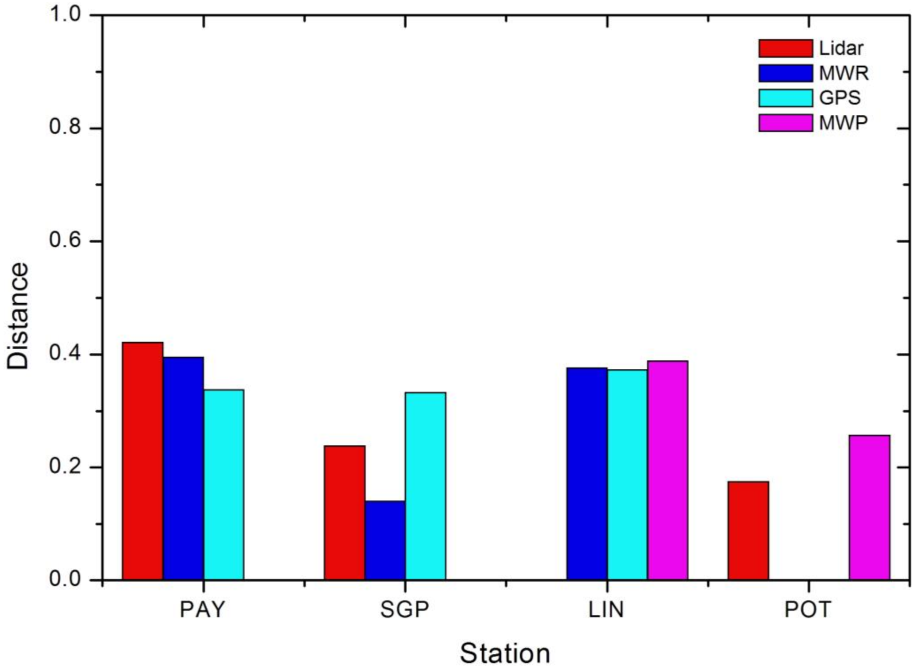

Figure 4. Comparison of the statistical distances between pairs of times series data retrieved for the instruments measuring integrated water vapor with respect to the time series obtained from the radiosondes at the Payerne (PAY), Southern Great Plains (SGP), Lindenberg (LIN) and Potenza (POT) sites.. The data set considered includes all available measurements in 2010–2012.

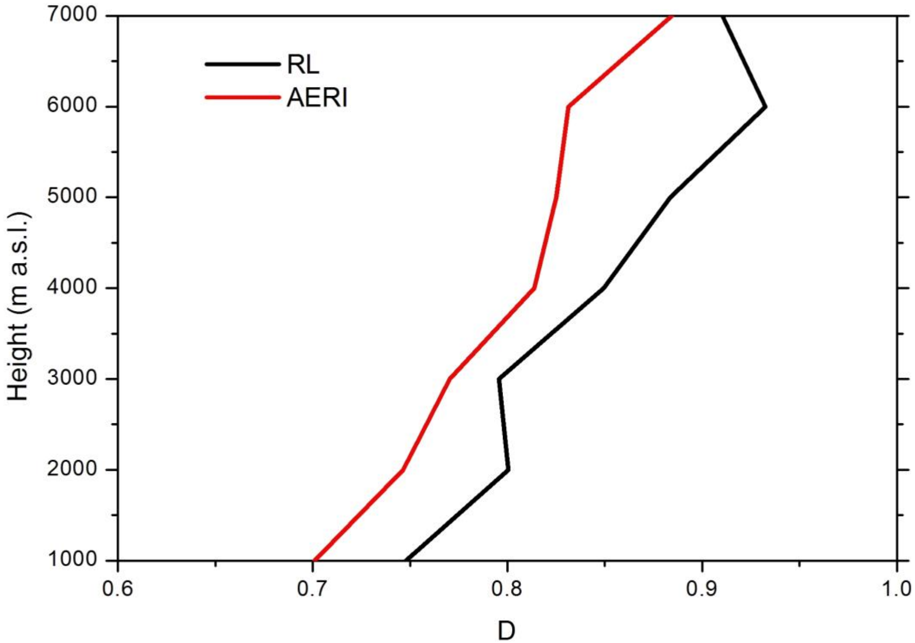

Figure 5. Comparison of the statistical distances retrieved for the three instruments (Raman lidar [RL], Atmospheric Emitted Radiance Interferometer [AERI], and RS92 radiosonde) measuring the water vapor vertical profile at the Southern Great Plains (SGP) site. The data set considered includes all measurements available at SGP (144 profiles) in the period 2010–2012.

AMTD
7, 6327–6357, 2014

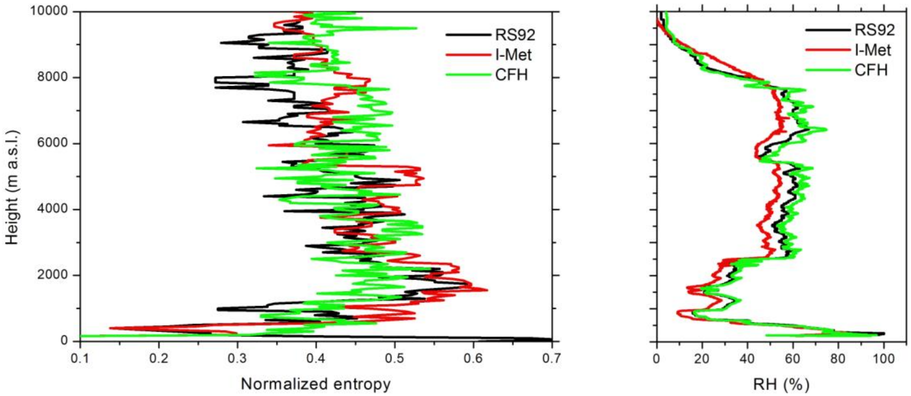

Figure 6. Comparison of the normalized entropy values (left) for the RS92 and Intermet Radiosondes (I-MET) and the Cryogenic Frost-point Hygrometer (CFH) measuring the in situ water vapor vertical profile at the Sodankyla site in 2010 (left panel); comparison of the relative humidity profiles in one case on 15 March 2010 (right panel).

AMTD
7, 6327–6357, 2014

Quantifying the value of redundant measurements at GRUAN sites

F. Madonna et al.

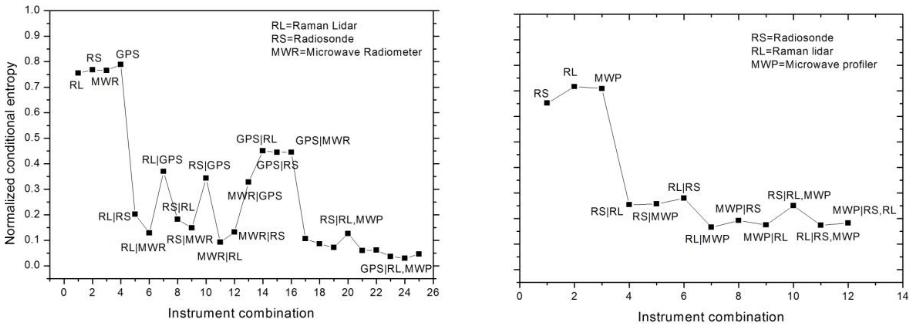

Figure 7. Comparison of the normalized conditional entropy values retrieved for all combinations of instruments measuring integrated water vapor at the Southern Great Plains site (left panel) and the Potenza site (right panel).

AMTD
7, 6327–6357, 2014

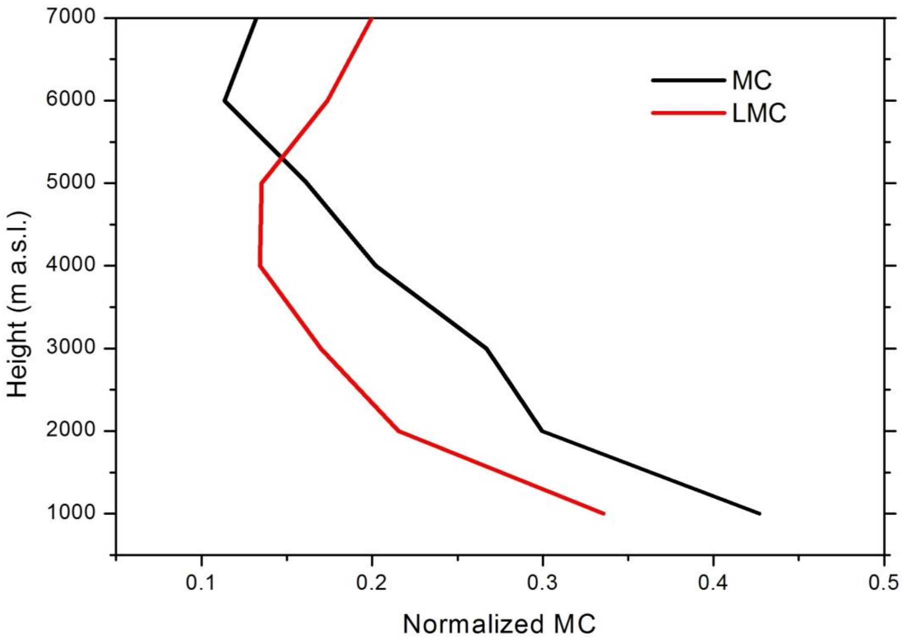

Figure 8. Comparison of the normalized mutual correlation for the Linear (LMC) and Non-Linear Cases (MC), calculated for the lidar and radiosonde data sets from the Potenza site.

AMTD
7, 6327–6357, 2014

Quantifying the value of redundant measurements at GRUAN sites

F. Madonna et al.

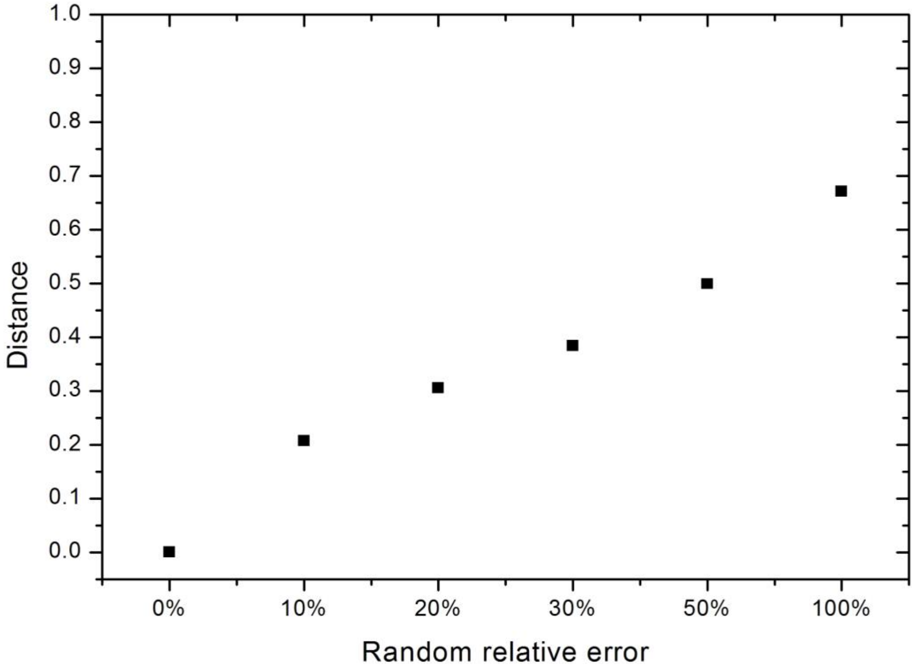

Figure 9. Statistical distance between the integrated water vapor time series retrieved from the radiosonde at the Lindenberg site and the corresponding time series obtained by adding random noise to the radiosonde time series to simulate the effect of increasing relative random uncertainty.

AMTD
7, 6327–6357, 2014