White Paper on GRUAN and Satellite Collocation criteria

Xavier Calbet, Tim Hewison, Tony Reale, Bomin Sun, John Dykema, Viju John, Nikita Pougatchev

EUMETSAT

Summary

Ground based in-situ, radiosondes or remote sensing measurements, as provided by GRUAN, due to their potential high measurement accuracy and high vertical resolution, can constitute valid reference measurements for Calibration and Validation (Cal/Val) of satellite observations and products, and particularly for infrared hyperspectral and microwave sounder instruments, the main subject dealt with in this paper, as this is one of the most demanding applications due to their high to moderate spectral resolution, moderate spatial resolution and higher to low sensitivity to cloud for infrared hyperspectral or microwave instruments respectively. Two critical issues have been identified to make these ground based measurements appropriate for Cal/Val of infrared hyperspectral space borne measurements: proper collocation of both sets of measurements and high accuracy in the measurements of the ground based instruments.

Adequate collocation depends strongly on the type of available ground based instruments, which will determine when and where the atmosphere is sampled relative to the satellite observations. The measurements from these instruments are usually combined in such a way as to obtain what is known as the best state of the atmosphere, such that it is directly comparable to the satellite observations. Typical collocation tools, which aim at integrating all these measurements into one atmospheric best state, range from various time and space interpolation or extrapolation to optimal estimation of the atmospheric state or the use of numerical weather forecast model fields to complement the measurements.

High accuracy observations are achieved by using state of the art sensors and data processing in the ground-based measurements. Typical ground based instruments are radiosondes, which measure with high accuracy and vertical resolution a small parcel of the atmosphere, i. e., assuming ascending velocity 5 m/sec and sampling rate every 2 sec. Among these, it is worth mentioning the Cryogenic Frost point Hygrometer (CFH) which has shown great accuracy for the measurement of humidity. They can be complemented with ground based remote sensing instruments, which typically provide a better time sampling at, potentially, the cost of lower accuracy or vertical resolution. This higher temporal sampling can be used either for homogeneous scene identification or for a better determination of the best state of the atmosphere.

As is common in physical sciences, measurement accuracy, error characterization, logistics and post-processing tools are complementary, in the sense that a suitably well-designed measurement procedure implies straightforward and easy post-processing and, on the other hand, a badly designed experiment can make the final data processing very difficult or, in the extreme cases, impossible. Many different measurement strategies can be devised, ranging from very sophisticated ones with many ground based instruments to more minimalistic ones comprised of, for example, just one radiosonde. The former could potentially provide a highly accurate best estimate of the atmosphere, while the latter would be very cost effective. Any of them could provide positive or negative results depending on its adequate design.

A minimalistic strategy that has provided good results in the past is to launch one RS92 sonde at satellite overpass time (± 30 minutes) within the satellite instrument field of view (e.g. ± 25 km). This strategy generally provides a difference between the infrared hyperspectral observed to the calculated (sonde data being input to a radiative transfer model) radiances within three sigma infrared hyperspectral instrument noise, which implies a small collocation and sonde measurement error as compared to the infrared hyperspectral instrument noise. A key component to this positive result is to process the sonde data adequately by correcting for all possible systematic error components, such as is done by the GRUAN processing.

A critically related subject is that of the estimation of the different type of errors committed when collocating reference measurements to satellite ones. Collocation uncertainty estimation can be key for measurement design and collocation strategies, as well as for conveniently selecting data for Cal/Val. To serve these

purposes, there is an effort currently under way which aims at estimating the collocation uncertainty on each individual case as a result of the "GRUAN-GSICS-GNSSRO WIGOS Workshop on Upper-Air Observing System Integration and Application".

In the remaining of this paper, the reasoning and justification behind these conclusions is described.

General Cal/Val Strategy

Calibration and Validation strategies, in general, follow a series of steps. These steps will be defined here serving as a reference for the rest of the document and will be explained in more detail in the following sections. The steps, listed in a logical sequential order, are:

1. **Collocation**. In this first step the reference measurements are collocated to the satellite observations. Usually a criteria based on proximity of both measurements in space and time is used.
2. **Pre-processing**. Before reference and satellite measurements are compared, a certain amount of pre-processing is usually needed. Interpolation or extrapolation of profiles to obtain a better co-location is standard practice. Possible corrections of the reference measurements also fall into this category, like for example sonde humidity bias corrections. Further selection criteria also fit into this category, such as restricting even further the time and space difference in the collocation, selecting homogeneous or clear sky scenes, etc.
3. **Consistency check**. This is a new proposed step, that is slowly becoming of pivotal importance in the Cal/Val process. It consists in comparing the observed radiances to the calculated ones. The latter is calculated by using the reference profile as input to a Radiative Transfer Model (RTM).
4. **Comparison**. Once the collocation and the quality of the reference measurement is well established, it can then be compared to the satellite observations. Conclusions can usually be drawn from these comparisons like retrieval error statistics, potential issues with the ground-based reference measurements, validation of the RTM, variation of standard deviation of differences with increased space collocation window, etc.

To highlight current solutions and issues with the steps involved in the Cal/Val strategy several examples are shown in the next sections.

Collocation

“How close is *close enough*?” This document is not going to give you a straight answer, because it’s not a straight question. “It depends” may not be a very satisfactory answer, but as we shall try to demonstrate, it depends not only on the uncertainty demands of the Cal/Val application, but also on the atmospheric conditions around the collocation. The natural scene variability, like that from the atmospheric profile, cloud or surface temperature/emissivity, often dominates the uncertainty of the comparisons, so limits are often placed on the maximum separation in space and time between the radiosonde and satellite measurements in the collocation.

Relatively relaxed collocation criteria are often adopted for comparisons with microwave instruments because of their broad field of view – e.g. 50 km/2 hr (Moradi *et al.,* 2013). Kitchen (1989) analysed UK radiosonde statistics to show that the radiosonde collocation errors of <1 K are achievable on spatial scales of 55 km, typical of the field of view of temperature sounding microwave radiometers. He also showed that temperature and humidity variations tended to propagate at speeds typical of advection on synoptic scales.

Generally, matching a sonde measurements to an individual satellite pixel is ambitious mainly because of the horizontal drift of sondes during their ascent and therefore some studies (e.g., Buehler et al., 2004) used a target area to represent a satellite measurement matching a sonde measurements. For example, an average of 10-30 pixels of microwave humidity sounders within a 50 km circle around the sonde station (which is normally the drift when a sonde reaches 200 hPa) has been used to match with simulated measurements from radiosonde profiles. The variability of these 10-30 pixels has also been used to weight the collocations, i.e., when the variability is less, atmosphere is more homogeneous and the collocation gets a higher weight when computing statistics and vice versa. They have also used a so called "displacement filter" which is calculated by multiplying the average wind in the free troposphere with the time difference between satellite and sonde measurement. When the displacement is larger than the size of the target area (50 km in case of studies mentioned above), such collocations are filtered out due to the fact that the satellite and the sonde have measured different airmasses.

John et al, 2005 and Moradi et al., 2013 have shown that these improved methods result in reducing the uncertainty in the bias estimates so that they can be used to deduce sensor dependent biases and specific biases such as radiation dry bias. Kottayil et al, 2012 have used similar methods to show the importance of GRUAN type corrections applied to RS92 humidity sensor measurements for a better agreement between satellite and sonde measurements.

The NOAA Products Validation System (NPROVS, Reale et al., 2012) supported by JPSS and operated at NOAA NESDIS office of SaTellite Applications and Research (STAR) provides routine data access, collocation and inter-comparison of multiple satellite temperature and moisture sounding product suites against global radiosonde and dropsonde observations and NWP products. The access and collocation of sonde and satellite data occur on a daily basis with the routine archiving of all collocated data at STAR. The minimum requirement for retaining a given sonde is that the collective temperature and moisture profiles extend vertically at least 5 km without a gap. Retained reports are further processed including analysis to identify suspicious observations, temperature and moisture profile features, and impacts when applying sampling constraints. The collocation approach is optimized for each satellite system to select a single ''closest'' sounding from each satellite that lies within 6 h and 150 km(250 km for GPS RO) of a given sonde. Furthermore, there is an expansion validation system named NPROVS+ which uses GRUAN and dedicated radiosonde data as the anchor to collocate satellite EDRs, but it also provides satellite radiance measurements within 500 km of the radiosonde for satellite retrieval algorithm development.

By analyzing radiosonde and collocated GPS RO data collected at NPROVS and taking into account of both radiosonde drift and GPS RO profile displacement, Sun et al. [2010] quantified the increase in standard deviation difference between GPS RO and radiosonde data with increasing spatial and temporal mismatches. They note that, globally, in the troposphere (850–200 hPa), the collocation mismatch impacts on the standard deviation differences for temperature are 0.35 K per 3 h and 0.42 K per 100 km and, for relative humidity, are 3.3% per 3h and 3.1% per 100 km, indicating an approximate equivalence of 3 h to 100 km in terms of mismatch impact, consistent with Kitchen (1989).

A comprehensive global climatology of radiosonde balloon drift distance and ascent time is presented in Seidel et al. (2011) with particular attention to GRUAN stations. Typical drift distances are a few kilometers in the lower troposphere, ~5 km in the mid troposphere, ~20 km in the upper troposphere, and ~50 km in the lower stratosphere, although there is considerable variability due to variability in climatological winds.

Conventional radiosonde temperature measurements in the troposphere are overall good for satellite Cal/Val, but they do appear to suffer radiation-induced biases in the upper troposphere and lower stratosphere. Sun et al. (2013) indicate that temperature measurement from May 2008 to August 2011 on average from most radiosonde types show a nighttime cold bias and a daytime warm bias relative to GPS RO dry temperature from the Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC). Most daytime biases increase with altitude and solar elevation angle (SEA). The global average biases in the 15–70 hPa layer are -0.05 ±1.89K standard deviation at night and 0.39 ± 1.80K standard deviation in daytime. Most conventional radiosonde humidity sensors tend to have a dry bias in the middle and particularly upper troposphere (5-8%) with daytime bias greater than the nighttime (Sun et al. 2010), but Russian sondes are an exception that show a wet bias of 2-3% in the mid-upper troposphere. Caution is needed for regional applications of sonde data. On the other hand, CFH instruments have shown a good performance for satellite Cal/Val (Calbet et al. 2011), in the sense of passing the consistency check described below.

Infrared hyperspectral instruments, like AIRS, IASI or CrIS, provide high spectral resolution measurements of Earth scenes as observed from space. These measurements can be converted, via what is known as Level 2 processing, into high accuracy and relatively high vertical resolution atmospheric profiles of temperature and humidity, which are usually referred as retrievals. The expected error in these retrievals is small enough such that collocation with ground based measurements has to be very precise. Pougatchev et al. (2009) have determined that, at least for the vicinity of the Lindenberg observatory, collocations of IASI and radiosonde data should not exceed on average the 30 min time difference and the 20 km spatial distance (Fig. 1). In the GRUAN intercomparison exercise (Calbet, 2014) the IASI fields of views where in the range of 25 km and 30 minutes from the radiosonde launch location.

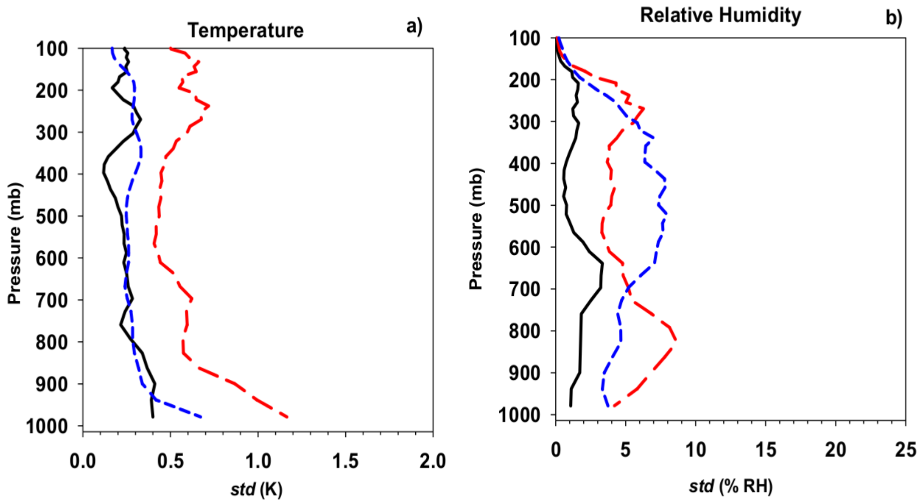

Figure 1: (a) Temperature retrievals and (b) Relative humidity retrievals. Square roots of diagonal elements of the covariance matrices of the: retrieval noise (solid black lines); temporal non-coincidence of 0.5 h (dashed blue lines); and spatial non-coincidence error of 20 km (dashed red lines). From Pougatchev et al. 2009.

There is considerable correlation between the spatial and temporal collocation criteria. Hewison (2013) found the spatial and temporal variability of radiances observed by an infrared geostationary imager to be comparable on scales between 5 km/5 min and 100 km/100 min (Fig. 2). These results suggest the variability on small scales is propagated with speeds much faster than those typical for advection, because of the influence of rapidly changing convective cloud in the tropics.

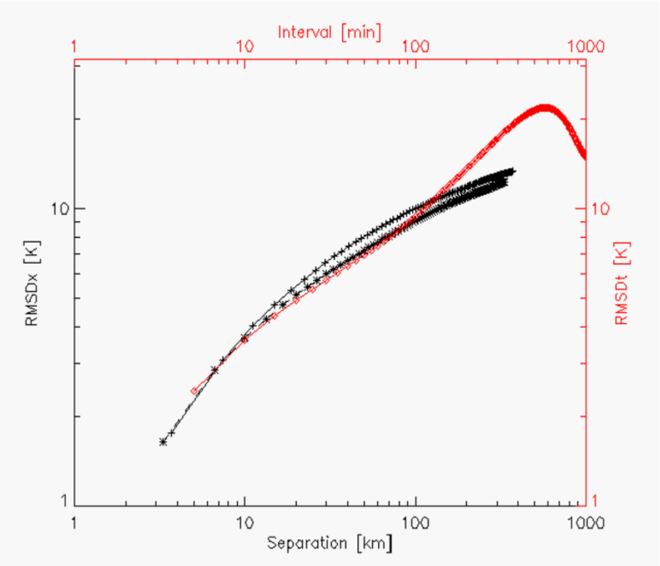

Figure 2- Variograms calculated as RMS differences in Meteosat-8/SEVIRI 10.8 µm brightness temperatures with time intervals from Rapid Scanning data (red diamonds, upper x-axis) and with spatial separation in North-South direction (black pluses, lower x-axis) and West-East direction (black stars, lower x-axis).

The above figures are statistical averages that can be applied in general, but care must be taken in special atmospheric circumstances when the atmosphere is quickly changing. Furthermore, systematic biases could be introduced if the observations are always made at a fixed observation time. For example, if radiosondes are always launched 30 min. before to satellite's morning overpass time, the surface and atmosphere will typically warm during this period.

Pre-processing

Some processing is usually needed before the reference measurements are compared to the satellite observations. Raw reference measurements sometimes have systematic errors which should be dealt with before using the data. It is also possible to restrict the number of reference atmospheric profiles measurements by selecting the ones which have been measured in more homogeneous or little varying atmospheric conditions, like for example, by means of the selecting the sondes which while airborne had small variability in accompanying ground-based GPS measurements. A very useful option for collocating two consecutive sonde measurements for a given satellite observation, is to interpolate them (Tobin et al. 2006). A simple solution is to find the closest in time and space field of view to the ground-based measurement.

In the case of the GRUAN intercomparison exercise (Calbet, 2014), where GRUAN data from Manus island is compared to IASI radiances, the radiosonde data was used straight away as it is delivered by GRUAN. The GRUAN processing already takes into account any systematic errors present in the RS92 measurements and no further processing is needed.

It is important to note that all methods used in the pre-processing step should have a sound physical basis. When further selecting the data, the applied criteria should be chosen to avoid as much as possible bias sample selection. All data should be treated with exactly the same procedures avoiding ad-hoc methods for particular fields of view. In the GRUAN intercomparison exercise the nearest field of view to the radiosonde launch location had to be selected (±25 km), a bigger spatial window did not provide useful results, as dictated by the consistency check.

When overpasses occur coincident with ground sites that supplement radiosonde launches (whether synoptic or dedicated) with continuous atmospheric profiling from ground-based remote sensing, it is possible to compensate for space and time mismatches with a variety of methods. One method that has been utilized to good effect at several sites (Lamont, Manus, Barrow) is the site atmospheric state best estimation (SASBE) method demonstrated first for AIRS (Tobin et al., 2006) and more recently for CrIS

(Nalli et al., 2013). The SASBE method can compensate for both temporal and spatial mismatch by using different classes of observations. The temporal compensation was initially obtained from retrievals from uplooking hyperspectral infrared sounders (Atmospheric Emitted Radiance Inteferometer; Knuteson et al., 2004a, 2004b, Feltz et al. 2003). This technique is easily generalized to other profiling instruments, such as Raman LIDAR for water vapor or microwave profilers for temperature and water vapor (Dykema et al., 2014).

Consistency check

A determination of the collocation and reference measurement error can be made on an individual field of view basis by comparing the hyperspectral observed radiances with the calculated ones (Calbet et al. 2011). Calculated radiances, in this context, are the radiances obtained from feeding the reference ground based measurement as input to a radiative transfer model (sonde-to-satellite approach which eliminates errors introduced during the ill-posed inversion of satellite measurements to obtain atmospheric profiles). If the observed minus calculated radiance difference exceeds the three sigma infrared hyperspectral instrument noise in the spectral regions of interest, the collocation and reference measurement error can quickly become much bigger than the random component of the retrieval error due to radiometric noise (Fig. 3 and 4), therefore rendering the validation of little value (Calbet 2012). In the extreme case where the time difference and spatial separation collocation criteria are too lax, the statistics of the comparison of the retrievals with the reference measurements will mostly reflect the collocation difference, being the retrieval accuracy determination when comparing to the reference profile negligible in the global comparison. The verification that the observed minus calculated radiances remain within three sigma of the hyperspectral instrument noise in the spectral region of interest is therefore truly a consistency check between the reference profile and the hyperspectral measurements. Seen from another point of view, if the difference between the observed and calculated radiances is big, it will be nearly impossible to obtain a retrieval that is similar to the reference profile. This consistency check provides a great confidence that the collocation is adequate and that the accuracy of the reference profile measurements is high enough for the validation. It then follows that the consistency check should be performed after the collocation of the observations is made and before making the comparison of the reference profiles (Calbet et al. 2013).

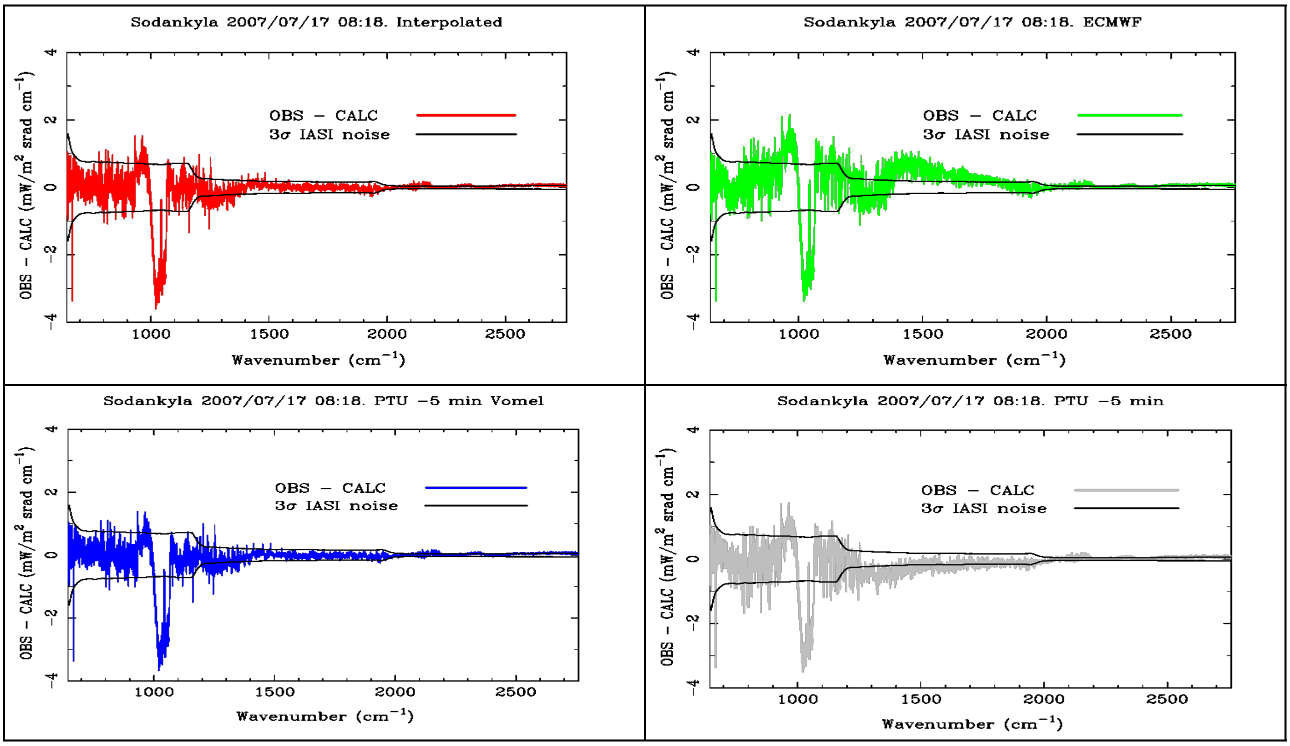

Figure 3: IASI observed minus calculated radiances (OBS-CALC) for the "Interpolated" (upper left in red), "ECMWF'" (upper right in green), "PTU -5 min"(lower right in grey) and "PTU -5 min Vomel" (lower left in blue) profiles. From Calbet (2012).

Ideally, the consistency check should not be part of the field of view selection in the pre-processing step. In this way, the logic behind the pre-processing is kept under physical arguments and the consistency check can be used as a test for the methods used in the pre-processing. In this way, the consistency check can constitute a powerful tool to determine whether pre-processing is effective for the given measurements. Unfortunately, this is not always possible and sometimes the consistency check needs to be used as a selection criteria even running the risk of introducing a bias selection in the final validation sample.

One main drawback for the consistency check which applies to hyperpsectral infrared sounders is that it can only be applied in clear sky situations because the radiative transfer models and cloud characterization are not mature enough for cloudy scenes. This would not be the case for the microwave sounders which would only have limitations for the consistency check under scenes with rain or thick cirrus clouds. Another drawback of this step is that undetected clouds for hyperspectral infrared instruments could also cause the consistency check to fail, but if it passed the test, it would introduce a systematic bias in the comparison dataset.

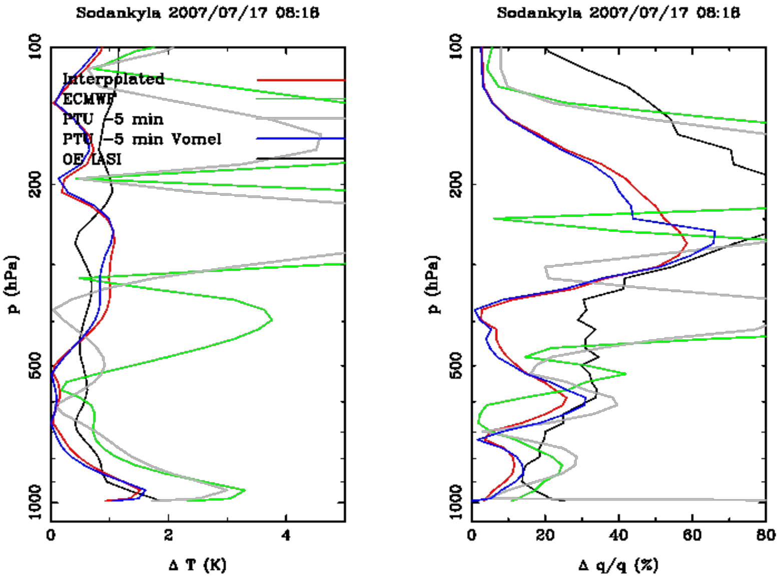

Figure 4: Retrieval error (black) and co--location and adequacy errors for the different reference profiles: "Interpolated" (red), "ECMWF'" (green), "PTU -5 min"(grey) and "PTU -5 min Vomel" (blue) profiles. From Calbet (2012).

In the GRUAN intercomparison with hyperspectral infrared exercise, seven out of eight well collocated sondes did pass the consistency check. This provides a high confidence that the pre-processing method used in this exercise is valid.

Comparison

This constitutes the final step of the Cal/Val exercise and where the conclusions are drawn. A typical conclusion is, for example, the determination of retrieval errors by plotting the statistics of the difference of the retrieved profiles versus the reference measurements. Another conclusion, which is mainly obtained from the pre-processing step and subsequent consistency check, is the determination of potential issues with the ground based measurements, such as verification of the humidity bias corrections used. Consistent differences between the observed and calculated radiances can also point to potential problems which are present in the radiative transfer model. Determination of how the difference between the reference measurement and the retrievals increase when the spatial collocation distance increases is also possible (Fig. 5). Many more conclusions could be derived from this thorough analysis of the data.

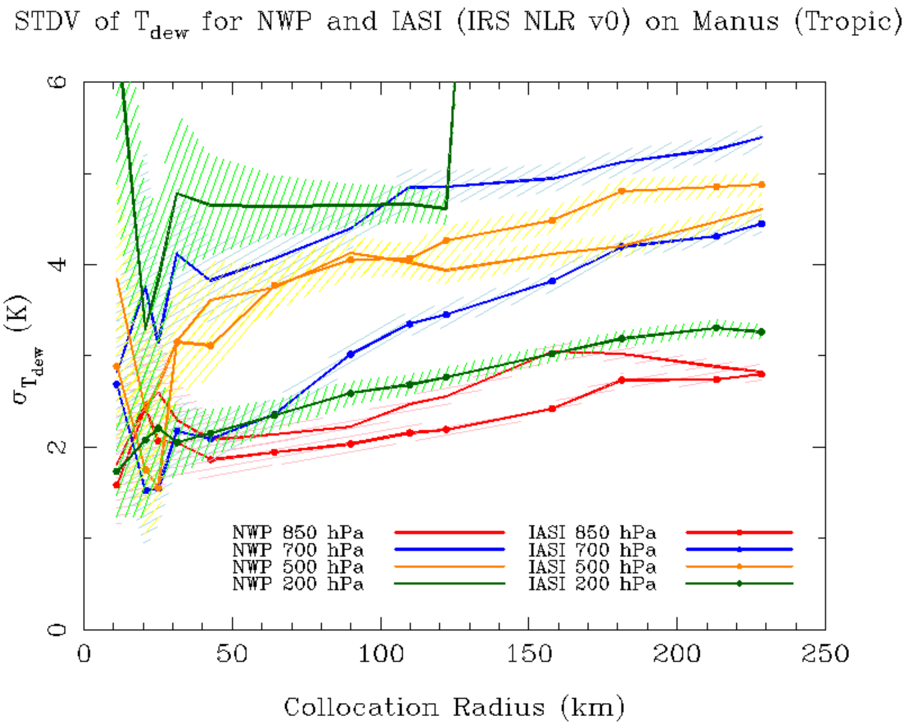

Figure 5: Standard deviation of the difference between the GRUAN sondes and IASI retrievals (solid lines) as a function of the average collocation radius as more IASI fields of view are included in the comparison. Solid/dotted lines show a similar standard deviation but for ECMWF profiles. Shading reflects the errors coming from the statistical comparison. From Calbet (2014).

Conclusions

It is clear that more effort is needed in this field. It is not always possible to have CFH humidity sensors available to obtain measurements due to their high cost and elaborate manipulation. More research should be devoted in obtaining a generalized humidity bias correction. An alternative option is to develop more cost effective and high accuracy humidity sensors, an effort undertaken by Vaisala with the RR01 model (Kivi et al. 2013). Added to that, in many circumstances, only one sonde will be launched at satellite overpass time, requiring generalizations of the Tobin et al. (2006) time interpolation that require further validation. New imaginative interpolation methods, maybe involving NWP models, may be needed.

Another approach is to use complementary measurements to detect the atmospheric conditions which are unfavourable for satellite validation. One potential example is to use complementary ground-based remote sensing instruments like GPS receivers or microwave radiometers co-located with the radiosonde launch site to detect a high variability in the atmosphere and therefore a very inhomogeneous atmosphere. This could be highly advantageous for characterising/filtering the atmospheric variability. Furthermore, high-sensitivity ceilometers can also be used to confirm cloud-free conditions over the launch site, which is essential for Cal/Val of infrared instruments. In future, water vapour LIDARs could also be used to modify the radiosonde profile to account for atmospheric changes in a similar way to that described above using multiple radiosondes. It may also be possible to use data from NWP models to characterise the spatial/temporal variability in the vicinity of each potential collocation in the same way. In a simple case this data could be used to reject cases with inhomogeneous conditions. A more sophisticated approach would be to develop an estimate of the collocation uncertainty on each potential collocation based on this time series. Potentially the concept could be further extended to eventually trigger the launch of high quality radiosondes automatically when conditions are most favourable for a satellite collocation.

A proper estimation of all the errors involved in the comparison of GRUAN sondes with satellite measurements is also critically needed. A deeper understanding in the origin of these errors would provide a better knowledge of the different factors that contribute to the total co-location error budget, making it useful in designing campaign measurements or creating new pre-processing methods. It would also provide a tool to trade-off experiment design, such as in the collocation criteria, to optimize the uncertainty. There are two ways to derive co-location errors. One of them is to determine them statistically, where the result would be an average or most probable error made when performing a co-location. Following this line of research are the papers by Pougatchev et al. (2009) and, more recently by, Fassò et al. (2013). Another line of research would be to determine the collocation uncertainty individually for each field of view, with the aid of nearby in space or time sondes or NWP models (Fig. 5). Along these lines, at the "GRUAN-GSICS-GNSSRO WIGOS Workshop on Upper-Air Observing System Integration and Application" a working group was set up to estimate with different methods the collocation uncertainty for each individual collocation case. The methods range from using several ground based measurements, NWP fields or spatial displacements of the state vector via linear methods to estimate the collocation uncertainty. Results from this work can be extremely useful in many fields, from designing Cal/Val campaigns, to deciding whether, for a given atmospheric condition, a radiosonde is worth while launching or to use it in the pre-processing or comparison step.

Future Work

Clearly much further work is needed in this area to better characterise the uncertainties in the use of radiosondes for satellite Cal/Val, and subsequently optimise the methodology for different applications. It is suggested that this work includes:

- Investigate potential methods to combine several ground based measurement methods to provide a best estimate of the atmosphere for comparison to the satellite data. Like for example, develop methods to modify radiosonde profiles using ground-based LIDAR or microwave remote sensing retrievals.
- Demonstrate the use of ground-based instruments as quality control and uncertainty characterisation of satellite-radiosonde comparison.
- Investigate potential of NWP models to characterise atmospheric variability.
- Support the effort to estimate the collocation uncertainty error, using several different methods, which is currently undertaken by the working group set up at the "GRUAN-GSICS-GNSSRO WIGOS Workshop on Upper-Air Observing System Integration and Application"

Bibliography

Buehler, S. A., M. Kuvatov, V. O. John, U. Leiterer, and H. Dier (2004), Comparison of Microwave Satellite Humidity Data and Radiosonde Profiles: A Case Study, *J. Geophys. Res.*, ***109***, D13103, doi:10.1029/2004JD004605.

Calbet, X., Kivi, R., Tjemkes, S., Montagner, F., and Stuhlmann, R. 2011: Matching radiative transfer models and radiosonde data from the EPS/Metop Sodankylä campaign to IASI measurements, Atmos. Meas. Tech., 4, 1177-1189, doi:10.5194/amt-4-1177-2011. http://www.atmos-meas-tech.net/4/1177/2011/amt-4-1177-2011.pdf.

Calbet, X. 2012. Assessment of Adequate Quality and Co–location of Sondes with Space Borne Hyperspectral Infrared Instruments to Validate Retrievals of Temperature and Water Vapor. Poster Presentation at The 18th International TOVS Study Conference (ITSC-18), 2012, Toulouse, France. http://cimss.ssec.wisc.edu/itwg/itsc/itsc18/program/files/links/10.07_Calbet_po.pdf.

Calbet, X., Crapeau, M., August T. 2013. Validation of ECMWF analyses and IASI L2 profiles with SALSTICE. Presentation at the EUMETSAT Satellite Conference 2013, Vienna, Austria. http://www.eumetsat.int/website/wcm/idc/idcplg?IdcService=GET_FILE&dDocName=PRESENTATIONS_VIENNA_20120919&RevisionSelectionMethod=LatestReleased&Rendition=Web with password: eumetsat.

Calbet, X., 2014. GRUAN and Satellite collocation. Presentation at the AIRS Science Team Meeting, March 17-19, 2014, Pasadena, California.

Dykema, J.A. (2014): Site Atmospheric State Best Estimate (SASBE). In 6th GRUAN Implementation-Coordination Meeting (ICM-6), Greenbelt, MD, 10-14 March 2014. (http://tinyurl.com/dykema-sasbe)

Hewison, T. J., 2013: An Evaluation of the Uncertainty of the GSICS SEVIRI-IASI Inter-Calibration Products", IEEE Trans. Geosci. Remote Sens., vol. 51, no. 3, Mar. 2013,doi:10.1109/TGRS.2012.2236330.

Kitchen, M. (1989), Representativeness errors for radiosonde observations. Q.J.R. Meteorol. Soc., 115: 673–700. doi: 10.1002/qj.49711548713

Fassò, A., Ignaccolo, R., Madonna, F., and Demoz, B. B. 2013 : Statistical modelling of collocation uncertainty in atmospheric thermodynamic profiles, Atmos. Meas. Tech. Discuss., 6, 7505-7533, doi:10.5194/amtd-6-7505-2013.
http://www.atmos-meas-tech-discuss.net/6/7505/2013/amtd-6-7505-2013.pdf.

Feltz, W. F., Smith, W. L., Howell, H. B., Knuteson, R. O., Woolf, H., and Revercomb, H. E. (2003). Near-continuous profiling of temperature, moisture, and atmospheric stability using the Atmospheric Emitted Radiance Interferometer (AERI). J. App. Met., 42, 584-597.

John, V. O. and S. A. Buehler (2005), Comparison of microwave satellite humidity data and radiosonde profiles: A Survey of European stations, *Atmos. Chem. Phys.*, *5*, 1843–1853, SRef-ID:1680-7324/acp/2005-5-1843, doi:10.5194/acp-5-1843-2005.

Kivi, R., Kujanpää , J., Aulamo, O., Heikkinen, P., Hassinen, S., Calbet, X., Montagner, F., and Vömel, H. 2009: Observations of water vapor profiles over Northern Finland by satellite and balloon borne instruments, in: Proceedings: 2009 EUMETSAT Meteorological Satellite Conference, 21–25 September 2009, Bath, UK, EUMETSAT P.55, 2009. https://www.eumetsat.int/website/wcm/idc/idcplg?IdcService=GET_FILE&dDocName=PDF_CONF_P55_S3_18_KIVI_P&RevisionSelectionMethod=LatestReleased&Rendition=Web.

Kivi, R., Kujanpää, J., Aulamo, O., Heikkinen P., Calbet, X., Montagner, F., Petteri, S., Jauhiainen, H., Lentonen, J., Heikki, T., Vömel, H., 2013. Vapor measurements over northern Finland. Poster at the EUMETSAT Satellite Conference 2013, Vienna, Austria.

Knuteson, R. O., Revercomb, H. E., Best, F. A., Ciganovich, N. C., Dedecker, R. G., Dirkx, T. P.,Ellington, S. C.; Feltz, W. F.; Garcia, R. K.; Howell, H. B.; Smith, W. L.; Short, J. F.; and Tobin, D. C. (2004). Atmospheric Emitted Radiance Interferometer. Part I: Instrument design. J. Atmos. Ocean. Tech., 21, 12, J. Atmos. Ocean. Tech., 21, 1763-1776..

Knuteson, R. O., Revercomb, H. E., Best, F. A., Ciganovich, N. C., Dedecker, R. G., Dirkx, T. P., Ellington, S. C.; Feltz, W. F.; Garcia, R. K.; Howell, H. B.; Smith, W. L.; Short, J. F.; and Tobin, D. C. (2004). Atmospheric Emitted Radiance Interferometer. Part II: Instrument performance. J. Atmos. Ocean. Tech., 21, 1777-1789.

Kottayil, A., S. A. Buehler, V. O. John, L. M. Miloshevich, M. Milz, and G. Holl (2012), On the importance of Vaisala RS92 radiosonde humidity corrections for a better agreement between measured and modeled satellite radiances, *J. Atmos. Oceanic Technol.*, *29*, 248–259, doi:10.1175/JTECH-D-11-00080.1.

Moradi, I., S. A. Buehler, V. O. John, A. Reale, and R. R. Ferraro (2013), Evaluating instrumental inhomogeneities in global radiosonde upper tropospheric humidity data using microwave satellite data, *IEEE Geosci. Remote Sens.*, *51*(6), 3615–3624, doi:10.1109/TGRS.2012.2220551.

Nalli, N. R., et al. (2013), Validation of satellite sounder environmental data records: Application to the Cross-track Infrared Microwave Sounder Suite, J. Geophys. Res. Atmos., 118, 13,628–13,643, doi:10.1002/2013JD020436.

Pougatchev, N., August, T., Calbet, X., Hultberg, T., Oduleye, O., Schlüssel, P., Stiller, B., Germain, K. St., and Bingham, G. 2009. IASI temperature and water vapor retrievals – error assessment and validation, Atmos. Chem. Phys., 9, 6453-6458, doi:10.5194/acp-9-6453-2009. http://www.atmos-chem-phys.net/9/6453/2009/acp-9-6453-2009.pdf.

Reale, A., B. Sun, F. H., Tilley, and M. Pettey (2012), The NOAA Products Validation System (NPROVS). J. Atmos. Ocean. Tech., 29, DOI:10.1175/JTECH-D-11-00072.1.

Tobin, D. C., Revercomb, H. E., Knuteson, R. O., Lesht, B. M., Strow, L. L., Hannon, S. E., Feltz, W. F., Moy, L. A., Fetzer, E. J., and Cress, T. S. 2006: Atmospheric radiation measurement site atmospheric state best estimates for atmospheric infrared sounder temperature and water vapor retrieval validation, J. Geophys. Res., 111, D09S14, doi:10.1029/2005JD006103.
http://onlinelibrary.wiley.com/doi/10.1029/2005JD006103/pdf.

Seidel, D. J., B. Sun, M. Pettey, and A. Reale (2011), Global radiosonde balloon drift statistics, J. Geophys. Res., 116, D07102, doi:10.1029/2010JD014891.

Sun, B., A. Reale, D. J. Seidel, and D. C. Hunt (2010), Comparing radiosonde and COSMIC atmospheric profile data to quantify differences among radiosonde types and the effects of imperfect collocation on comparison statistics, J. Geophys. Res., 115, D23104, doi:10.1029/2010JD014457.

Sun, B., A. Reale, S. Schroeder, D. J. Seidel, and B. Ballish (2013), Toward improved corrections for radiation induced biases in radiosonde temperature observations, J. Geophys. Res. Atmos., 118, doi:10.1002/jgrd.50369.