Global Criteria for Tracing the Improvements of Radiosondes over the Last Decades

Pierre Jeannet[1)], Carl A. Bower[2)], and Bertrand Calpini[3)]

[1-3)]Federal Office of Meteorology and Climatology, aerological station, PB 316, CH1530 Payerne, Switzerland. Tel. +41 26 662 62 11. Fax. +41 26 662 62 12, Pierre.Jeannet@meteoswiss.ch

[2)]National Weather Service, 1325 East-West Highway #4312, Silver Spring, Maryland

Abstract

In 2004, the CIMO expert team on upper-air systems intercomparisons decided to elaborate global criteria for tracing the improvements of radiosondes. A list of candidate performance criteria were proposed as well as the way for tracing their time evolution from the six international radiosonde comparisons organized between 1984 (UK) and 2005 (Mauritius). Geopotential height around the 10 hPa level is the criterion with the largest improvements over the two last decades. The GPS technology allowed an improvement of an order of magnitude in the quality of radiosonde geopotential heights at 30 km altitude. Large improvements have been achieved for pressure sensors, but the GPS technology constitutes a better way to improve the accuracy of pressure measurements in the stratosphere. Large improvements have also been achieved for temperature: an improvement by a factor of 3 at 30 km altitude is reported. Upper-air relative humidity measurements are most challenging and their performances in the middle troposphere at the time of the first radiosonde comparisons were very limited. The Mauritius results document a large improvement over any relative humidity sensing system in previous WMO Radiosonde intercomparisons, especially for very negative temperatures encountered in the middle and upper troposphere. The WMO international radiosonde comparisons, as well as all the other similar experiments, played a key role in the improvement of the radiosondes measurements.

1 Introduction and method

Within the framework of its reporting on the performance of upper-air observations, the CIMO expert team decided in 2004 to address the following task: “Develop performance measures to demonstrate the continuous improvement in the quality of upper-air observations”. The required action was to “elaborate global criteria for tracing the improvements, based on previous intercomparisons and recent radiosonde developments, and including remote sensing”. The report focusing on the radiosonde improvements in the last two decades is presently under review by the CIMO expert group. Appropriate criteria have been established which allow tracing the improvements in the quality of radiosonde observations.

Tracing the instrumental improvements requires an appropriate method and specific criteria. Different methods could be followed for tracing the radiosonde performance improvements, by either: (1) using the previous radiosonde comparison reports, (2) comparing radiosonde measurements with model values, (3) enquiry to the NMHSs, or (4) extracting numbers from the open literature. The first one has been selected. The first international WMO comparisons had been conducted in 1950 and 1956 (World Comparisons of Radiosondes, Payerne, Switzerland). Due to the technology used at that time, their results can only be compared in the troposphere with those of the sounding systems emerging in the early 1980s. In 1984, when the WMO Phase I international radiosonde comparison took place, the participating radiosondes were able to reach the 10 hPa (31 km) pressure level slightly more than 50 percent of the time. The planning of these experiments had been carefully defined by the new CIMO expert group, as well as the data processing and the statistical methods [Hooper, 1983)]. These guidelines have been generally respected and improved, during successive Phases. Consequently, their reports are organized in a similar way and most of the key statistical results can be found in all reports. The intercomparisons with a CIMO report took place in 1984 in the UK (Phase I), 1985 in the USA (Phase II), 1989 in the former USSR (Phase III), 1993 in Japan (Phase IV), 2001 in Brazil (Phase V), and 2005 at Mauritius (Phase VI). Besides them, a few other special international campaigns also took place in this time period, such as in the UK (PREFERS, 1992), and in the USA (Wallops Island, 1995, humidity sensors).

The statistical parameters (systematic biases, standard deviations, etc) based on differences between the measurements obtained with different types of radiosondes for simultaneous measurements represent a valuable tool for comparison over the last two decades. Each of the Phase I to VI intercomparisons used "link radiosondes" in order to define one reference value (working reference) for comparing it with the measurements of all participating radiosondes. Although the link sonde approach has shortcomings, it allows a straightforward use of the IOM reports.

Candidate quality performance criteria for temperature and geopotential altitude are presented in Table 1. They rely on comparison of simultaneous time-paired measurements. The selection was made with the objective of a small number of criteria in order to trace the improvement of radiosondes over the years with a straightforward data analysis. Criteria related to mean differences between sondes (bias) correspond to systematic measurement errors. They are more sensitive to radiosonde measurement problems than criteria based on the standard deviation of the differences between radiosondes. They allow the origins of radiosonde deficiencies to be determined. Standard deviation of the differences between radiosondes complements the information provided by the mean difference between radiosondes. If the standard deviation is smaller than the bias, it allows assigning measurement errors to a systematic problem in the radiosonde design and/or in the data processing. The standard deviation may also help identify error sources in the radiosonde reproducibility.

Table 1. Temperature and geopotential candidate criteria for tracing the improvements of radiosondes.

<table>
<tr><th>Criteria</th><th>Remarks</th></tr>
<tr><td>Temperature difference around*) 10 or 30hPa, @ night/day time</td><td rowspan="2">The 10 hPa level is the highest standard level in the TEMP messages. Reaching a high quality standard around this level is a demanding task. Temperature errors are different during night and daytime (noon). A higher data sample is found around 30 hPa than around 10 hPa, particularly in the first Phases.</td></tr>
<tr><td>Standard deviation of the temperature differences around*) 10 or 30 hPa, @ night/day time</td></tr>
<tr><td>Geopotential difference around*) 10 or 30 hPa, @ night/day time</td><td rowspan="2">Geopotential measurements from a radiosonde accumulate errors from other parameters (temperature, pressure, etc.) between surface and this level. Recent advances in GPS positioning have brought major upgrade on this criteria.</td></tr>
<tr><td>Standard deviation of the geopotential differences around*) 10 or 30 hPa, @ night/day time</td></tr>
<tr><td>Geopotential difference around*) 100 hPa, @ night/day time</td><td rowspan="2">The 100 hPa level is the primary level used in the quality control of upper air data based on comparison with numerical model outputs.</td></tr>
<tr><td>Standard deviation of the geopotential differences around*) 100 hPa, @ night/day time</td></tr>
</table>

*) These approximate pressure levels ("around") require an explanation. The first WMO radiosonde comparisons defined 15 pressure categories in the comparison of simultaneous measurements. The 10 hPa category considered all measurements between 8.4 and 11.9 hPa, as defined by the link sondes. The 30 hPa category was more exactly centred at 32 hPa (24.5 – 41.5). The 100 hPa category range was 84 – 119 hPa. This ensured that the statistics were relying on a sufficient number of time-paired measurements. In the more recent radiosonde comparisons, 2 km wide altitude categories were introduced instead of the previous ones. The altitude category that included the wanted pressure level was then used.

We illustrate the method on the basis of the example of the first criteria in Table 1. The systematic temperature differences around 10 hPa in the night-time were reported in an Excel sheet. The bias values were extracted from Figures or Tables of the six IOM reports, without any additional processing. Each value (bias of one sonde type during one radiosonde comparison) was given with the Figure or Table number as well as report number from where it was taken. Almost 30 radiosonde types and versions have been intercompared at least once over all six comparisons. The values of this Excel sheet were then represented with symbols in a corresponding figure. However, the radiosonde results appear anonymously in the figure, as the aim is not to find out the "best radiosonde", but to demonstrate that a general and continuous improvement in the quality of upper-air observations occurred over the last 20 years. The horizontal axis is a time axis covering the last 25 years. On the vertical axis, the span of the bias values is more important than their exact positions in relation to the zero point, as the reference is a relative one. In every comparison an outlier point may strongly increase the span. Basic statistical parameters complete the individual results, e. g. the envelope of the extreme values (maximum, minimum), the maximum span (difference between maximum and minimum), as well as the standard deviations of the biases. Other statistical parameters could be added, such as the average biases and the mean absolute biases. Nevertheless, one should be aware that all statistical parameters in this study have a somewhat limited statistical significance, as less than 10 radiosonde types (including different post-processings for the same sonde) were engaged in each comparison.

2 Results related to geopotential height

Up to a few years ago, radiosonde geopotential heights were mostly calculated with the hydrostatic equation. This method needs the pressure, temperature and humidity profiles and combines their errors into the calculated geopotential heights. Nowadays, the newer radiosondes use the GPS technology, directly measure geometric height and convert it to geopotential height. Geopotential altitude measurements are highly demanding. A 1 hPa error at 10 hPa corresponds to a 600 m geopotential error. Although reporting meteorological parameters at pressure levels (considered as true values) alleviates the errors of radiosonde measurements, this report is devoted to the accuracy of radiosonde measurements and only compares truly simultaneous measurements.

Figure 1 illustrates these improvements by pointing out in an anonymous manner all radiosonde biases of the geopotential altitude around the 10 hPa level, using the method described in chapter 1. In Phase II, a high precision radar was used as altitude reference. It demonstrated the real altitude errors the radiosonde was making. In Phase V, for the first time, GPS was introduced for height measurements on two sondes and the comparisons included in Figure 1 use these GPS results as reference. In Phase VI, there was still sondes measuring height on the basis of pressure sensors, but only the three full GPS radiosondes are documented in Figure 1. The envelope on Figure 1 started with a span of 1000 – 2000 m in the first three comparisons. In Phase IV, better sensors and better calibration curves reduced this span to approximately 500 m. In the Brazil campaign (Phase V), 8 years later than Phase IV campaign, some additional improvement were demonstrated, but fewer balloons reached 10 hPa during this campaign than during the previous ones and the next one and this hampered the comparisons above 30 km. The move to the GPS technology brought a new standard in geopotential measurement accuracy. In 2001, this technology was in an introductory phase. In 2005, it proved having reached its full potential. The GPS radiosondes are nowadays able to measure geopotential altitude at 31 km with an average agreement of about 20 meters (cf. Mauritius report). They reach the same absolute accuracy over their entire altitude range.

Figure 2 complements the results of Figure 1 with the corresponding estimated random errors of the geopotential altitudes around 10 hPa. It demonstrates that the reproducibility of the geopotential measurements improved as much as their accuracy. In the 1980s, radiosonde systems had random errors up to 1000 meters in the altitude range above 30 km. Both the poor accuracy and the poor reproducibility of the radiosondes were responsible for their very limited performances in the middle stratosphere. In the 1990s, a noticeable improvement was reached from improvements in the radiosonde technology and the altitude errors have been reduced to 100 – 200 meters near 30 km. In the last years, the GPS technology allowed a new improvement of an order of magnitude in the accuracy and reproducibility of radiosonde geopotential heights.

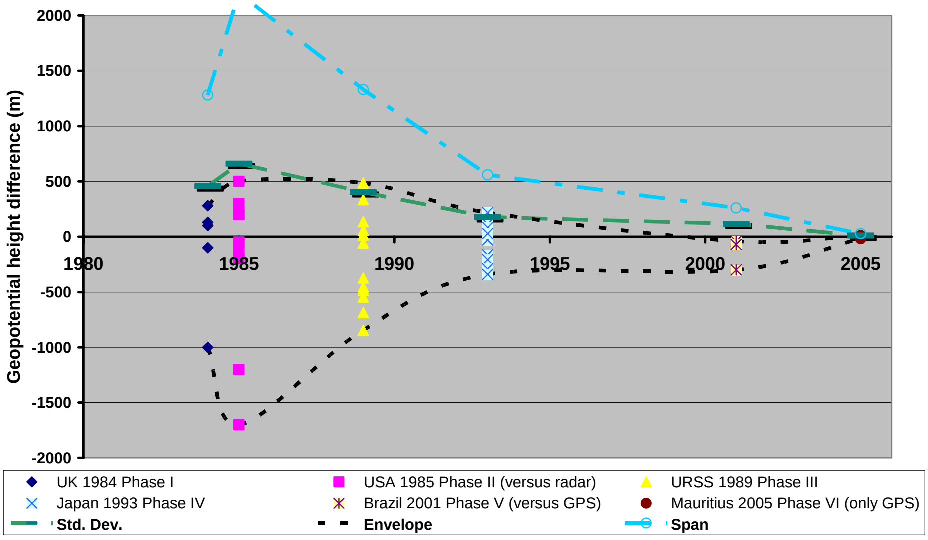

Figure 1. Bias of the geopotential altitude around 10 hPa for the six WMO Radiosonde Comparisons (simultaneous measurements). The two dotted black lines represent the envelope of all individual results, which is converted into a span with the dash-dotted blue line. The horizontal green bars on the dashed green curve correspond to one standard deviation of the biases for each comparison.

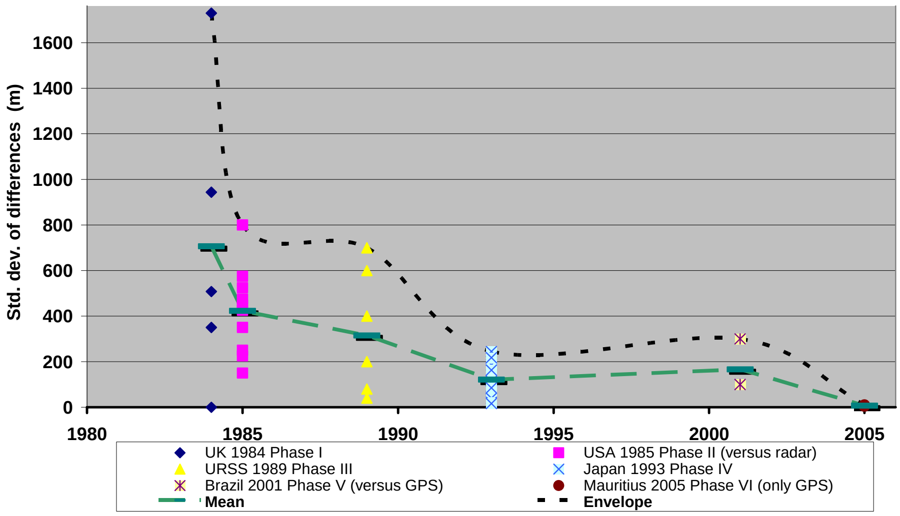

Figure 2. Estimated random errors of the geopotential height measurements around 10 hPa for the six WMO Radiosonde Comparisons (simultaneous measurements). The dotted black line represents the envelope of all individual results. The horizontal green bars on the dashed green curve correspond to the mean random errors.

3 Results related to temperature

In the 1980s temperature sensors were still of varying types: bimetal element with mechanical link, coiled tungsten wire (resistive element), thermistor rod, thermistor bead, thermocapacitive bead, etc. Later on the first two types were not used in the WMO comparisons and new sensors appeared. Modern temperature sensors are often much smaller than those in use in earlier years and they are placed outside of the radiosonde box. Due to the different night- and day-time behaviour of the radiosonde temperature sensors from the sunlight and infrared radiation, the radiosonde intercomparisons were systematically performed under both conditions (night and day). As far as possible, they also captured the daily cycle with soundings at different solar elevations.

Figure 3 illustrates these improvements by pointing out in an anonymous manner all individual radiosonde biases around the 10 hPa level. Both midnight and daytime results are presented on the same graph, in order to emphasize the long-term evolution of the envelope of the negative and positive bias. Within the last 20 years, the technology evolution of the temperature sensors as well as more sophisticated data processing (improved sensor coatings, radiation correction algorithms, improved sensor booms, calibration fits, removal of statistical bias, etc.) has allowed large improvements. Nevertheless, they are not as large as in the case of the geopotential measurements. An improvement by a factor of roughly 3 emerges from Figure 3. The direct sun radiation generally has a larger influence than the IR radiation balance at night, although it only heats – and never cools – the sensor. The Mauritius results in Figure 3 depict a negative outlier due to a thermometer with white coating producing a radiative cooling at night time.

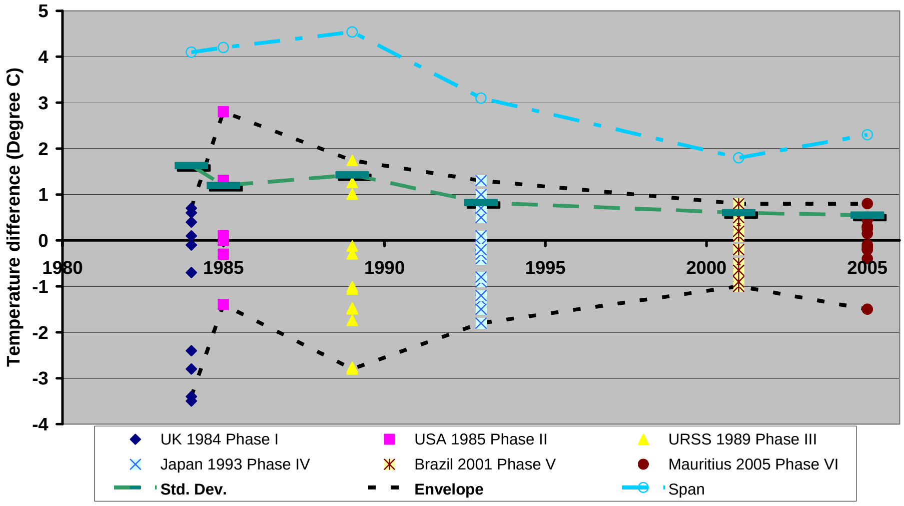

Figure 3. Night and day time temperature bias around 10 hPa for the six WMO Radiosonde Comparisons (simultaneous measurements). The two dotted black lines represent the envelope of all individual results, which is converted into a span with the dash-dotted blue line. The horizontal green bars on the green dashed curve correspond to one standard deviation of the biases of each comparison.

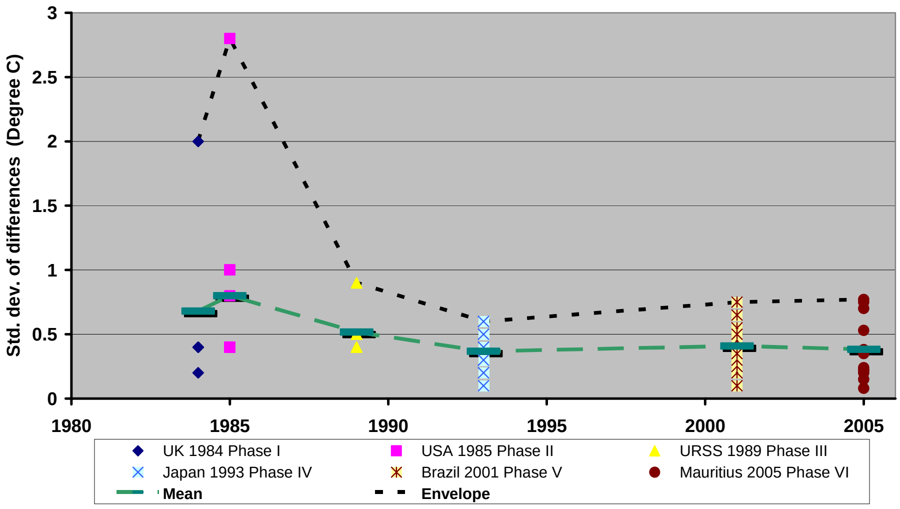

Figure 4. Estimated random errors of the temperature measurements around 10 hPa for the six WMO Radiosonde Comparisons (simultaneous measurements). The dotted black line represents the envelope of all individual results. The horizontal green bars on the dashed green curve correspond to the mean random errors.

Figure 4 completes the results of Figure 3 with the corresponding estimated random errors of the temperature around 10 hPa (31 km). The improvement between the 1980s and the 1990s is large. The Brazil and Mauritius results are similar. However, this does not reflect the general results, as the radiosondes agreed more closely together in the low stratosphere and in the troposphere in the Mauritius campaign than in Brazil. At Mauritius, two of the radiosondes had daytime random errors less than 0.2 K at heights up to 30 km, whereas the other ones did reach this performance only up to 16 km. A redesign of their temperature sensor mount would minimize the fluctuations from air that has passes over surrounding sensor support structures.

4 Conclusions of the report

Geopotential height around the 10 hPa level is the first selected criterion with very large improvements over the two last decades. The GPS technology allowed an improvement of an order of magnitude in the quality of radiosonde geopotential heights at 30 km altitude. At Mauritius, all the GPS height measurements agreed on average to within ±20 m from the surface to 34 km.

Large improvements have been achieved for **temperature**: an improvement by a factor of 3 at 30 km altitude is reported. The best high quality radiosondes performed very well in the last experiment in 2005.

Large improvements have also been achieved for **pressure** sensors, but the GPS technology constitutes a better way to improve the accuracy of pressure measurements in the stratosphere.

Upper-air relative **humidity** measurements are most challenging. Relative humidity has been only partially treated in this report, due to their limited performance in the middle troposphere at the time of the first radiosonde comparisons. New systematic calculations on the basis of the original data sets would also be necessary in order to apply the same method as for temperature. The Mauritius results document a large improvement over any relative humidity sensing system in previous WMO Radiosonde intercomparisons, especially for very negative temperatures encountered in the middle and upper troposphere.

Wind has not been studied in this report, as it would require more specific criteria. However, it is well recognized that large improvements have been achieved during the last 20 years. All the GPS radiosondes in the Mauritius intercomparison can measure winds accurately enough to any height to satisfy climatological requirements, given that percentage of missing data is low.

Some final remarks complement these results:

- Although large improvements in the quality of the radiosondes have been achieved in the last two decades, it is not easy to quantify the overall improvements in a synthetic way. The method used in our report focuses on a limited number of criteria and provides for them a clear demonstration of large improvements in the performance of radiosondes in the last 20 years.
- The WMO international radiosonde comparisons, as well as all the other similar experiments, played a key role in the improvement of these measurements. Their results and the technology improvements have been extensively used by the manufacturers in order to successfully improve their products.

5 References

Hooper, A.H., J.F. Ponting, 1983: Analysis Schemes For Treating Data from Radiosonde Intercomparisons International Organizing Committee for Radiosonde Intercomparison 1984, Report of the First Session, Annex I.

Hooper, A.H., 1986: WMO international radiosonde comparison, Phase I, Beaufort Park, U.K., 1984, WMO/TD-No. 174, Instruments and Observing Methods Report No. 28, World Meteorological Organization, Geneva, 118 pp.

Schmidlin, F.J., 1988: WMO international radiosonde comparison, Phase II, Wallops Island, USA, 1985, WMO/TD-No. 312, Instruments and Observing Methods Report No. 29, World Meteorological Organization, Geneva, 113 pp.

Nash, J., F.J. Schmidlin, 1987: WMO international radiosonde comparison, (U.K. 1984, U.S.A. 1985) Final report, WMO/TD-No. 195, Instruments and Observing Methods Report No. 30, World Meteorological Organization, Geneva, 103 pp.

Ivanov, A., A. Kats, S. Kumosenko, J. Nash, and N. Zaitseva, 1991: WMO international radiosonde intercomparison phase III (Dzhambul, USSR, 1989) final report. WMO/TD-No. 451, Instruments and Observing Methods Report No. 40, World Meteorological Organization, Geneva, 135 pp.

Yagi, S., A. Mita, and N. Inoue, 1996: WMO international radiosonde intercomparison phase IV (Tsukuba, Japan, 1993) final report. WMO/TD-No. 742, Instruments and Observing Methods Report No. 59, World Meteorological Organization, Geneva, 130 pp.

da Silveira, R., G. Fisch, L. Machado, A. Dall'Antonia, L. Sapucci, D. Fernandes, and R. Marques, 2006: WMO international radiosonde intercomparison phase (Alcantara, Brazil, 2001). Executive summary of the GPS radiosonde comparison, Brazil, see WMO web page, under CIMO.

Nash, J., R. Smout, T. Oakley, B. Pathack, S. Kurnosenko, 2006: WMO Intercomparison of High Quality Radiosonde Systems (Vacaos, Mauritius, 2005) final report. WMO/TD-No. xxx, Instruments and Observing Methods Report No. yy, World Meteorological Organization, Geneva, 118 pp.