Preliminary Results of WMO Intercomparison of high quality radiosonde systems, Mauritius, February 2005

B. Pathack
Mauritius Meteorological Services, Vacoas

J. Nash and R. Smout
Met Office, Exeter, UK
and
S. Kurnosenko
Scientific software consultant, US

This paper provides a brief survey of the initial results from the WMO Intercomparison of High Quality Radiosonde Systems, Mauritius. All radiosondes that took part in the test proved to be of good quality, although problems were identified with some systems which will need to be addressed in the long term. The most suitable contenders for climate temperature monitoring work were the Vaisala and Sippican radiosondes. Both specialised reference systems (Meteolabor Snow-White and Sippican 3-Thermistor) needed further development to be fully reliable as working references.

1. Introduction

The WMO Intercomparison of High Quality Radiosonde Systems consisted of 62 successful radiosonde comparison flights, performed between 7 and 25 February 2005 at the headquarters of the Mauritius Meteorological Services, Vacoas.

This test was organised by the Expert Team on Upper- Air Systems Intercomparisons, chaired by Dr. J. Nash. Mauritius Meteorological Services volunteered to host this test at CIMO-XIII. February 2005 was chosen for the test by the Expert Team to allow the radiosonde relative humidity sensors to be tested in both wet and dry conditions.

B. Pathack performed a variety of tasks as Project Manager. Organising import and export of equipment proved a major effort, but all equipment was delivered on time. Provision of facilities for the test included the installation of a hydrogen generator to facilitate filling 2000g balloons, stabilised power supply for the ground system computers and Internet connections for the participants. The number of international participants present in Mauritius at any time during the test was about 15. The typical number of people from Meteorological Services Mauritius involved in the test during one week was more than 20. 4 teams of 3 persons were trained to prepare balloons, provide surface observations and manage the launch of the balloons. Radiosonde support rigs were assembled in advance from green bamboo canes obtained locally. Training in balloon handling and comparison launch procedures were provided by J. Nash and R. Smout.

Mauritius Meteorological Services provided technical support to participants throughout the test. Repair work was performed on the power supplies of the cloud radar brought from the UK, without which measurements would not have been obtained.

2. **Radiosondes tested**

Six main radiosonde types were tested [method of height determination indicated]:

Vaisala RS92	(Finland)	[pressure sensor and GPS height]
Graw DFM-97	(Germany)	[pressure sensor and GPS height]
MODEM M2K2	(France)	[GPS height only]
Sippican MKIIA	(USA)	[GPS height only]
Meisei RS-01G	(Japan)	[GPS height only]
Meteolabor, SRSD-C34	(Switzerland)	[Hypsometer pressure sensor]

InterMet Systems withdrew from the test in autumn 2004.

Vaisala, Modem and Meisei radiosondes and Vaisala, Graw and Sippican radiosondes were flown together as two groups with either Meteolabor or three thermistor radiosondes included as a working reference. This arrangement had been agreed at the International Organizing Committee by the HMEI representatives, since all manufacturers wished to have the opportunity to compare with Vaisala. Fig.1 shows preparation for launching the Vaisala-Sippican-Graw radiosonde group with a 3 thermistor radiosonde.

Fig.1 Graw and Sippican 3 thermistor radiosondes ready for a night time comparison of the Vaisala-Sippican-Graw radiosonde group on the bamboo cross support rig.

All radiosondes tested were operating in the band 400.5 to 405.5 MHz. It would have been possible to fly all the radiosondes supported by one balloon if the frequency stability and bandwidth of the Sippican transmitters had been similar to the other radiosondes. Only one set of measurements was lost from poor radiofrequency reception /interference, when the first SRS was flown at too low a frequency.

Vaisala radiosondes were flown on all test flights, with 59 out of 62 successful. Thus, it is convenient to use the Vaisala measurements to link the performance of the radiosondes in the two groups. Launch times were separated by about 5 hours to allow enough time to generate hydrogen for the 2000g balloons, but this separation was shortened at night in the second and third weeks. Thus, the two daytime flights were launched at 09.00 and 14.00 local time, so that solar elevation was similar in the stratosphere for both groups. Night time launches were at 19.00 and between 22.00 to 23.30 local time.

24 Sippican MKII, 3 thermistor radiosondes were flown, [5 at night, 19 in the day] to provide a "working reference" for temperature. The Snow- white chilled mirror hygrometer was successfully deployed as a "working reference" for dew point/ relative humidity measurements on 34 flights.

During the second half of the test MODEM flights were operated by staff from Mauritius, see Fig.2, and the Graw system was operated by staff from the UK at the request of the respective manufacturers.

Fig.2 Staff from the teams of Mauritius Meteorological Services working with the MODEM system during the second half of the test.

50 flights reached higher than 30 km and sufficient flights ascended to heights above 34 km to provide useful comparisons up to this level. The balloon performance was judged as good given the rainy conditions and the presence of thick upper cloud at night for part of the test. Local staff only received a short period of training before starting the test, but coped well with unpredictable launch conditions with low level winds varying significantly between launches.

3. Data processing

The processing software used for this test was provided by S. Kurnosenko. This was an updated version of the RSKOMP software used to analyse results from Phases III and IV of earlier WMO Radiosonde Comparisons.

Sergey Kurnosenko managed the data input from the files provided by the manufacturers. The workload associated with data entry was increased by the large number of last moment modifications made to proposed file formats by most manufacturers in the test. The comparison data base consists of samples extracted at 1s intervals from the files provided by the manufacturers, after modification of the extraction software on site in Mauritius.

The attempt to use GPS timing as a method of synchronising samples did not work because of a lack of consistency in the use of GPS time between the systems. In practice, data samples were synchronised by matching temperature and relative humidity profiles near the ground using the WVIEW software. The adjustment procedure works well with temperature and humidity data sampled at 1 s intervals. The timing adjustment procedure may not work so well for pressure, especially near the ground where sensed values may have been adjusted by software to a different value to be consistent with a different launch time.

Input data were checked by the WMO supervisory team as soon as possible following the flight. Problems with systems were discussed with the specific teams, e.g. the filtering of the Japanese GPS measurements and a solution agreed. The aim was to ensure that data represented correct functioning of the systems deployed in Mauritius. For some of the systems, this entailed ensuring that algorithms for converting GPS geometric height to geopotential height used the correct value of g for Mauritius.

In some cases, launch procedures were modified to try and prevent damage to the more fragile radiosondes and to ensure that other systems did not lose GPS lock during the launch procedure.

Test procedures and early results were reviewed towards the end of the first week by all the participants. The team leaders agreed that test procedures were satisfactory.

Some data problems became more obvious towards the end of the trial and this required some rework of the observations after the final flight.

- Vaisala reprocessed daytime temperature measurements using a different editing filter.
- Meisei recomputed temperatures because incorrect corrections had been applied to night time measurements during the test.
- Meteolabor reprocessed geopotential heights because of errors in the height computation software.
- MODEM reprocessed geopotential height computations since an incorrect value of local g had been used for the geometric height to geopotential conversion.

Statistical processing was based on the WSTAT program supplied by S. Kurnosenko. The data were edited by the Chairman of the IOC, before the statistics were processed. Editing was required mostly by the two specialised sensing systems, where elimination of the various occasional Snow white failure modes [excessive instability in dewpoint measurement at upper levels, contamination in the Snow white duct leading to dewpoints that were higher than air temperature in upper layers in daytime, loss of the water film on the chilled mirror in very dry layers], and thermistor calibration errors/ inter-channel radiofrequency offset problems for the 3 Thermistor radiosondes.

When a temperature sensor becomes wet in passing through cloud, the sensor is cooled on emerging into a drier layer above the cloud, as the water evaporates. The Vaisala sensor was least sensitive to this problem. The other manufacturers ought to consider using a hydrophobic coating on the temperature sensor to minimise the significant errors that follow sensor wetting. The measurements in the layers where this wetting error happened were hidden and not used in the following statistics.

4. Simultaneous temperature comparisons

Figures 3(a) and (b) show the results of temperature comparisons at night. Meisei, Sippican and Vaisala measurements agreed to within ± 0.3 K from the surface to 31 km. At the lowest temperatures [-80 deg C] in the upper troposphere, Graw and SRS temperatures had calibration discrepancies of about +0.5 K. In the case of Graw the discrepancies were much smaller than the night time errors in the previous WMO GPS Radiosonde test in Brazil, where errors had been larger than 1K. Only, the MODEM radiosonde had a temperature sensor coated with white paint. This sensor was in error by more than 2 K at 30 km, with at least 1K the result of cooling by infrared radiation.

At night the standard deviations of the differences between each radiosonde and Vaisala were consistent with a reproducibility of temperatures from sensors with aluminized coatings of better than 0.2 K from the surface up to 30 km.

Figures 4(a) and (b) show the results for day time temperature comparisons. The absolute values of the three thermistor measurements may be offset by up to ± 0.2 K from truth. Significant numbers of errors in individual thermistor measurements had to be offset by comparison with the results from the other radiosondes. However, the 3 thermistor measurements provide an accurate representation of the variation in the vertical of the correct temperature in the stratosphere.

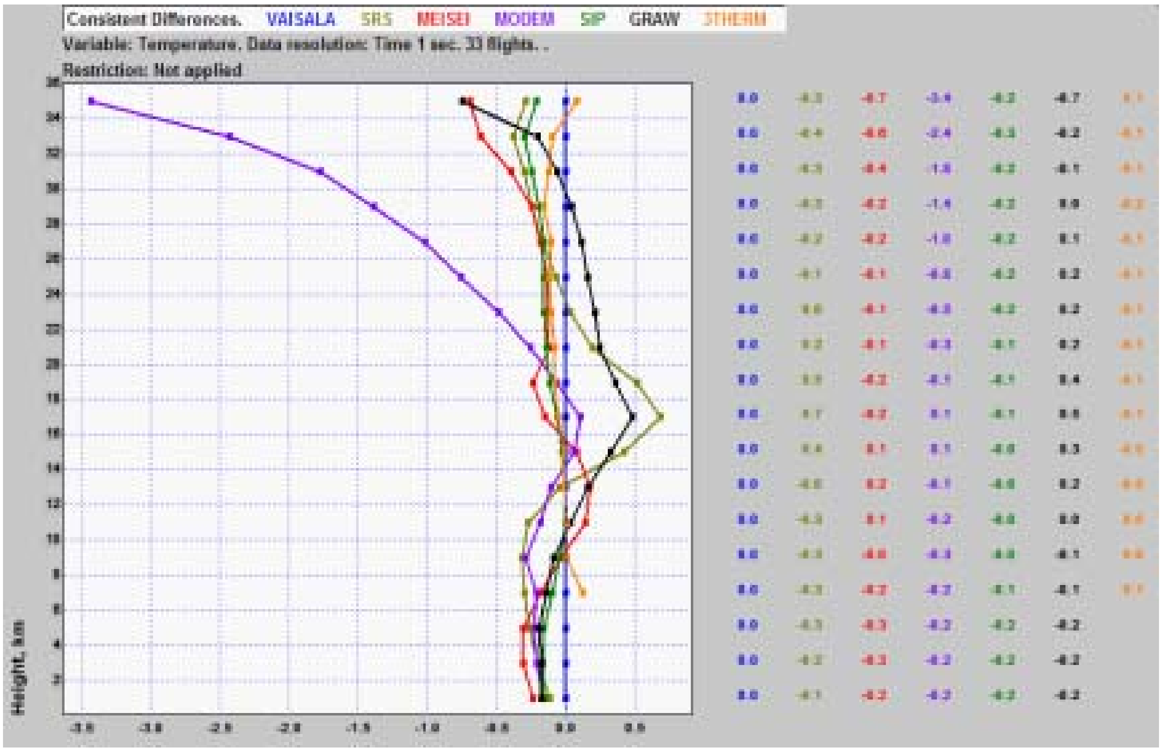


Fig.3 (a) Systematic bias between simultaneous temperatures [K] at night using Vaisala as a working reference

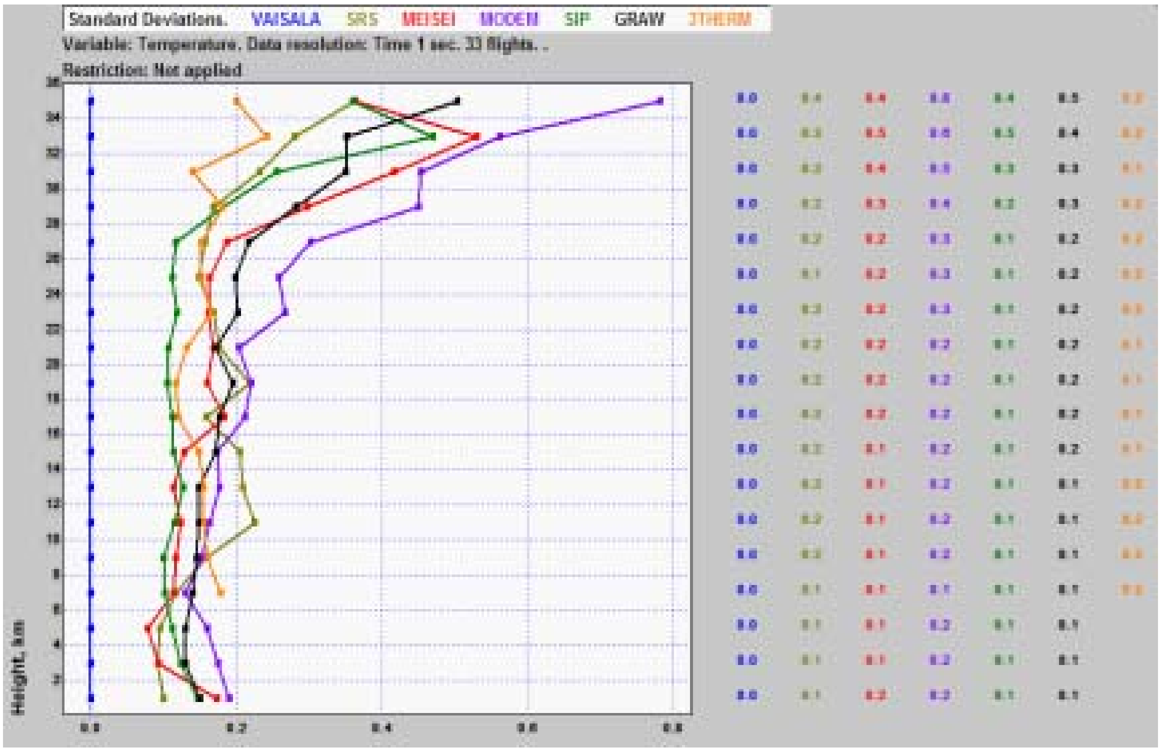


Fig. 3(b) Standard deviations between simultaneous temperatures [K] at night using Vaisala as a working reference.

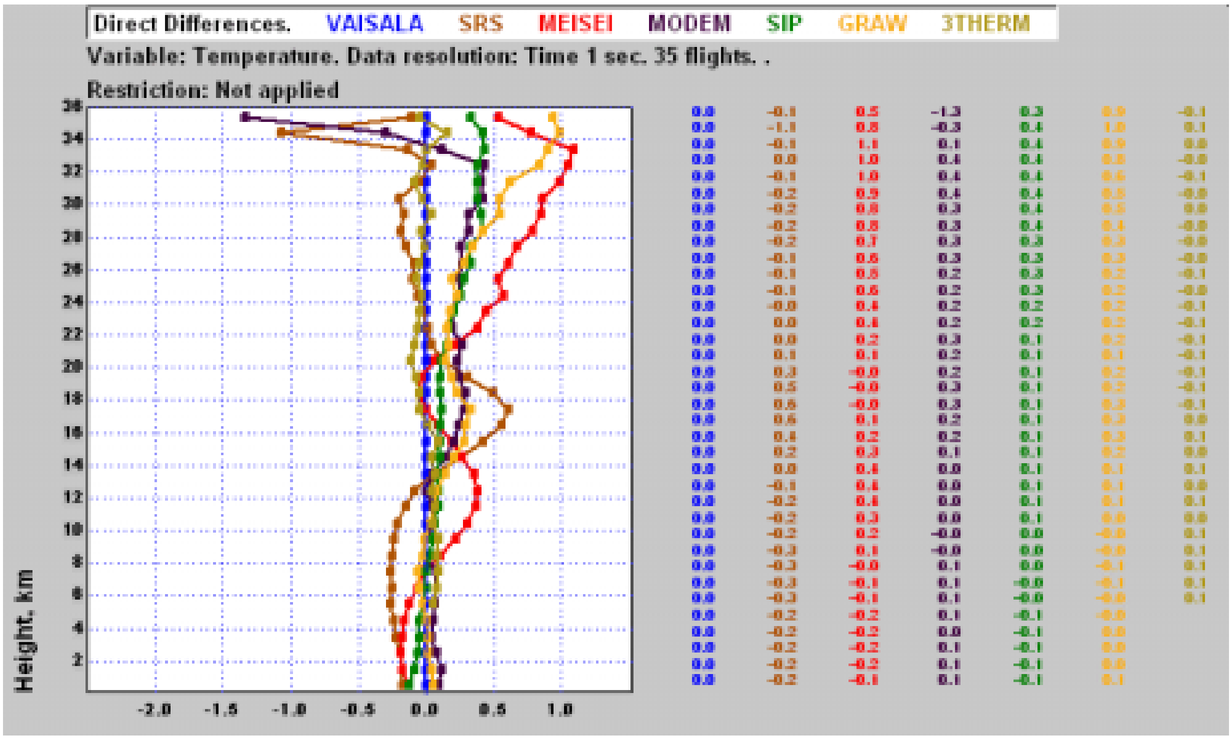


Fig.4 (a) Systematic bias between simultaneous daytime temperatures [K] using Vaisala as a working reference

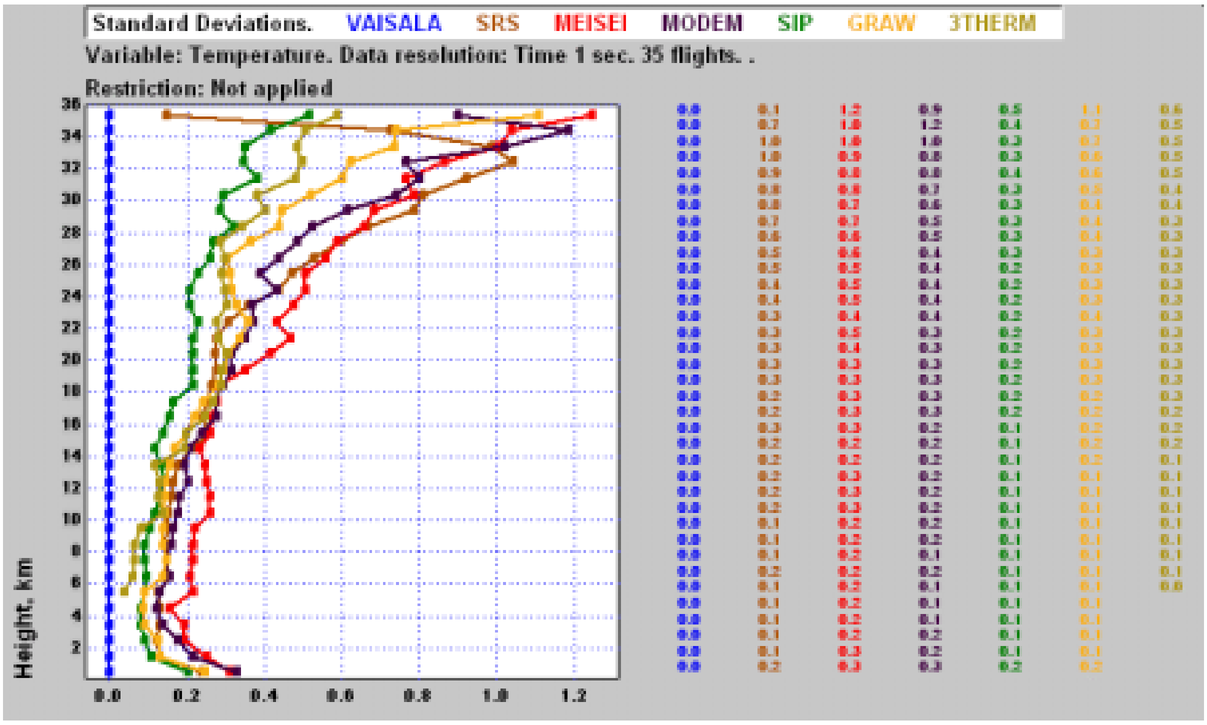


Fig. 4(b) Standard deviations between simultaneous daytime temperatures [K] using Vaisala as a working reference.

Vaisala make the smallest daytime temperature correction (about 0.6 K at 10 hPa). SRS and Sippican make corrections of just over 1 K at about 30 km. From Fig 4 (a), Vaisala and SRS corrections produce results very close to the three thermistors at upper levels. The Sippican measurements diverge from the three thermistors at 30 km so the Sippican corrections should probably be larger by about 0.3 K at 30 km. Modem temperature corrections are about 2 K at upper levels. Meisei daytime temperature corrections, about 2.5 K at 30 km, were larger than most of the other radiosondes. With the upper cloud conditions experienced in Mauritius, Meisei temperature corrections needed to be larger by at least 0.7 K at 10 hPa.

Temperature errors in daytime measurements fluctuate in the short term as the radiosondes rotate in flight, with the period of the predominant rotation between 10 and 15 s. These fluctuations increase with height and affect all the radiosondes to some extent, including Vaisala. Raw daytime Vaisala measurements have significant error fluctuations at upper levels. Air passing over the Vaisala sensor support is warmed and if this air then passes over the temperature sensor, positive temperature error pulses result. In Mauritius, the magnitude of the temperature pulses was about 1 K at 30 km if the radiosonde was rotating smoothly under the support rig. These pulses are larger than occur in individual Vaisala flights where the radiosonde motion is more random. In the reported Vaisala data, the temperature pulses are largely filtered out by Vaisala processing software. The new filter used in processing the final Vaisala data in Mauritius is not yet in operational use. The original Vaisala data in Mauritius used the existing operational Vaisala filter. At heights above 28 km these original Vaisala temperatures showed larger standard deviations in the differences with Sippican and three thermistor than are now found in Fig. 4(b). The standard deviations in Fig. 4(b) are consistent with the reproducibility of Sippican and Vaisala daytime temperature measurements being better than 0.2 K at heights up to 28 km. The random errors in the other radiosonde daytime temperature measurements became larger than 0.2 K above 18 km. At 32 km, the random errors in daytime temperatures had increased to between 0.3 and 1 K depending on the radiosonde design.

Overall, the two most suitable radiosonde temperature measurements for climate monitoring both day and night were Vaisala and Sippican. Three thermistor radiosonde measurements can also give very high quality measurements if the system is implemented carefully, but the Sippican system used in Mauritius needs further development to be reliable as an absolute reference.

A combination of Vaisala GPS with a suitable operational version of the Sippican GPS radiosonde would be recommended for best measurement quality for high performance climate/satellite monitoring.

5. <u>Simultaneous geopotential height comparisons</u>

The simultaneous height comparisons from this test demonstrate that GPS height measurements give geopotential heights that are more accurate than the best pressure sensors at all heights above 16 km and are of similar accuracy to pressure sensor measurements at heights below 16 km. The systematic bias of all the geopotential heights relative to the Vaisala GPS height measurements are shown in Fig. 5 (a).

All the GPS height measurements agreed on average to within ± 20 m from the surface to 34 km. At 30 km pressure sensors were in error by values between -70m (Vaisala) up to +120m (SRS). Both Graw GPS and Sippican geopotential had standard deviations relative to Vaisala of less than 10m. Thus, GPS heights are suitable to replace geopotential from pressure sensors at all heights, i.e. a pressure sensor is no longer a necessity for a best quality radiosonde.

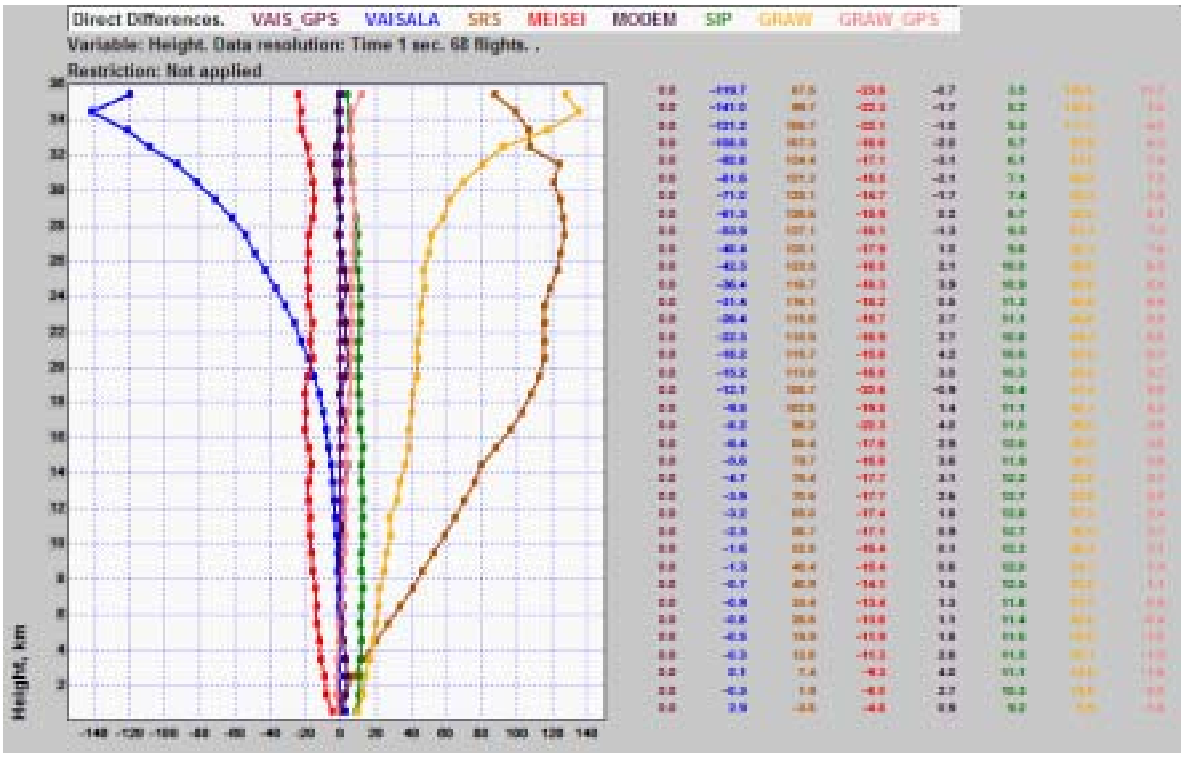


Fig. 5 (a) Systematic bias between simultaneous geopotential height measurements [gpm], using Vaisala GPS measurements as a working reference. Vaisala, SRS and Graw are heights derived from high quality pressure sensors

The reproducibility of the GPS geopotential heights at 32 km is an order of magnitude better than the reproducibility of the heights from the best pressure sensors, see Fig. 5(b) . Thus, temperature errors caused by height errors in radiosonde output will become negligible with the new GPS height measurements, even at pressures lower than 5 hPa.

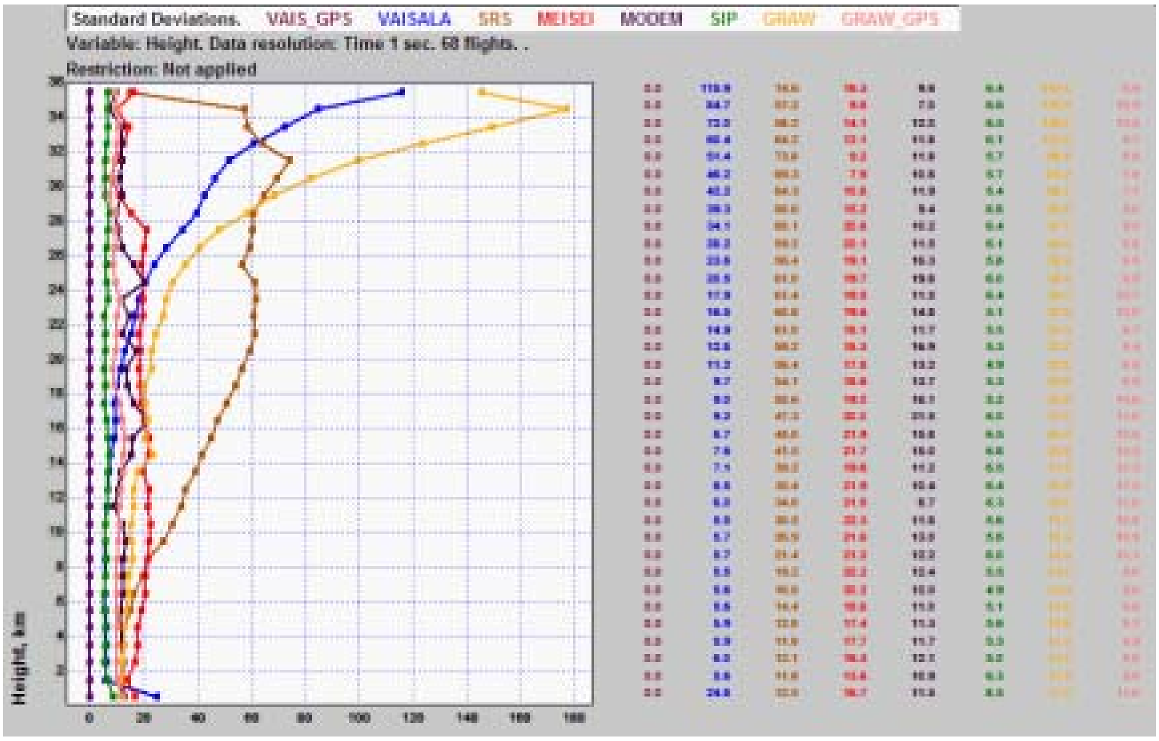


Fig. 5 (b) Standard deviations between simultaneous geopotential height measurements [gpm] using Vaisala GPS measurements as a working reference. Vaisala, SRS and Graw are heights derived from high quality pressure sensors

6. Simultaneous pressure comparisons

Fig. 6 shows the results of the simultaneous pressure comparisons from the data base. Two Modem flights where water/ice apparently shunted the temperature sensor for part of the flight giving very large negative temperature anomalies were excluded Four out of 34 SRS pressures were also judged atypical and excluded.

The spread of systematic differences in pressure close to the ground may have partly been the results of the time adjustment procedure used to synchronise temperature and relative humidity and winds, since all systems were using a similar surface pressure, but the launch times used were not always coincident

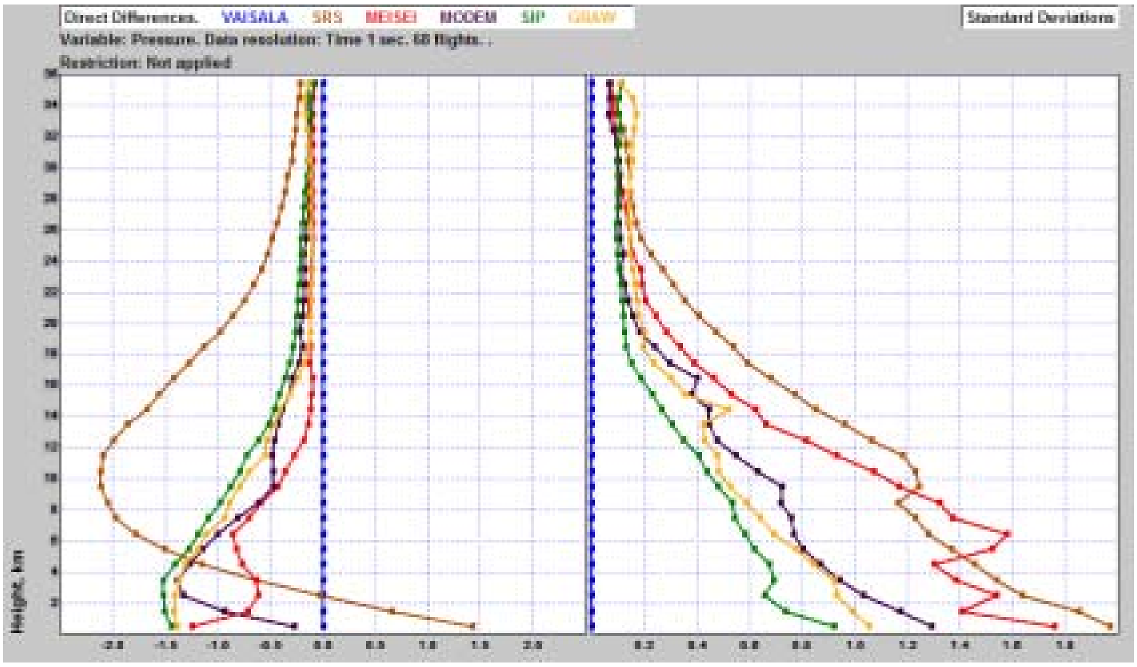


Fig.6 Systematic bias and standard deviations of simultaneous comparisons between pressure measurements [hPa].

7. Simultaneous relative humidity comparisons

The systematic differences between the relative humidity sensors is presented as a function of height for 5 relative humidity bands, with daytime and night time results presented separately. The bands were 75 to 95 per cent in Figs. 7(a) and (b), 55 to 75 per cent in Figs. 7(c) and (d)) and 35 to 55 per cent in Figs 7(e) and (f), 15 to 35 per cent in Figs 7(g) and (h) and 0 to 15 per cent in Figs 7 (i) and (j)

In Fig 7(a) it can be seen that most of the relative humidity measurements at high humidity were within ± 4 per cent of the average used. Meisei measurements below 2 km were an exception with a large negative bias of greater than 8 per cent shortly after launch. Fig.7 (b) contains an estimate of the daytime relative humidity measurements referenced to the same reference as at night. It has been assumed that there is little day-night difference in Snow-white measurements in this height range. This assumption gives a day-night difference in the other radiosonde measurements consistent with the day-night differences indicated by initial comparisons of radiosonde integrated water vapour amount (IWV) with GPS water vapour measurements

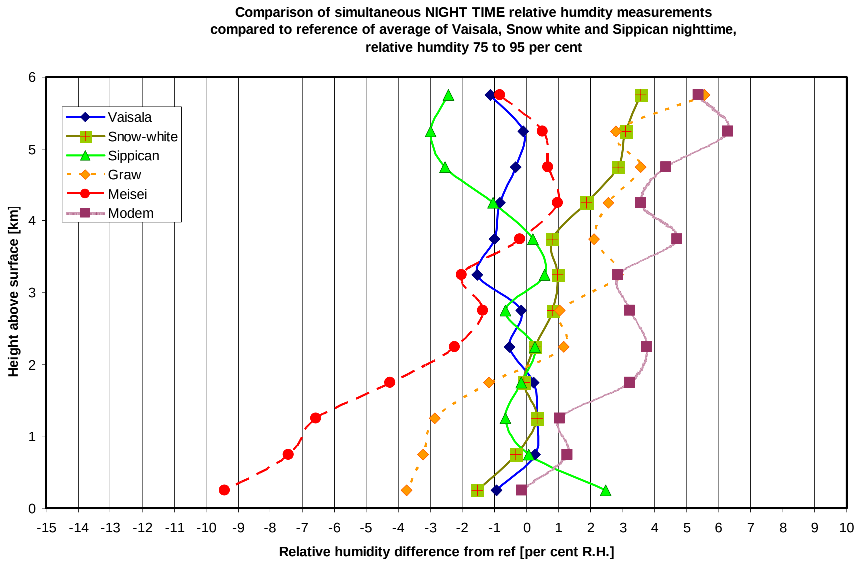


Fig. 7(a) Systematic bias for night time relative humidity , range 75 to 95 per cent, referenced to the average of Vaisala, Snow white and Sippican.

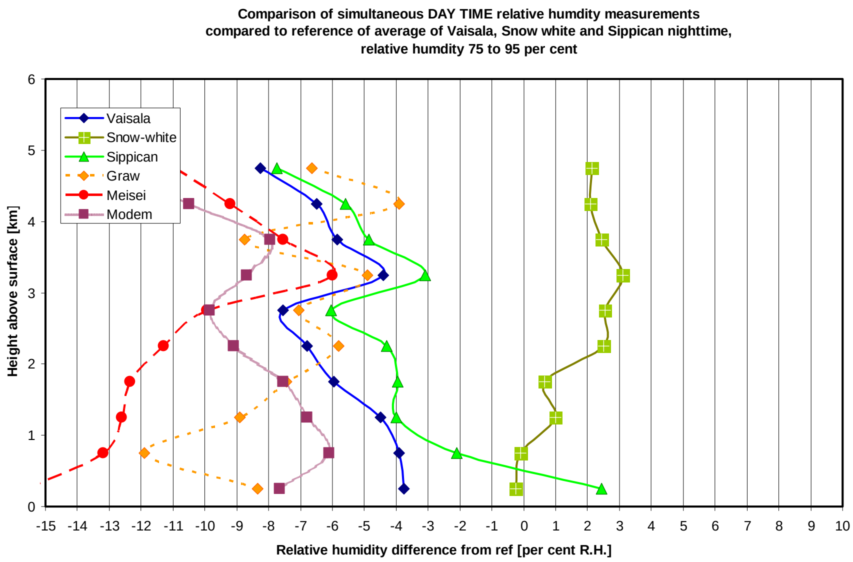


Fig. 7(b) Systematic bias for daytime relative humidity, range 75 to 95 per cent , referenced to the night time average of Vaisala, Snow white and Sippican,

Fig.7 (b) shows Sippican relative humidity had the smallest day-night difference with daytime measurements low by between 0 and 4 per cent relative to night. The largest day-night differences were present in Modem measurements with daytime low by between 8 and 16 per cent compared to night.

Day-night differences in Vaisala measurements varied from about -3 per cent near the surface to about -7 per cent at about 5 km. The Vaisala RS92 radiosondes used in Mauritius had improved protection against solar heating with the white glue and the bare copper near the sensors both aluminized, in contrast to current production models.

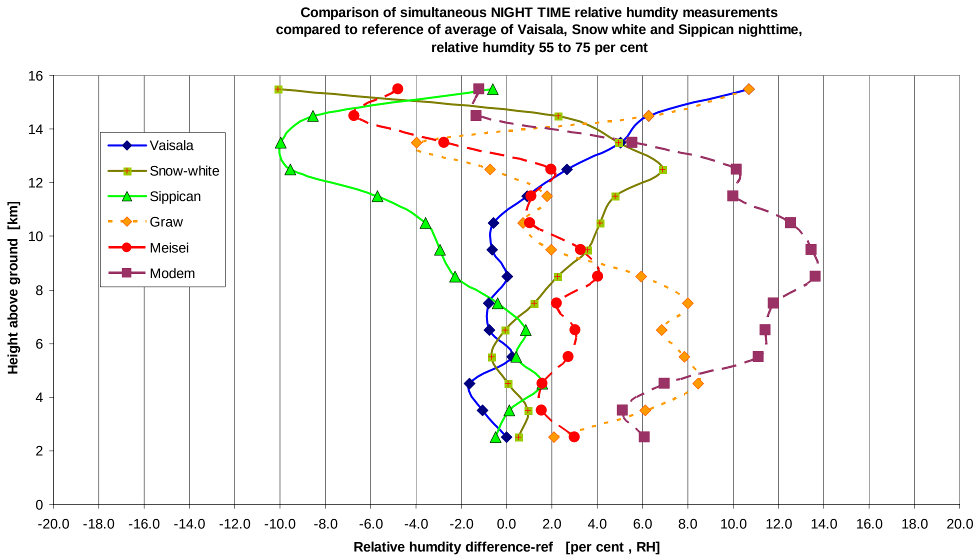


Fig. 7(c) Systematic bias for night time relative humidity , range 55 to 75 per cent, referenced to the average of Vaisala, Snow white and Sippican.

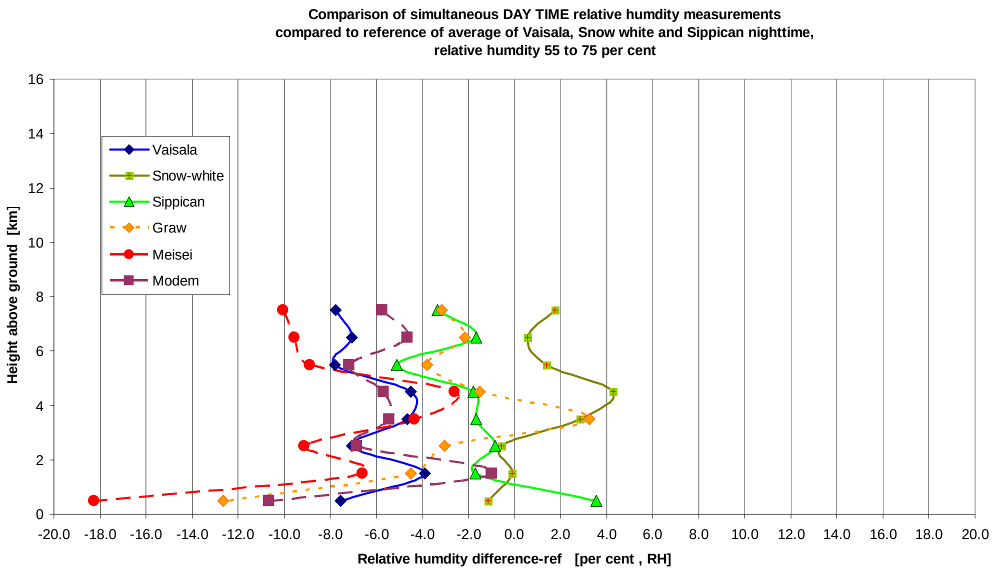


Fig. 7(d) Systematic bias for daytime relative humidity, range 55 to 75 per cent referenced to the night time average of Vaisala, Snow white and Sippican.

Comparison data from the relative humidity range 55 to 75 per cent, see Figs 7(c) and (d), were available over a much greater height range than for the highest relative humidity. This humidity range includes observations in middle and upper cloud.

Fig.7(c) shows that Meisei, Sippican, Snow white and Vaisala measurements agreed with the average to within ± 4 per cent from 2 to 11 km, i.e. for all temperatures down to -40 deg C. Exceptions were the positive bias of Modem measurements at night, more than 10 per cent for much of the height range, and to a lesser extent positive bias in Graw measurements. Above 11 km Vaisala and Snow white measurements agreed to 4 per cent down to temperatures of about -70 deg C. but Snow white measurements were much lower (20 per cent) than Vaisala at the lowest temperature. The reasons for the negative bias in Snow White relative to Vaisala can be seen in a typical individual comparison plot from the test.

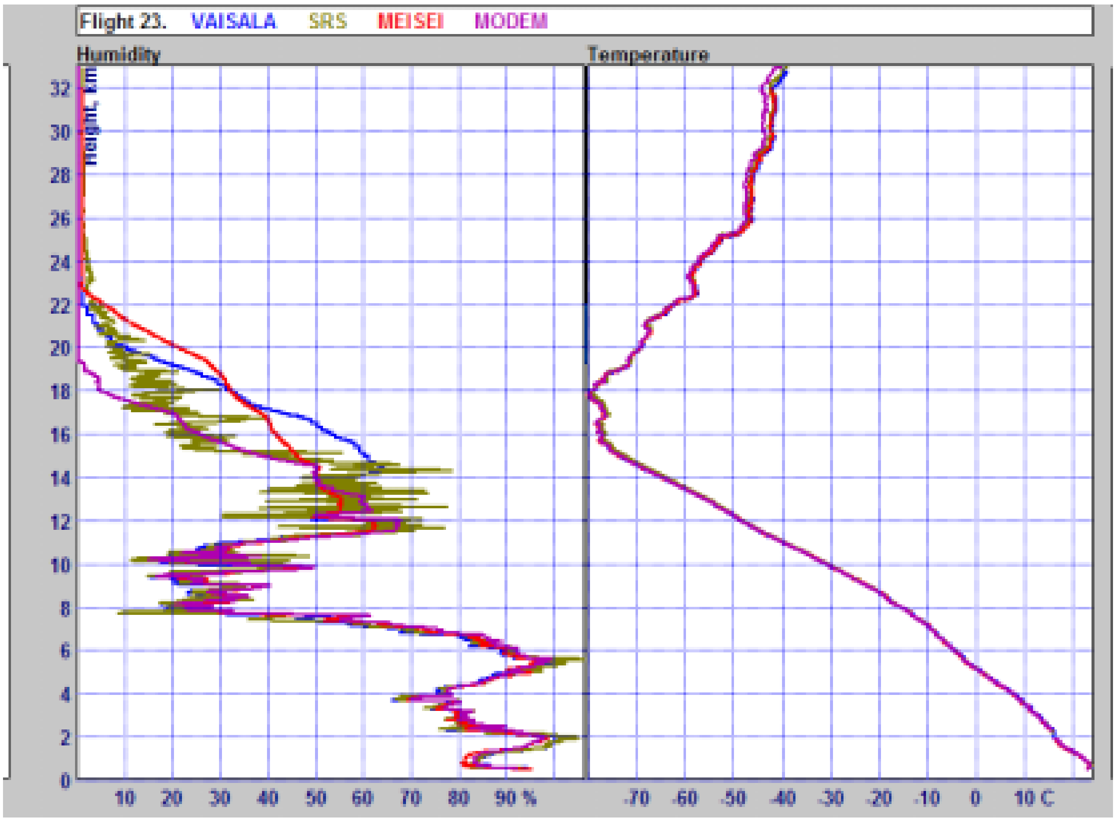
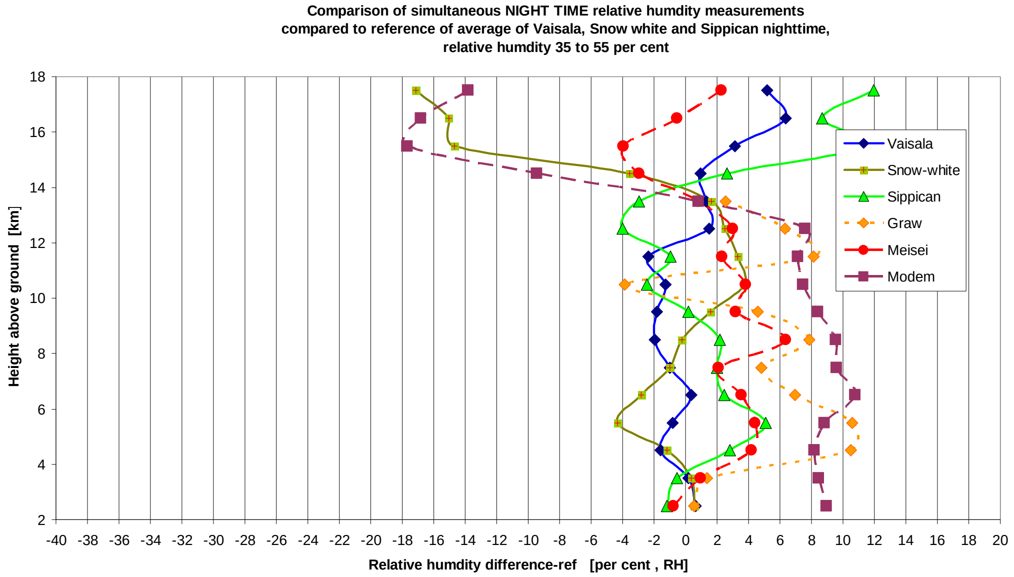


Fig.8 Night time relative humidity and temperature comparison between the Vaisala, Meisei –Modem group and SRS Snow White

In Fig. 8, the Vaisala relative humidity does not seem to fall fast enough after emerging from an upper cloud, possibly because of contamination from the cloud. Pulse heating of the Vaisala humidity sensors was limited to temperatures greater than -40 deg C on this flight.

Daytime relative humidity measurements shown in Fig.7(d) were again offset low relative to night time measurements , with Sippican showing the smallest day-night difference, 0 to -2 per cent, and Modem the largest of -12 to -17 per cent. In the case of Modem the night time measurements were very different from the other radiosondes, but the daytime measurements were similar to most of the other radiosondes. Vaisala day-night difference increased from about -4 per cent at 2 km to -6 per cent at 6 km.



.Fig. 7(e) Systematic bias for night time relative humidity, range 35 to 55 per cent, referenced to the average of Vaisala, Snow white and Sippican

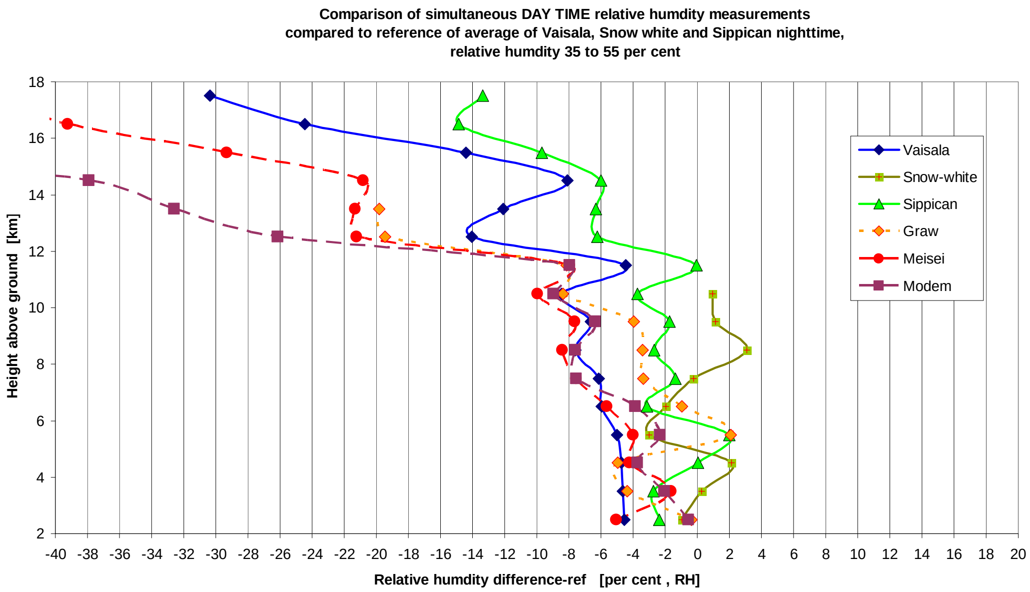


Fig. 7(f) Systematic bias for daytime relative humidity band 35 to 55 per cent referenced to the night time average of Vaisala, Snow white and Sippican.

The humidity range 35 to 55 per cent allowed measurements to be compared both day and night from 2 to 18 km; see Figs. 7(e) and 7(f). This range contains a significant number of samples from daytime upper cloud. Fig. 7(e) shows that Meisei, Sippican, Snow-white and Vaisala mostly agree to within ±5 per cent of the average at all heights up to 14 km. Modem and Graw again had a positive bias at night relative to the others. At heights above 14 km, Modem agrees most closely with Snow white. These two radiosondes had a negative bias of around 20 per cent relative to the other radiosondes. The magnitude of day-night differences deduced from Fig. 7(e) appear to range from Sippican, near zero up to 14 km, to Modem in the range from -11 to -16 per cent up to 14 km. Vaisala day- night differences were about -4 at 2 km, -6 at 10 km , and higher than 10 per cent at 14 km.

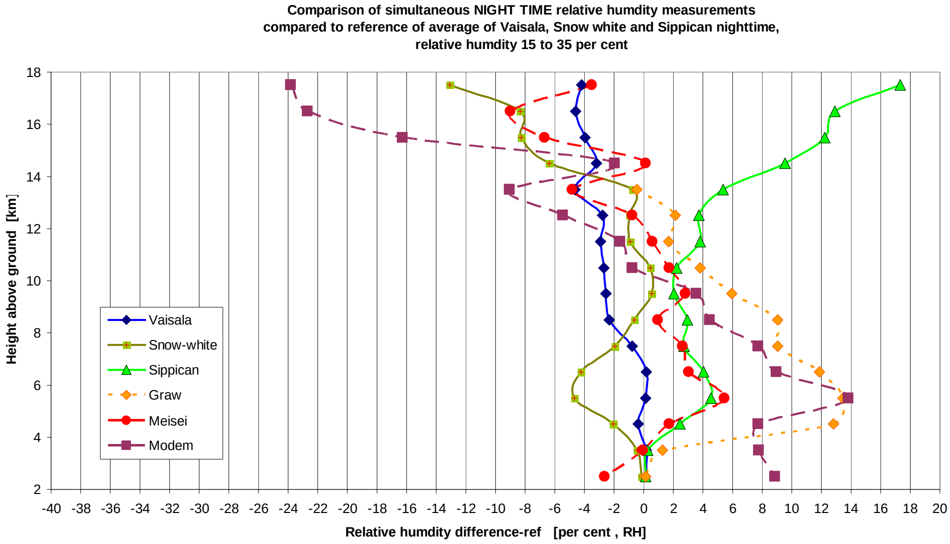


Fig. 7(g) Systematic bias for night time relative humidity, range 15 to 35 per cent, referenced to the average of Vaisala, Snow white and Sippican

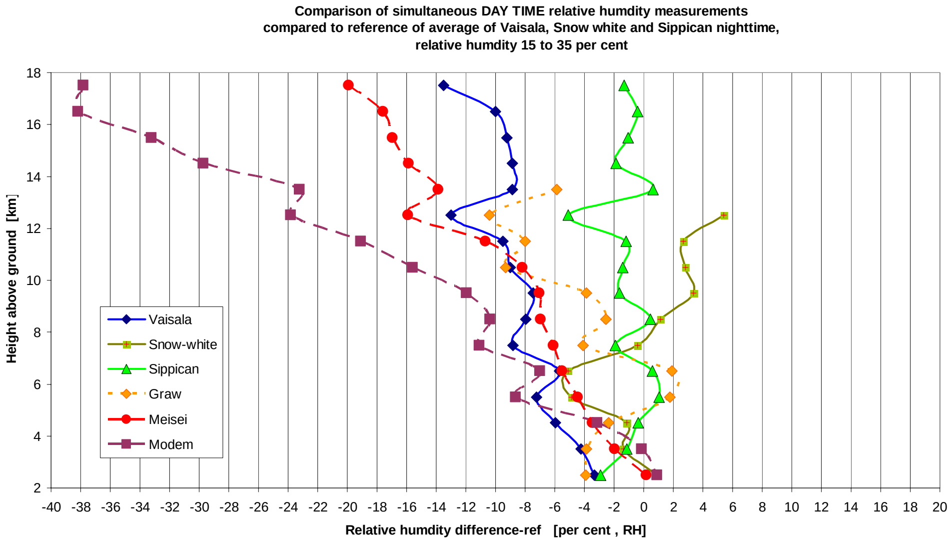


Fig. 7(h) Systematic bias for daytime relative humidity, range 15 to 35 per cent referenced to the night time average of Vaisala, Snow white and Sippican.

Figs. 7(f) and (g) show the night and day comparisons for the relative humidity range 15 to 35 per cent. At night all the radiosondes apart from Graw and Modem agree within ± 5 per cent from 2 to 14 km. At heights above 14 km, the differences between Vaisala and as Snow white were smaller than at higher humidity. However, Sippican shows a more pronounced positive bias than in higher humidity categories. Day-night differences for Vaisala in this low humidity range were about -3 at 2 km increasing to about -6 at 14 km.

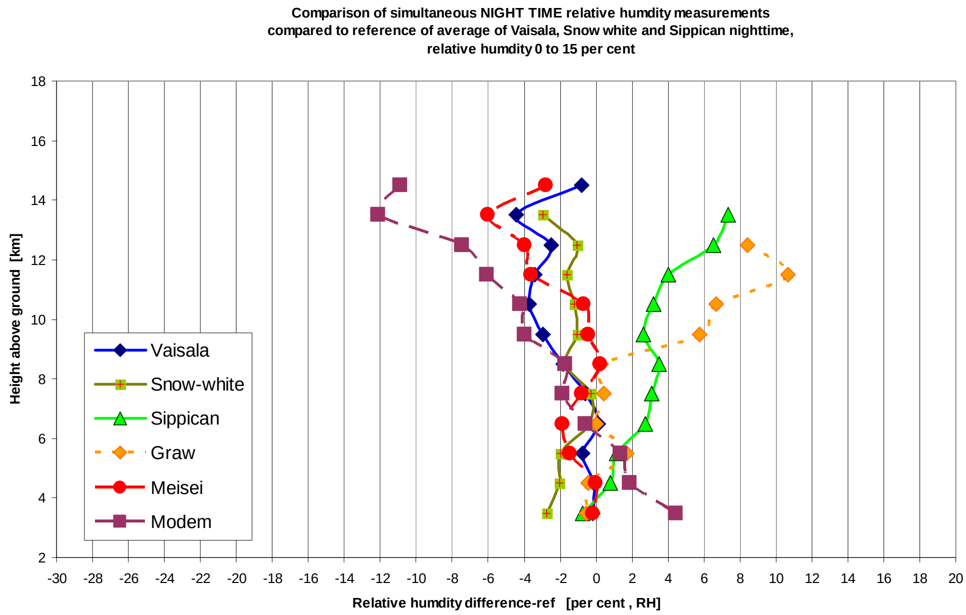


Fig. 7(i) Systematic bias for night time relative humidity, range 0 to 15 per cent, referenced to the average of Vaisala, Snow white and Sippican

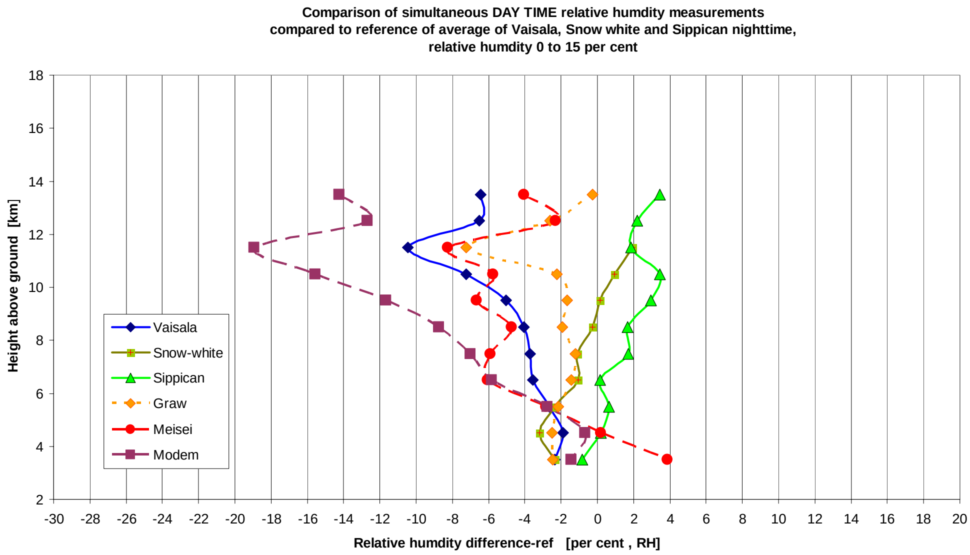


Fig. 7(j) Systematic bias for daytime relative humidity, range 0 to 15 per cent referenced to the night time average of Vaisala, Snow white and Sippican.

In this lowest humidity range, Meisei, Sippican, Snow-white and Vaisala again agree within ± 4 per cent between 3 and 12 km. Day-night differences were generally small.

The Sippican calibration for very low temperatures needed some improvement since there was a low bias at high humidity, see Fig. 7(c) and high bias at low humidity, see Figs. 7(g). Also, at heights between 18 and 20 km, Sippican had a positive bias of about 27 per cent relative to Vaisala at night for the relative humidity range 0 to 15 per cent.

Fig.9 (a) shows the standard deviations of the relative humidity differences with respect to Vaisala in the layer 0 to 4 km . Apart from Graw the magnitude of the standard deviations are consistent with random errors in relative humidity in the range 1 to 4 per cent, with lowest random errors found at night.

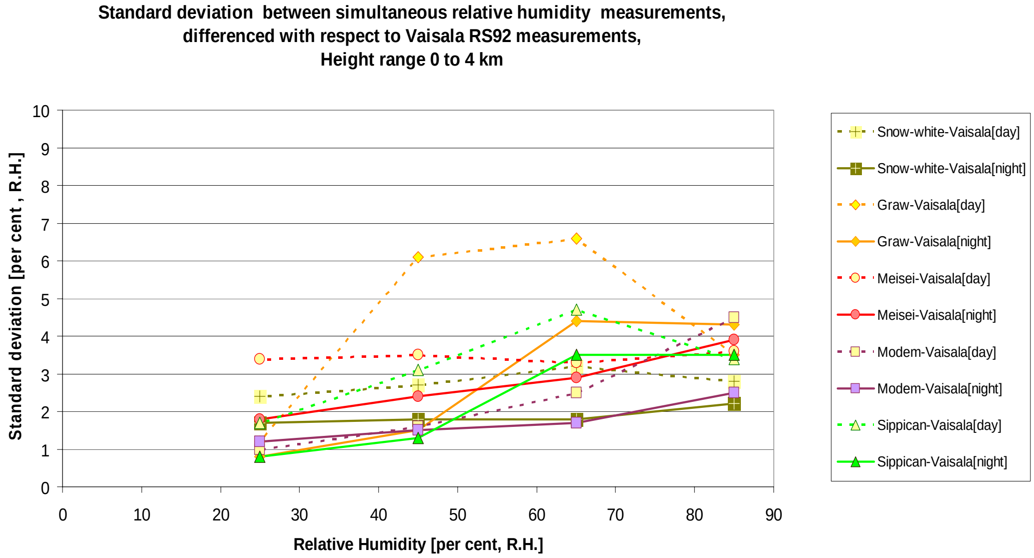


Fig 9(a) Standard deviations of differences with respect to Vaisala, 0 to 4 km

Fig. 9(b) and (c) show that the random errors of the relative humidity sensors increased slightly for the height range 8 to 12 km to values between 2 and 6 per cent, but only Meisei shows low standard deviations relative to Vaisala in the height range 12 to 16 km. This probably means that the random errors in Meisei measurements were similar in nature to those of Vaisala. The larger standard deviations with respect to Snow white measurements were caused by errors in Vaisala as well as random errors in Snow white, see Fig.8. Thus , it is probable that random errors in Vaisala measurements at temperatures lower than -70 deg C were in the range 5 to 10 per cent.

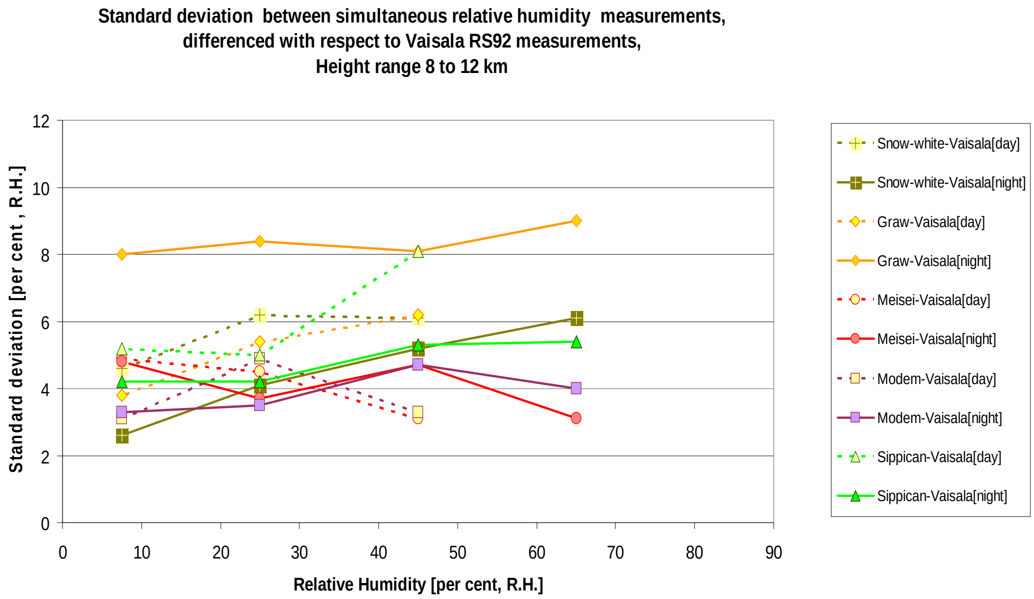


Fig 9(b) Standard deviations of differences with respect to Vaisala, 8 to 12 km

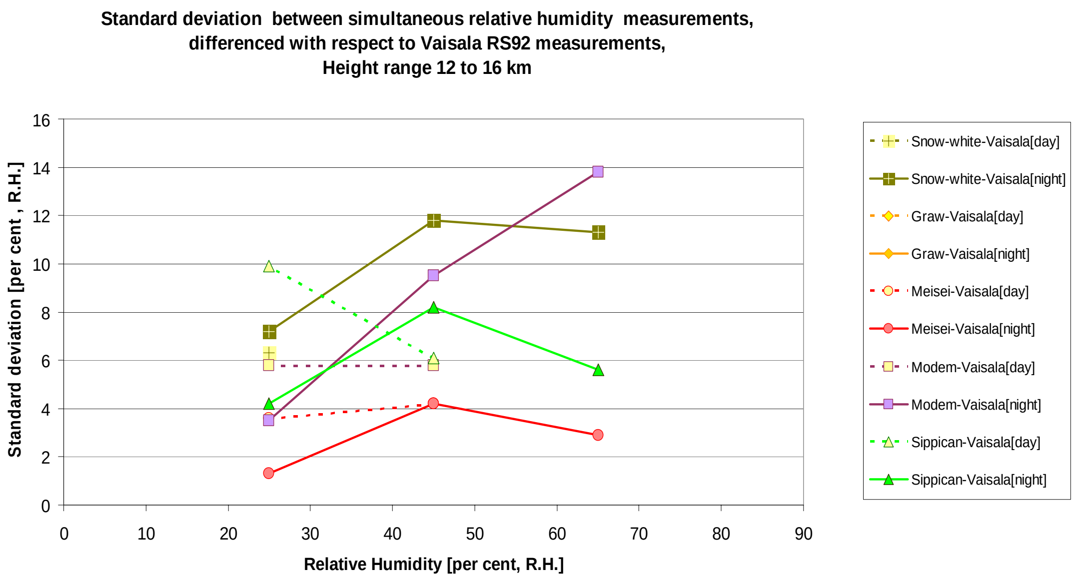


Fig 9(c) Standard deviations of differences with respect to Vaisala, 12 to 16 km

8. Simultaneous wind comparisons

There were no significant problems with this generation of GPS wind measurements. The main differences between the systems see Figs.10 (a) and (b) for comparisons between U and V wind components, arose from the different types of filtering used to remove the pendulum motion of the radiosonde under the balloon. The filtering of the Meisei measurements averaged over too long a period to give optimum performance in the stratosphere, and some of the test flights were too long for the battery design, so some Meisei measurements deteriorated in quality at the uppermost heights.

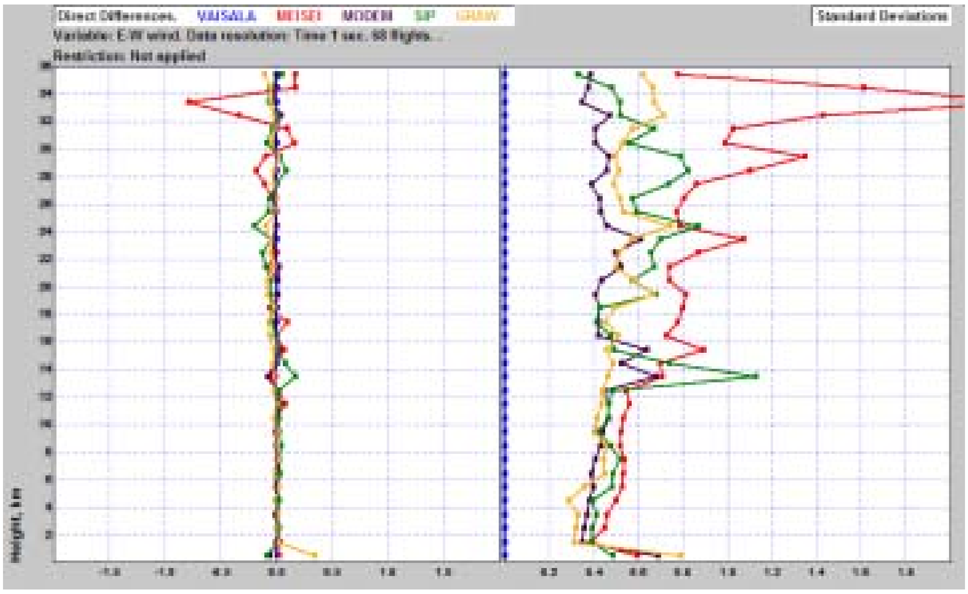


Fig. 10(a) Systematic bias and standard deviations of simultaneous comparisons between U components [ms^{-1}].

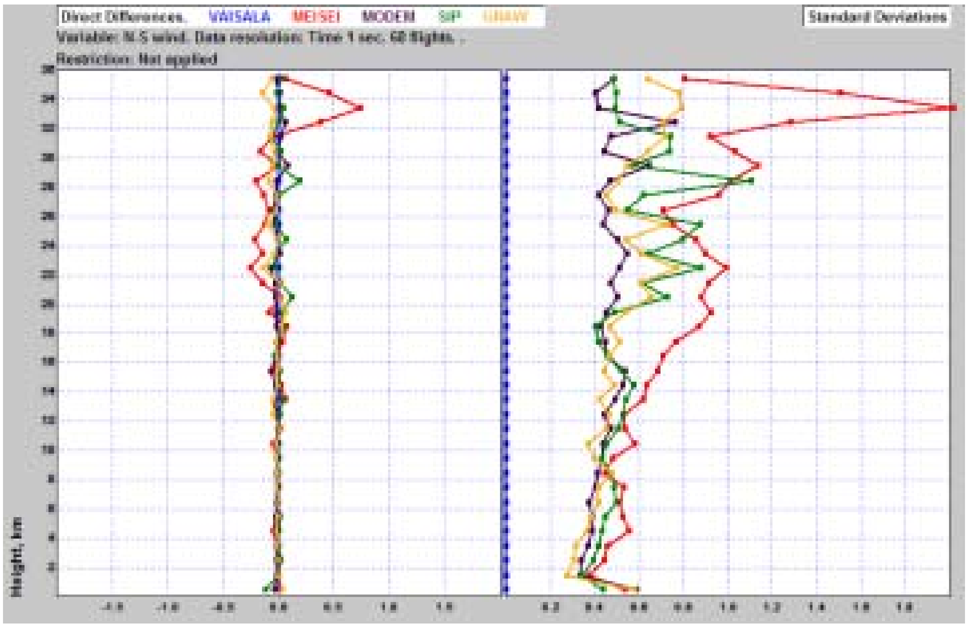


Fig. 10(b) Systematic bias and standard deviations of simultaneous comparisons between V components [ms^{-1}].

Typical random errors in wind component [u, v] measurements must have been less than or equal to 0.3 ms^{-1} for all systems at all heights apart from Meisei in the stratosphere. Systematic bias between measurements from different systems was negligible. These results were obtained with minimal editing of the wind profiles by the WMO Supervisors.

Thus, it is concluded that the new generation of GPS radiosondes should be capable of very accurate wind measurements in tropical locations, with minimal missing data. This will be true even when there are strong upper winds as in Mauritius with wind speed higher than 40 ms^{-1} at heights around 30 km.

Acknowledgements

This test could not have been successfully implemented without a very large amount of work from many people, often at unsocial hours, including

- Many staff from the Meteorological Services, Mauritius,
- the six manufacturers,
- colleagues representing the COST meteorological projects in Europe [H. Richner, Switzerland: T. Oakley, M. Oldfield and D. Lyth (UK)]
- support from the WMO secretariat through M. Ondras.