Investigation into the consistency of radiosonde and GPSRO observations and their application in NWP at the Met Office

26/03/2010

Forecasting R&D Technical Report No. 541

by Michael P. Rennie

Document History

	Name	***Position***	***Date***	***Comment***
Prepared by:	Michael Rennie	Scientist SA	26/03/2010	Final version
Edited by:	Chris Burrows	Scientist SA	31/10/11	Removed internal links
Reviewed by:	Dave Offiler	Manager SASG	February 2010	Amendments made
Reviewed by:	Rick Rawlins	Manager Global DA	February 2010	Amendments made
Approved by:	John Eyre	Head SA	February 2010	Amendments made

Contents

Abstract

This report analyses radiosonde temperature, humidity and pressure data (that are available for assimilation at the Met Office) in detail, by comparisons with the global model and with GPSRO refractivity data. The majority of radiosonde models show good consistency with the global model and GPSRO — Vaisala RS92 are the most consistent. We found that a significant amount of biased radiosonde temperature data is being assimilated in the stratosphere (biases provided with data, not in Met Office processing), particularly over Russia, which may be degrading forecasts and also be limiting GPSRO's ability to accurately control stratospheric temperature biases. It appears the OPS quality-control check against background forecast temperatures should be made stricter in the upper troposphere and stratosphere. Also a positive bias for the OPS processing of radiosonde temperature observations onto model levels in the stratosphere (around 30 km) of around 1 K was found. This is due to the layer-mean radiosonde temperatures being misplaced in the vertical, since the model θ levels are not at the centre of surrounding ρ levels in geopotential height space, when the layer-mean method assumes they will be applied at the centre of the layer. Also we found that radiosonde data provided with relatively low vertical resolution (low number of significant levels) may be over-weighted in the assimilation system.

1 Introduction

GPSRO (Global Positioning System Radio Occultation) observations have been operationally assimilated into Met Office NWP models since late 2006 (Rennie, 2010). GPSRO is a limb-geometry remote sensing technique, whereby the time delay of GPS radio signals that have passed through the limb of the Earth's atmosphere (as received on a low-earth orbit satellite), are used to determine vertical profiles of measurements related to refractive index. An in-depth description of the technique is given by Kursinski et al. (1997).

GPSRO measurements have small systematic errors compared to other available observations that provide pressure, temperature, geopotential height and humidity information. The data also have a high vertical resolution — typically quoted as less than 1 km. However, the observations have relatively poor horizontal resolution, of several hundred kilometres; also the interpretation of the observations as local vertical profiles is an approximation which has decreasing accuracy in the moist lower-troposphere, where the atmosphere tends to be more spherically asymmetric. Also, GPSRO observations have a dependency on total pressure, temperature and water vapour pressure (through the formula for microwave-frequency refractive index), therefore in variational data assimilation we rely on background covariances to separate the components.

Radiosondes provide in-situ measurements of the atmosphere. Modern radiosondes measure or calculate the following variables: total pressure, altitude, geographical position (latitude/longitude), temperature, relative humidity, wind speed and direction. For this investigation we are interested in temperature and humidity measurements, with pressure as the vertical co-ordinate; variables which can be forward modelled to GPSRO equivalent observations, in particular for this report: microwave-frequency refractive index.

GPSRO observations produced by independently developed receivers and processing techniques have been shown to be consistent to a high degree e.g. see von Engeln et al. (2009), Schreiner et al. (2007) and Rennie (2010); this high level of consistency is lacking between many types (and even factory batches) of radiosondes — see for example Steinbrecht et al. (2008), which shows biases in the calculated geopotential heights (compared to GPS-derived heights available on these radiosondes) varying between different Vaisala RS92 manufacturing batches. Even so, state-of-the-art radiosondes such as Vaisala RS92 and Sippican MARK II show small biases when compared with GPSRO, as will be shown in this report. However, those radiosondes deployed in Russia and previously India (in 2010 India changed to using MODEM radiosondes, which are of better quality) are prone to fairly large biases. These biases will be analysed by comparison with GPSRO in this report. He et al. (2009) showed that temperature measurements for Vaisala RS92 and Shanghai Radio radiosondes agree very well with FORMOSAT-3/COSMIC GPSRO data, whereas MRZ (Russia) and VIZ-B2 (USA) showed large biases probably related to diurnal radiative effects.

This investigation was performed due to concerns that GPSRO may be biased relative to some types of radiosondes, particularly in the stratosphere. We wish to understand whether the biases are in the data *per se* or in the assimilation method (processing/forward model). This topic is also of interest since the standard verification system for Met Office NWP trials is to a large extent based upon comparisons of RMSE (root-mean-square errors) differences between forecasts and the assimilated radiosonde observations. Recent investigations by other NWP centres indicate that aircraft measurements are a source of bias in the troposphere. This however is beyond the scope of this report.

Section 2 describes the methods used to analyse the data, section 3 shows results from the investigation, section 4 discusses the effect of processing/interpolation of radiosonde temperature information onto model levels and the final section provides conclusions.

2 Methods

Observation and NWP model data are written out to NetCDF files in a branch of the Met Office OPS (Observation Processing System) which outputs arrays of background fields and observation data within the GPSRO and radiosonde OPS subroutines. The NetCDF files are read into IDL and processed when required to generate the plots in this report. We output the radiosonde data before and after the OPS temperature correction (for radiation biases) and the processing of observation data onto model levels has been performed. See OSDP5 (2008) for the details of the radiosonde processing.

The approach of the investigation is to:

1. Compare radiosonde data to their NWP model equivalents on a profile-by-profile basis, before and after OPS processing.

2. Examine statistics of observation minus background ($O - B$) departures for different types of radiosonde, splitting the data into day and night-time, before and after OPS processing (also done in refractivity space for comparison to GPSRO $O - B$ statistics).

3. Comparing coincident (in space and time) radiosonde-derived microwave-frequency refractivity and GPSRO refractivity data.

The report uses (unless stated otherwise) the 70 level global model (Met Office UM) that went operational in November/December 2009 following PS22.

Some of the plots produced in this report are available in near real-time on the GRAS SAF website. See:

`http://www.grassaf.org/monitoring/index.php`

2.1 GPSRO observations

As discussed in the introduction, GPSRO measurements are ultimately derived from the time delay of GPS radio signals, due to the presence of the Earth's atmosphere. From the excess Doppler shift of the radio signal, the amount of bending of the rays (refractive bending angles of the radio waves as a function of impact parameter) can be derived assuming the atmosphere is spherically symmetric (which is a very good approximation for a large fraction of atmospheric conditions). Hence an Abel transform integral is used to calculate refractive index as a function of geometric height above the WGS-84 reference ellipsoid, given the bending angles and impact parameters. *A priori* climatological information is merged with the observations to reduce noise and extend bending angles for the upper boundary of the Abel integral. However its influence is not significant at the levels where comparisons against radiosondes will be made i.e. up to $\sim$30 km. The geometric heights of the refractive index product can be converted to geopotential heights using the gravity potential model that is appropriate to the ellipsoid e.g. see http://mtp.jpl.nasa.gov/notes/altitude/altitude.html.

The GPSRO refractivity N, $N = 10^6(n-1)$ where n is the refractive index, is typically provided to users on a $\sim$200 m vertically spaced grid, although the grid can vary from occultation to occultation and by processing centre. Therefore, we interpolate the refractivity onto a fixed 200 m spacing geopotential height grid using the IDL INTERPOL function with log-linear interpolation (given refractivity is an approximately exponentially decaying function with height, the dry term of refractivity being linearly proportional to density).

2.2 Radiosonde observations

Radiosonde pressure, temperature, dew-point temperature (which is calculated by the data providers from the measured relative humidity) and geopotential height information are the data of interest for this report that are provided to the Met Office. Some radiosondes now provide GPS-derived altitude to locate the measurements in the vertical. However these data are not available for those stored in the MetDB as yet — this method is more accurate than calculating geopotential heights using the radiosonde pressure, temperature and dew-point. Some radiosonde models calculate the pressure profile using the GPS-height, the ground pressure and the vertical profile of temperature and humidity. This state-of-the-art method significantly increases the accuracy of the pressure-data while reducing the costs as a pressure-sensor is not needed anymore — however, it does assume hydrostatic equilibrium. The dew-point temperature is converted back to relative humidity in the OPS for data assimilation.

If the radiosonde GPS-derived heights were consistently available it might be beneficial to change the radiosonde assimilation method from assimilating temperature and relative humidity using pressure to locate the observations vertically in the model, to assimilating temperature, pressure and humidity as a function of geopotential height — with similarities to the GPSRO method of assimilating refractivity as a function of geopotential height.

The radiosonde temperature and dew-point are provided on significant levels. Significant levels allow data at points in between to be reconstructed to sufficient accuracy by linear interpolation of temperature in log-pressure, whereas the geopotential height values are generally provided (if at all) on standard pressure levels (1000, 925, 850, 700, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, 20 and 10 hPa). The Manual on Codes documentation, see World Meteorological Organisation (WMO) (2009), provides a fairly complicated explanation for deriving significant levels, but the main points are that: As far as possible, these additional levels shall be the actual levels at which prominent changes in the lapse rate of air temperature occur. The accuracy of the linearly interpolated value between significant levels should agree with the observed temperature by less than $1\,°\mathrm{C}$ in the troposphere and $2\,°\mathrm{C}$ in the stratosphere. This seems like a fairly relaxed constraint given the size of typical errors encountered in modern NWP.

Given GPSRO's high vertical sampling, it is unsatisfactory to use only the radiosonde data on standard levels for comparisons. Therefore, we calculate the radiosonde geopotential heights independently by summing thicknesses between significant levels according to the method outlined below.

The radiosonde dew-point is converted to water vapour pressure, p_w, using the Goff-Gratch formulae with the radiosonde dew-point temperature, T_d, as input (details of calculation are not shown). The same procedure is done in the OPS when converting dew-point temperature to relative humidity. The virtual temperature, T_v, is then calculated as follows:

$$T_v = \frac{T}{1 + (\frac{M_w}{M_d} - 1)x_w} \tag{1}$$

where $x_w = p_w/p_{tot}$ is the water vapour volume mixing ratio, p_{tot} is the total air pressure measured on the radiosonde, T is the temperature measured on the radiosonde, $M_w = 18.01534\,\mathrm{gmol}^{-1}$ is the molar mass of water and $M_d = 28.9644\,\mathrm{gmol}^{-1}$ is the molar mass of dry air.

The geopotential heights are then calculated using the hypsometric equation: equation 2, which combines the hydrostatic equation with the ideal gas law, where $R_d = 287.0531\,\mathrm{JK}^{-1}\mathrm{kg}^{-1}$ is the specific gas constant for dry air, $g = 9.80665\,\mathrm{ms}^{-2}$ is the standard value for the gravitational acceleration, $\overline{T_v}$ is the layer mean virtual temperature (calculated as the average of layer boundary temperatures) and p_1, p_2 are pressures bounding the layer.

$$H_{12} = \frac{R_d}{g}\overline{T_v}\ln\left(\frac{p_1}{p_2}\right) \tag{2}$$

We also need to use the first available geopotential height provided with the observation, as this includes the station geopotential height.

Radiosonde-derived microwave-frequency refractivity is then calculated using the empirical formula, requiring total pressure, temperature and water vapour pressure, with constants $k_1 = 77.6\,\mathrm{K/hPa}$ and $k_2 = 3.73 \times 10^5\,\mathrm{K}^2/\mathrm{hPa}$ (Smith and Weintraub, 1953), the same values used in GPSRO forward model.

$$N = k_1 \frac{p_{tot}}{T} + k_2 \frac{p_w}{T^2} \tag{3}$$

For monitoring purposes we then calculate the specific humidity (mass-mixing ratio):

$$q = \frac{x_w M_w}{x_w M_w + (1 - x_w) M_d} \tag{4}$$

The specific humidity and water vapour pressure values are set to zero above ∼20 km since the algorithms do not apply in this range and spurious dew-point temperature data above these altitudes often led to unrealistically high humidity values — which then adversely affects the calculated refractivity.

The temperature, specific humidity and radiosonde-derived refractivity are interpolated onto the same 200 m spaced geopotential height grid as for the GPSRO data, allowing comparisons to be made. Linear interpolation with geopotential height is used for temperature and specific humidity, log-linear for refractivity. There may well be significant errors introduced in the interpolation of radiosonde data that is provided at a particularly low vertical resolution — this should not be an issue for most types of radiosondes (given that significant levels are designed for accurate interpolation to other levels in between).

Also written out to NetCDF files are the temperature after OPS processing of radiosonde data onto model levels (including the radiation correction, if applied), along with the QC (quality control) flags and final PGE (probability of gross error) values, which determine whether the observations would be assimilated in VAR — note we use the latest OPS stationlist, so the QC should be appropriate to that time. These methods are explained in OSDP5 (2008).

Swinbank and Wilson (1990) explain the procedure for the OPS processing of radiosonde temperatures onto model levels. Radiosonde temperatures (with given pressures) are linearly interpolated in log-pressure onto the surrounding model ρ level pressures. Then a summation of thicknesses between ρ levels is done using the observation temperatures and the ρ level interpolated observation temperatures, using the hypsometric equation (see equation 2). The summed thickness between ρ level pressures are then converted to a layer-average observation temperature (actually potential temperature, since it uses an Exner version of the hypsometric equation). The θ observation profile is then assimilated on the model θ levels (between the surrounding ρ levels). The method ensures that the geopotential thickness across a model layer is consistent with the thickness between a pair of corresponding pressure levels from a temperature sounding. In practice it looks very similar to a linear interpolation of temperature in geopotential height, as can be seen in the plots of the next section. For monitoring purposes we assess the temperatures i.e. the potential temperature values converted back to temperature, given the appropriate Exner values.

The solar elevation angle of the radiosondes are calculated, and used to determinate whether it is a day-time (greater than 0 degrees) or night-time (less than 0 degrees) ascent — some investigations use a stricter criterion for night/day e.g. less than -10/greater than 10 degrees).

3 Monitoring Results

This section shows results from the analysis performed of the data using the methods described in the previous section.

3.1 Examination of Individual Radiosonde Ascents

Here we provide some examples of radiosonde ascents which epitomise the data that is provided for operational assimilation in VAR.

Figure 1 shows a radiosonde ascent with the equivalent operational global model background data for comparison. The top-left of the figure shows the temperature information: 'Obs temperature' is the radiosonde significant and standard level data as a function of pressure (as supplied by the data centres), 'Obs dew-point' is the corresponding dew-point temperature (as supplied by the data centres), 'Global model temperature' is the temperature on θ levels of the global model 3–9 hour background forecast and 'OPS processed obs temp' is the observation temperature after the OPS processing onto model θ levels. Note the OPS processed potential temperature observation value is converted to temperature using:

$$T = \theta \Pi_\theta \tag{5}$$

Where $\Pi_\theta = \left(\frac{p_\theta}{p_0}\right)^\kappa$ is the Exner interpolated to θ levels using the OPS routine Ops_CXP (which performs a linear interpolation of Exner with geopotential height, from ρ to θ levels), p_0 is the reference surface pressure and $\kappa = R_d/c_p$, where c_p is the specific heat capacity of dry air at standard temperature and pressure, at constant pressure. The bottom right-hand plot of Figure 1 uses the same data, but now plotted as a function of the calculated geopotential height for the observation and the θ level geopotential height for the model, rather than pressure.

Note that the pressure plot shows qualitatively the same behaviour as the geopotential height, which implies there is no inconsistency in the geopotential height calculation (e.g. no vertical displacement of the observation profile relative to the background profile); at least as seen on this scale. Also, note there does not appear to be any inconsistency between the raw observation temperature data and the OPS processed onto model level temperature data (for this example at least).

The particular example in Figure 1 is a Vaisala RS92/Autosonde (Finland) radiosonde. They are the most widely used radiosonde (nearly five times as many deployed as the next most abundant: Sippican), and are used most commonly in Europe. For this example, the temperature agrees particularly well with the model up to ∼23 km (∼30 hPa), where the observation data stops. Also note the high vertical resolution of the data particularly in the stratosphere — high vertical frequency of significant levels. The top-right plot shows the observation specific humidity (as derived from the dew-point temperature) along with the model specific humidity — again, the agreement is good. Of course the model and the observations are imperfect, and we expect some departures due to random and systematic measurement/model errors and also representativeness errors. The lower-left plot shows the radiosonde-derived refractivity (on a logarithmic scale) as a function of pressure, again with the OPS processed (orange) and background model (blue) equivalents (try zooming-in to the PDF version if the plots are hard to see when printed).

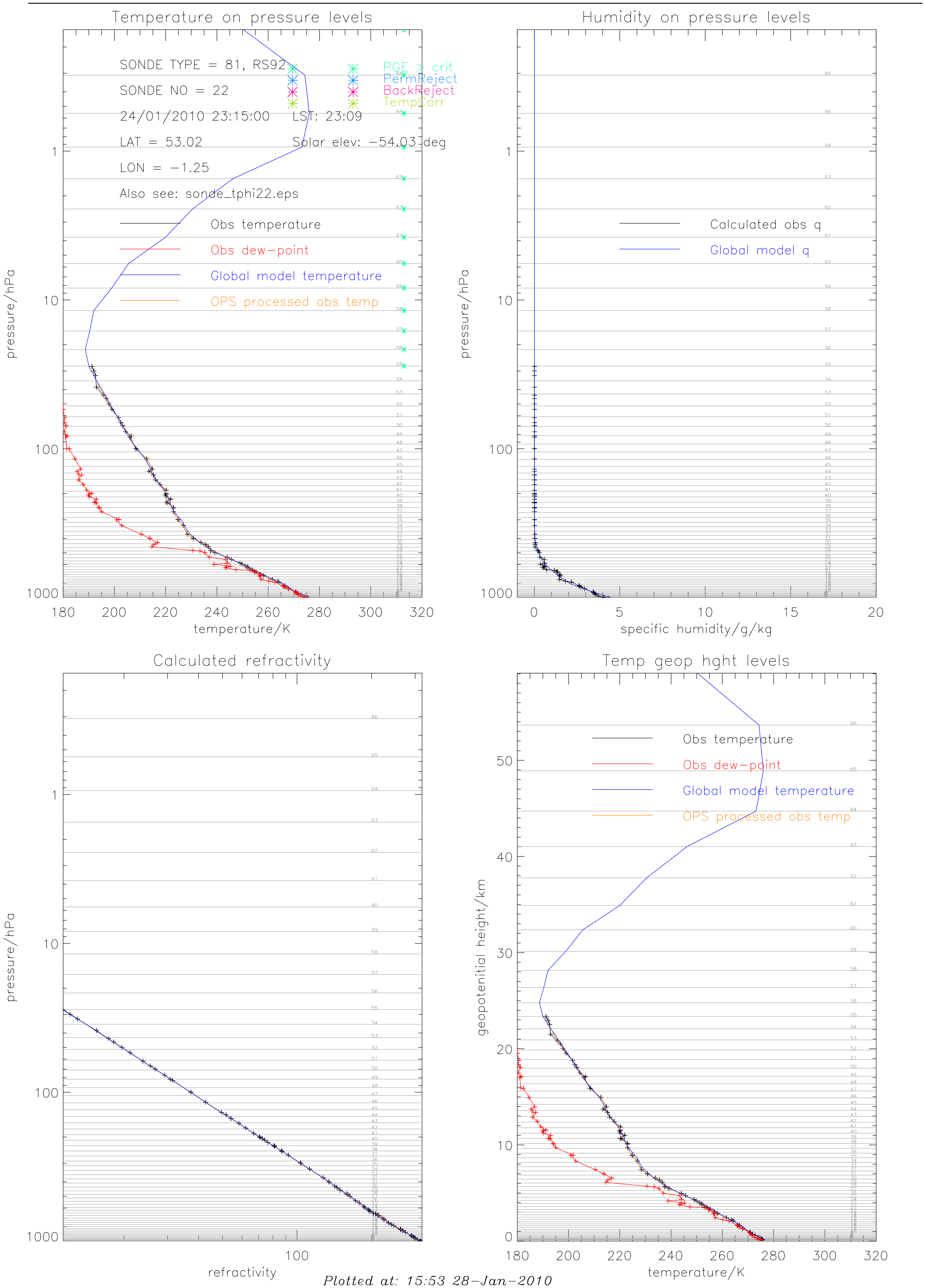


Figure 1: A good quality radiosonde profile of temperature and dew-point temperature — a Vaisala RS92 over England. Plot explained in main text. *Zoom-in on PDF version to see detail.*

Figure 2 is a Sippican MARK II (with chip thermistor, pressure and carbon element) radiosonde. This type is employed mostly in the USA. This ascent also appears to be of good quality given its similarity to the background forecast. There are some departures from the background humidity below 300 hPa and in the stratospheric temperatures, however this behaviour is often seen — probably an example where the assimilation of the data will provide benefit. Again, there is an abundance of significant levels in the stratosphere, revealing lots of vertical structure. The dew-point temperatures in the upper-troposphere and stratosphere look unrealistically large though, and would give unrealistically large specific humidities if the calculation was not stopped at $\sim$20 km. Within VAR radiosonde humidity information is cut-off in the vertical when the temperature drops to about -60°C. Therefore the unrealistic dew-point information should not be assimilated. Note also there is an indication of the OPS-processed radiosonde temperatures being slightly positive biased compared to the input data at the highest altitudes.

Figure 3 shows a radiosonde profile of type AVK-MRZ-ARMA (Russian Federation), used widely over Russia in a fairly dense network. Experience from observing many ascents suggests this is a case where the observation temperatures are showing significant biases (by up to $\sim$5 K), particularly in the stratosphere. One should note that not all AVK radiosondes are as biased as this. However a significant proportion of them do show what appears to be biased stratospheric temperatures. This radiosonde also provides data at a poor vertical resolution (very few significant levels compared to RS92 or Sippican). Perhaps this is because they are only just meeting the WMO criteria for significant levels (whereas other radiosondes types exceed the WMO criteria by a large margin) or perhaps because there were not any 'measured' fluctuations deemed significant between these levels due to instrument problems. Such a lack of vertical structure is probably not representative of the true atmosphere and not greatly suited to the model's vertical resolution.

As discussed in section 2.2, the OPS processes the observation data e.g. temperature, onto all the θ levels between the bottom and top of the ascent. In the low vertical resolution AVK example from Figure 3, there are (in the upper troposphere) three θ levels between adjacent observation levels. Since in VAR the processed obs (θ level radiosonde observations) are treated independently i.e. vertical error correlations are ignored, increasing the vertical resolution of the model would increase the weight given to the observations in VAR — the same two observations of temperature would be interpolated to a larger number of θ level observations. This effect was noticed by Rick Rawlins (DAE, pers. comm.) in the change of the global model from 50 to 70 vertical levels, from the larger radiosonde cost functions. Given that it appears the AVK radiosondes' significant levels are not realistic, especially when compared to RS92 and Sippican, we should probably reduce their influence by giving them lower weight in VAR.

Perhaps a better assimilation method would be for the data providers to provide the raw data. One could then interpolate the model to the observation vertical locations and assume some vertical correlation of the errors for a profile, in a similar manner to the GPSRO forward model, which interpolates model refractivity to the observation geopotential heights. This method would stop the weighting of radiosondes from becoming inappropriate following changes to the model vertical resolution. However, the vertical interpolation is part of a forward model, thus it would make the VAR code for radiosondes slightly more complicated, and it would complicate the monitoring of radiosonde data from VAR stats (which is done on model levels). Perhaps as a simple alternative there could be a method of weighting the radiosonde errors according to the typical vertical resolution for that type of radiosonde.

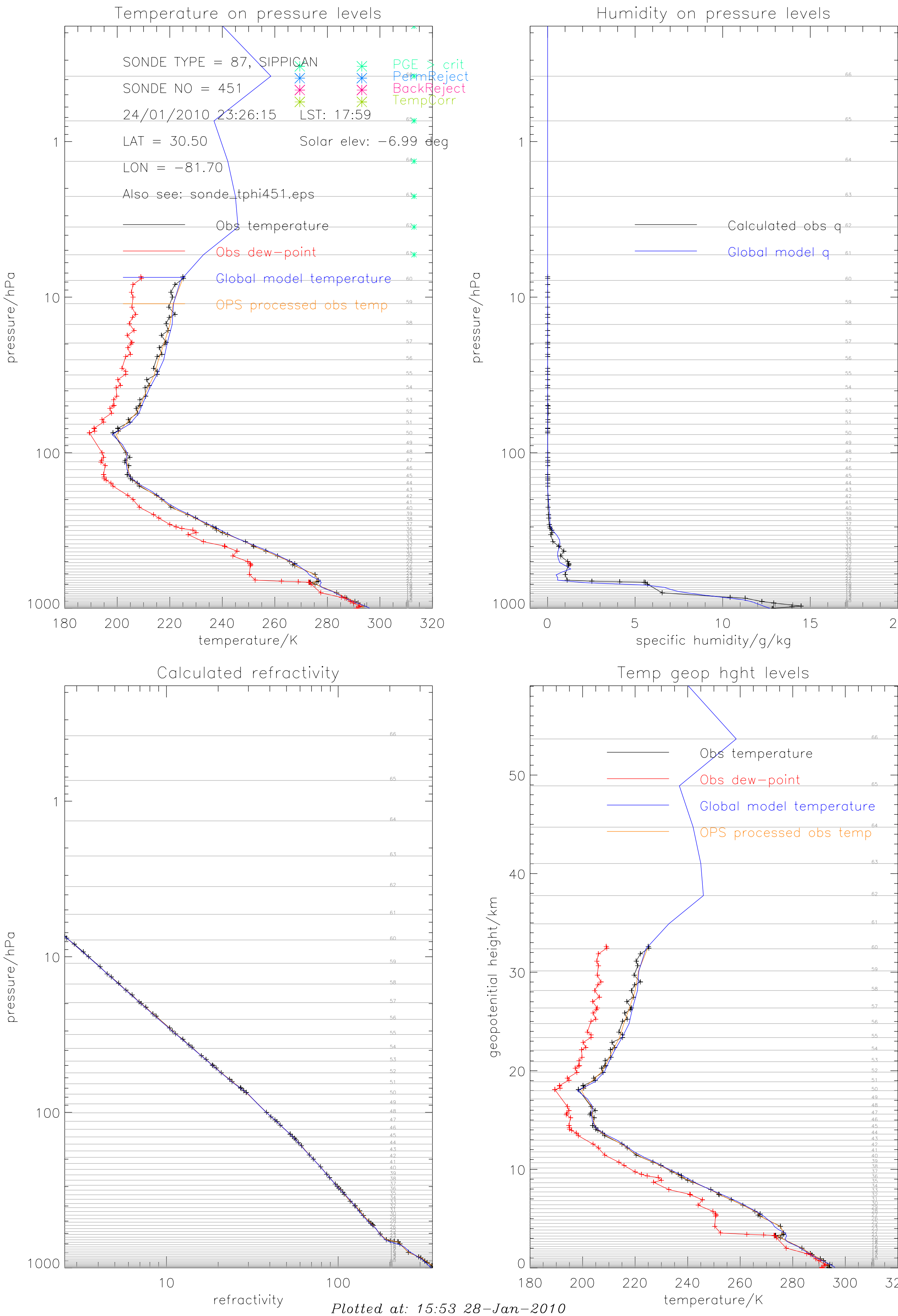


Figure 2: A good quality radiosonde profile of temperature and dew-point temperature — a Sippican over Florida (USA).

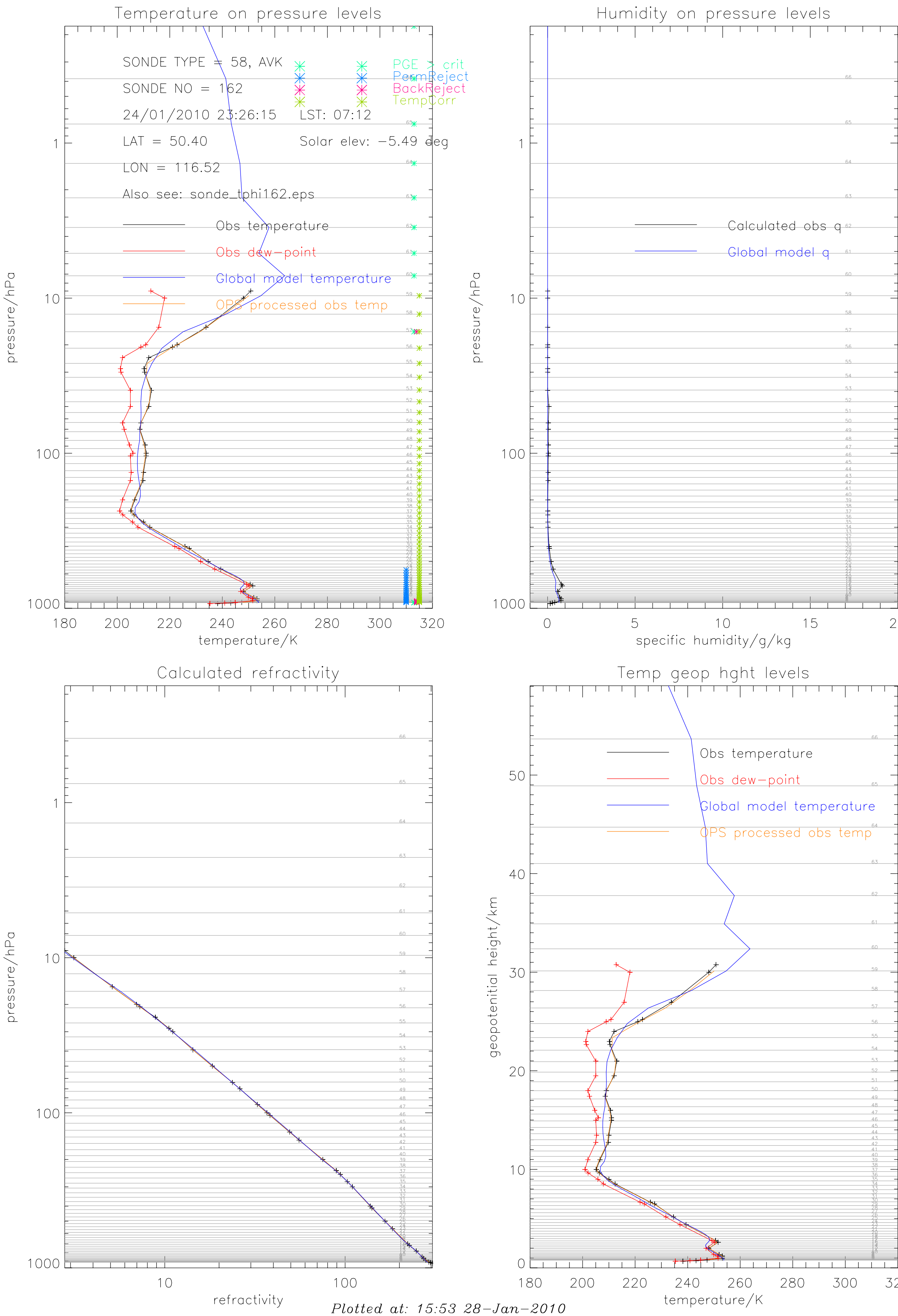


Figure 3: Poor quality temperature measurements from a radiosonde ascent — an AVK over Eastern Russia.

The AVK in Figure 3 does have the Permanent Reject flag set (via the stationlist) from the surface to model level 23 (around 4 km) as indicated by the blue asterisks to the right of the temperature plot, so this data would not be assimilated. However data from 4–30 km would be assimilated, apart from data on model level 57 which has triggered the Background Reject flag get (which is shown by pink coloured asterisks) — this implies the QC for background checks is slightly too lenient. Therefore the bulk of the temperature data will have been assimilated in VAR. Note a temperature correction was applied for this ascent (green asterisks), but the temperature correction did not appear to improve the situation (this was a night-time ascent, hence prone to some infrared cooling — but the measurements actually appear to be too warm in the stratosphere). Generally, the temperature correction for radiation does not seem to improve biased temperatures by any noticeable amount.

The radiosonde in Figure 4 is also an AVK radiosonde. Here the temperature is consistently larger than the background by a significant amount (this behaviour was not seen for other radiosondes in close proximity). The warm bias seems to appear above the temperature inversion at 1 km and remains so up to the top of the ascent at 16 km. This data was not rejected by the QC — again the QC must be too lenient. This radiosonde also has a very low number of significant levels. There is a gap in the significant levels centred around 4 km with seven model levels in between, even though the model background suggests a notable curvature to the temperature lapse-rate.

The temperature observations may not be biased *per se*, rather the corresponding pressure associated with each temperature measurement may be biased, therefore the vertical assignment of each temperature observation will be misplaced. In fact, this is the more likely explanation for the AVK biases rather than a radiation bias, given that the correction of temperature for solar radiation effects tends to provide much smaller changes in temperature than the biases seen and the bias was positive rather than negative at night-time.

Figure 5 shows a MODEM (M2K2-P, France) radiosonde which clearly went wrong between 3–10 km. It should be noted that data from MODEM radiosondes are usually of good quality. The QC is doing a good job in rejecting the data in this region — the PGE value is set greater than critical, so this portion of the profile would not be assimilated. However, data between model levels 43–49 would be assimilated even though the lower part of the profile went catastrophically wrong. We think if such gross errors are found then the whole ascent should be rejected to be safe — the temperatures that would be assimilated look to be negatively biased.

It appears that in general the OPS Background QC should be made stricter, given that ascents apparently with biases of $\sim$5 K are passing through to VAR, especially in the stratosphere. It should be noted that cases where the QC perhaps could do better are not hard to find, any 0Z or 12 Z run will provide several examples. Changes made in parallel suite (PS) 23 have not improved the QC significantly (although the assumed observation errors are more suitable), as is shown by the example in Figure 6 — again an AVK radiosonde which is not rejected by QC, despite a consistent negative bias against the global model with height. The bias appears to start at the surface as shown in the tephigram of the same data: Figure 7. Also notice how the dew-point appears to be positively biased compared to the background.

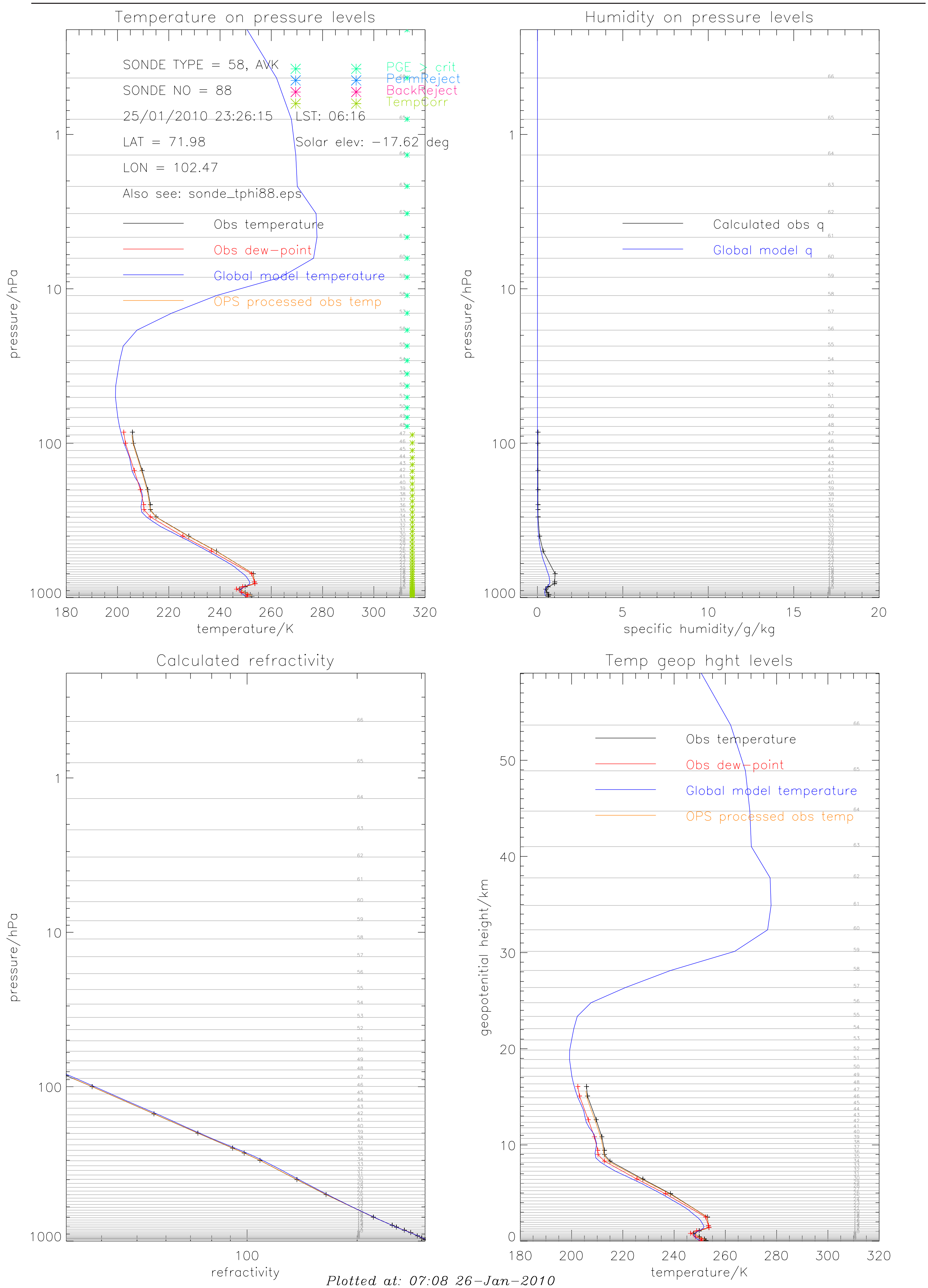


Figure 4: Poor quality temperature measurements from a radiosonde ascent — an AVK over Northern Russia.

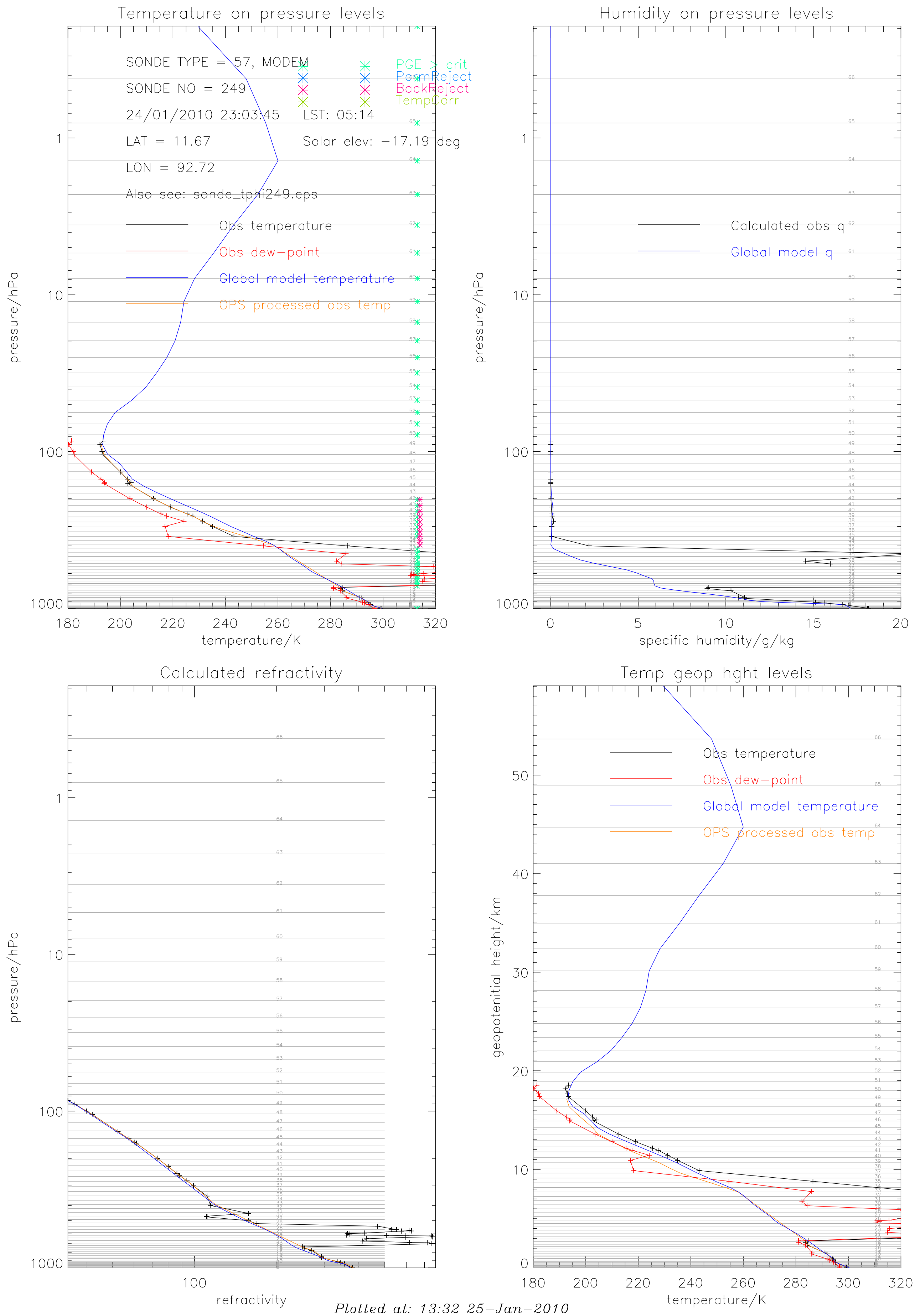


Figure 5: Poor quality temperature measurements from a MODEM radiosonde in the Andaman Islands.

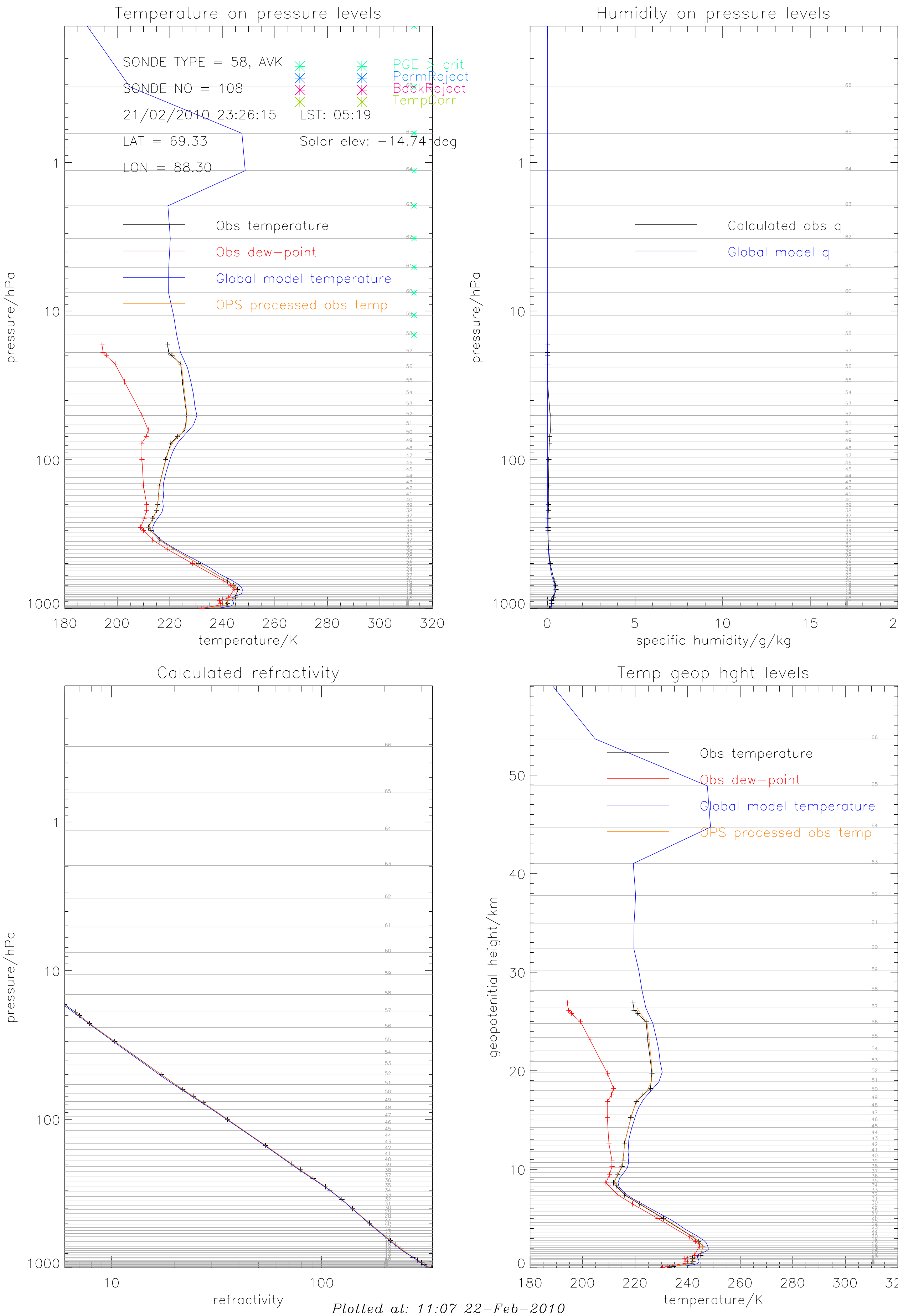


Figure 6: Poor quality temperature measurements from an AVK radiosonde from Northern Russia. Using the PS23 stationlists.

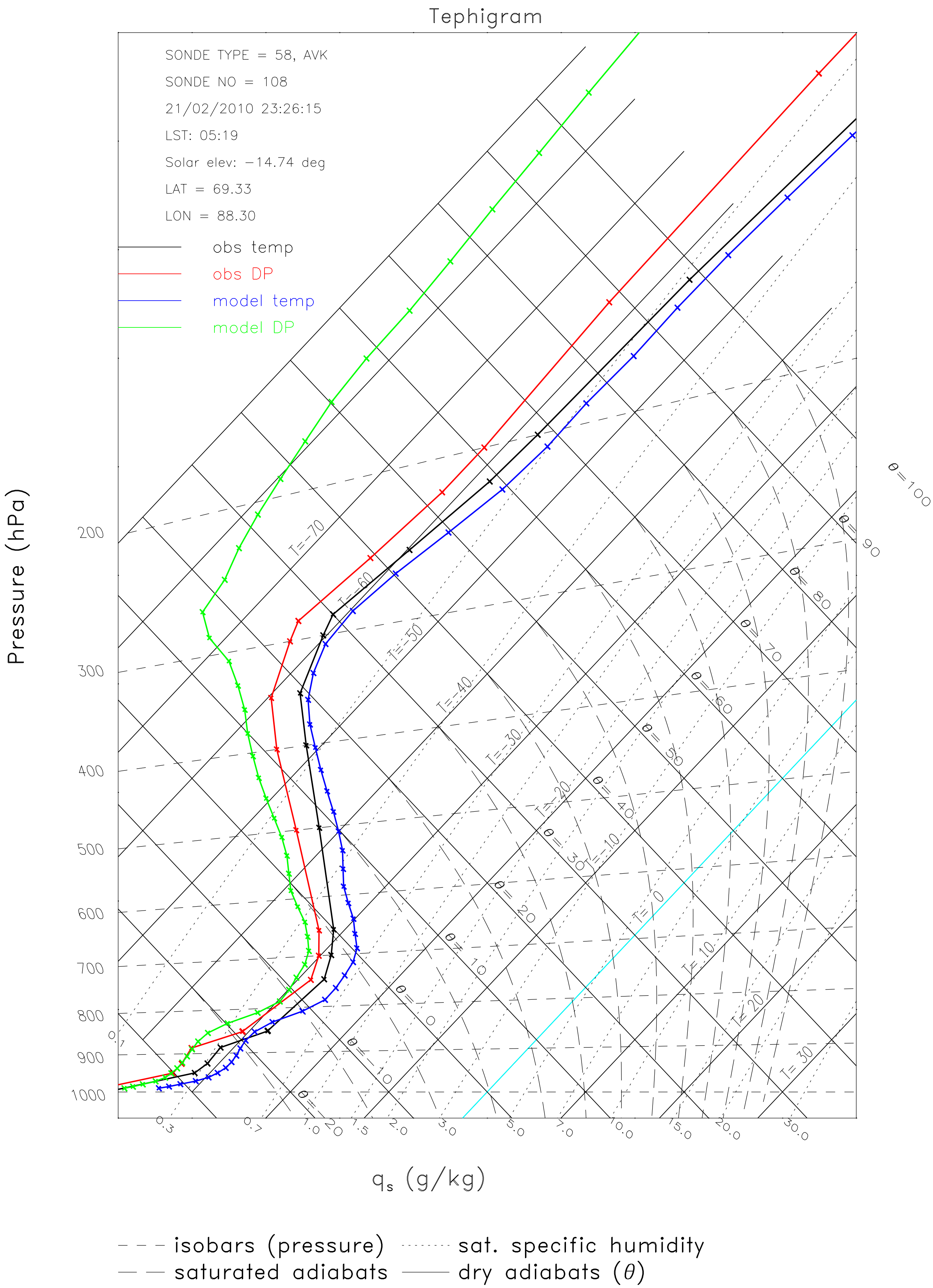


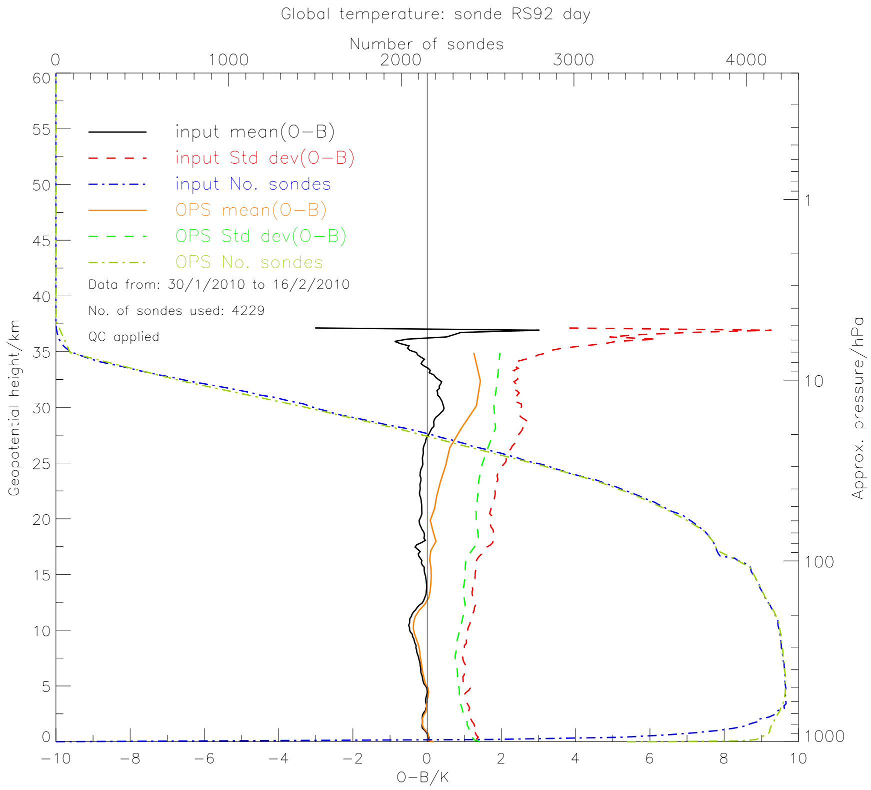
Figure 7: Tephigram of Figure 6 data.

3.2 $O - B$ statistics

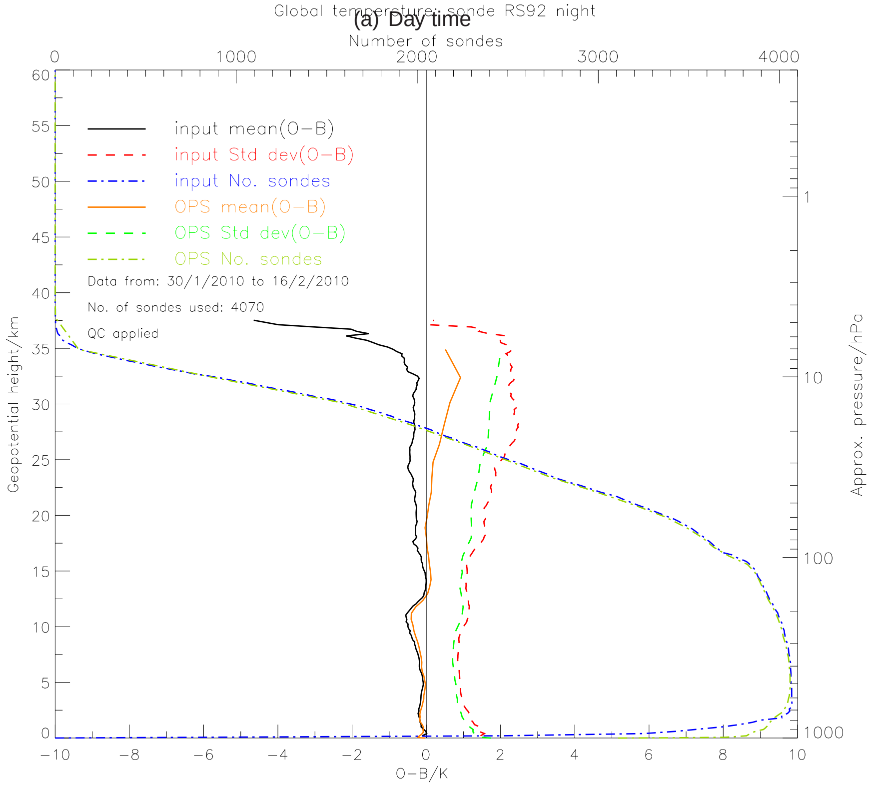
Statistics of observation minus background departures ($O - B$), using two weeks of data (30 January–16 February 2010) can help to identify any issues with a particular type of radiosonde. One should note that the biases and standard deviations for radiosondes vary to some extent from month to month, reflecting the meteorological conditions at the radiosonde site — so these plots typify the February 2010 period.

Figure 8 shows the temperature $O - B$ statistics (in Kelvin) for RS92 radiosondes for day-time and night-time. The Figure plots both the radiosonde data as interpolated with geopotential height from significant levels to a 200 m grid (using IDL INTERPOL) and the OPS processed temperature data (on θ model level heights). One can see in that below 15 km there is a high consistency in the vertical temperature bias structure for day and night. Above 15 km there are differences between day and night-time in the mean, presumably due to the effect of solar radiation not being completely corrected. Note that a temperature radiation correction is not applied to RS92 within the OPS, a correction is apparently made at source.

The effect of the OPS processing of temperature onto model levels (orange line is the bias) appears to increase the temperature of the observation data by up to $\sim$1 K above 15 km for all times relative to the unprocessed data, an effect which must be related to subtle differences in the assumptions made for the vertical interpolation/processing methods employed for temperature — this becomes more evident when the model level vertical spacing increases (removing the QC applied to the input temperature data did not alter the mean). Rick Rawlins (per. comms.) has previously found biases between different methods for the conversion of θ level potential temperature to temperature — the biases depend on the method chosen for interpolating pressure from ρ to θ levels. The methods for processing/interpolation are discussed in section 4.



(a) Day time



(b) Night time

Figure 8: Temperature $O - B$ statistics for Vaisala RS92 radiosondes

Figure 9 is equivalent to Figure 8, but for AVK (coverage over Russia) radiosondes. We can see that the AVK have a larger day/night difference than RS92. Note a large fraction of AVK are rejected by the OPS QC, especially in day time between 5–10 km. AVK radiosondes temperatures are bias-corrected in the OPS, which might explain the departure of the OPS processed bias line (orange) to more negative values below 25 km relative to the significant level data (black line), whereas it was to more positive values for RS92.

Figure 10 shows the same type of plot, but for the French MODEM radiosondes. There is a good consistency in the bias between day and night-time. As for RS92, the OPS processed temperature bias is shifted more positively relative to the input data by up to 1 K above 15 km.

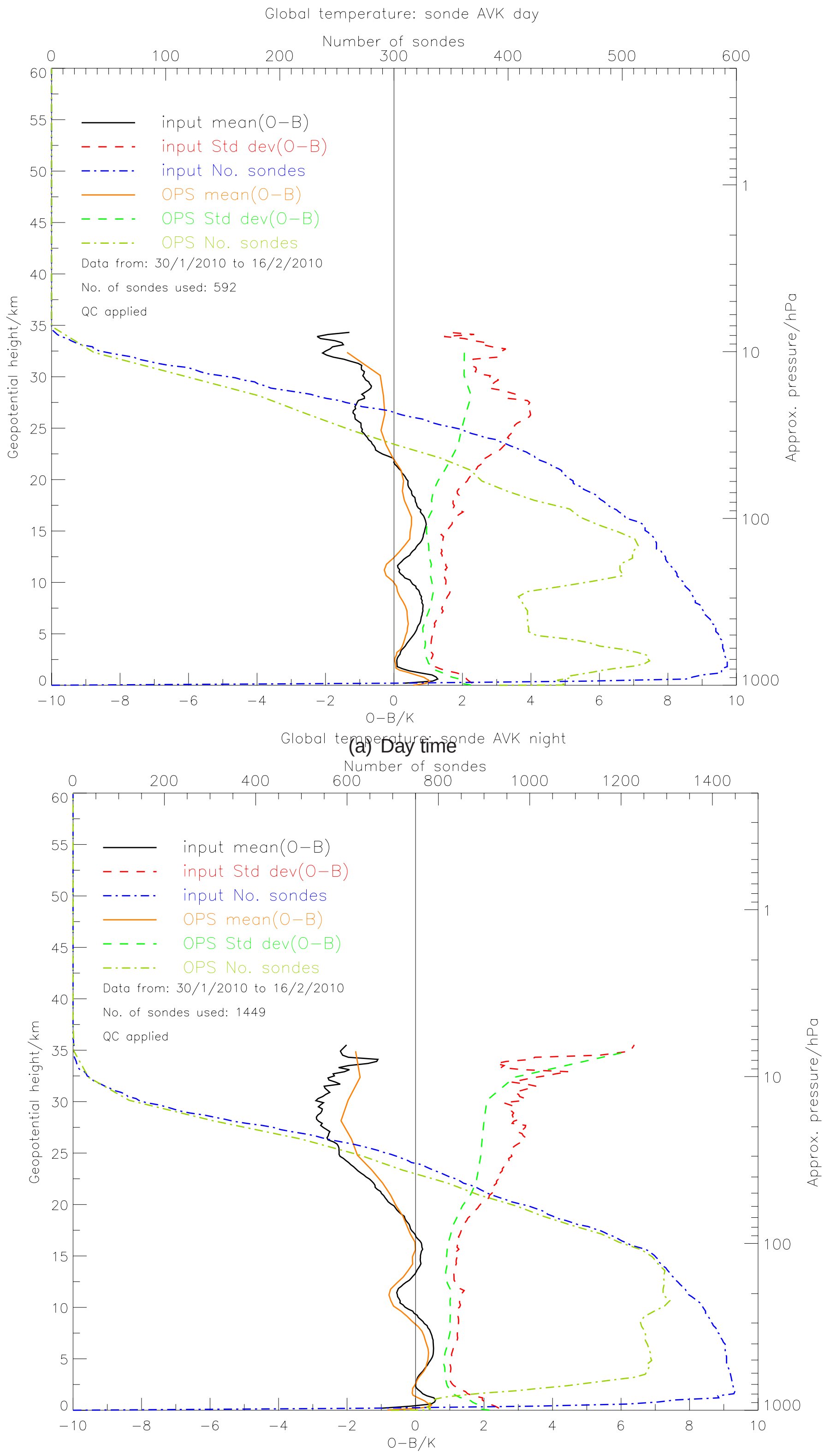
Figure 11 is a similar plot to Figure 8, but for refractivity $(O-B)/B$ rather than temperature $O-B$ (refractivity being an exponentially decaying function benefits from being normalised by the B). A similar result to the temperature statistics is seen regarding the similarity of the day/night statistics for the majority of the profile. However, there is a fluctuation in low level (0–5 km) refractivity bias between day/night, which is due to a varying day/night biases in radiosonde humidity. It appears the radiosondes are biased dry during the day below 5 km by up to $\sim$0.25 g/kg (figure not shown). Assimilating biased radiosonde humidity measurements is believed (at the time of writing) to be causing UM dry bias.

There is tendency for more negative biases in the refractivity derived from the OPS processed temperature measurements in the stratosphere compared to the input data — positive temperature bias would lead to negative refractivity bias, due to the inverse relationship. The lower level more positive biases for the processed data must be related to humidity, which we have not investigated rigorously.

Figure 12 shows the AVK refractivity $(O-B)/B$ plots; it seems that an accumulation of the temperature bias leads to an increasing geopotential height bias higher up the profile, particularly in day-time for the input data. The accumulation of bias in geopotential height calculations is exaggerated in refractivity space because refractivity decreases exponentially with height, thus e.g. a 25 m bias in height leads to a $\sim$0.35% bias in refractivity assuming a 7 km scale-height, whereas temperature increases/decreases at a smaller rate with height. There is not a varying day/night-time bias in the lower troposphere for AVK, probably because humidity is very low over Russia in January, therefore it has a negligible effect on refractivity.

Figure 13 is the same type of plot, but for French MODEM radiosondes. These radiosondes look well behaved in refractivity space above 10 km; below 10 km the day/night bias difference due the humidity part of refractivity is quite significant. Looking at $O-B$ statistics for specific humidity (figure not shown) for day/night-time the diurnal range of the mean is $\sim 0.7\,\mathrm{g/kg}$ at 2 km.

From Figure 14 we can see that the FORMOSAT-3/COSMIC (FORMOSAT-3 - Taiwan's Formosa Satellite Mission #3 COSMIC - Constellation Observing System for Meteorology, Ionosphere & Climate, data provided by UCAR, hereafter referred to as COSMIC) GPSRO stratospheric bias against the global model is most similar to the RS92 refractivity $(O-B)/B$ bias (note the negative bias at

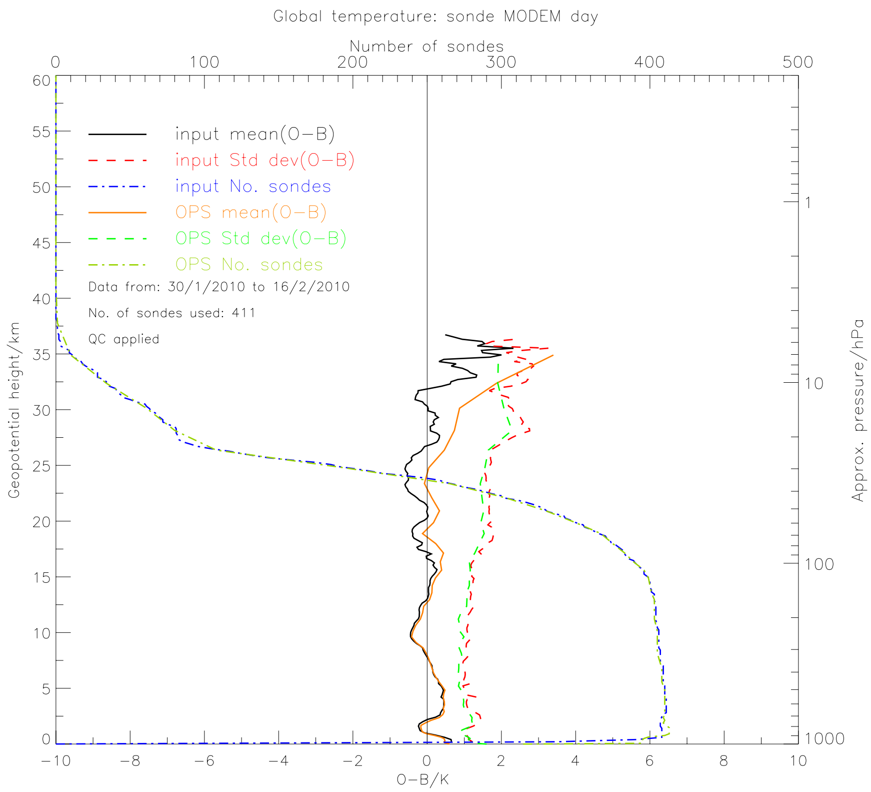


(a) Day time

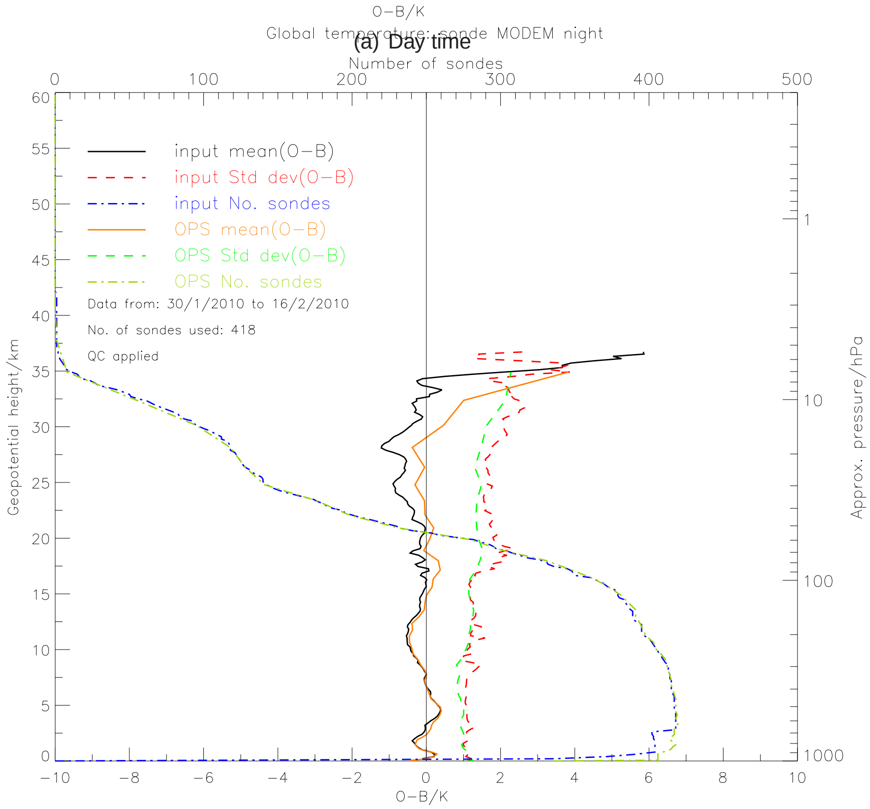
(b) Night time

Figure 9: Temperature $O - B$ statistics for Russian AVK radiosondes

∼30 km), suggesting RS92 have the best stratospheric biases of those shown, since GPSRO data is known to have relatively small biases up to 40 km.

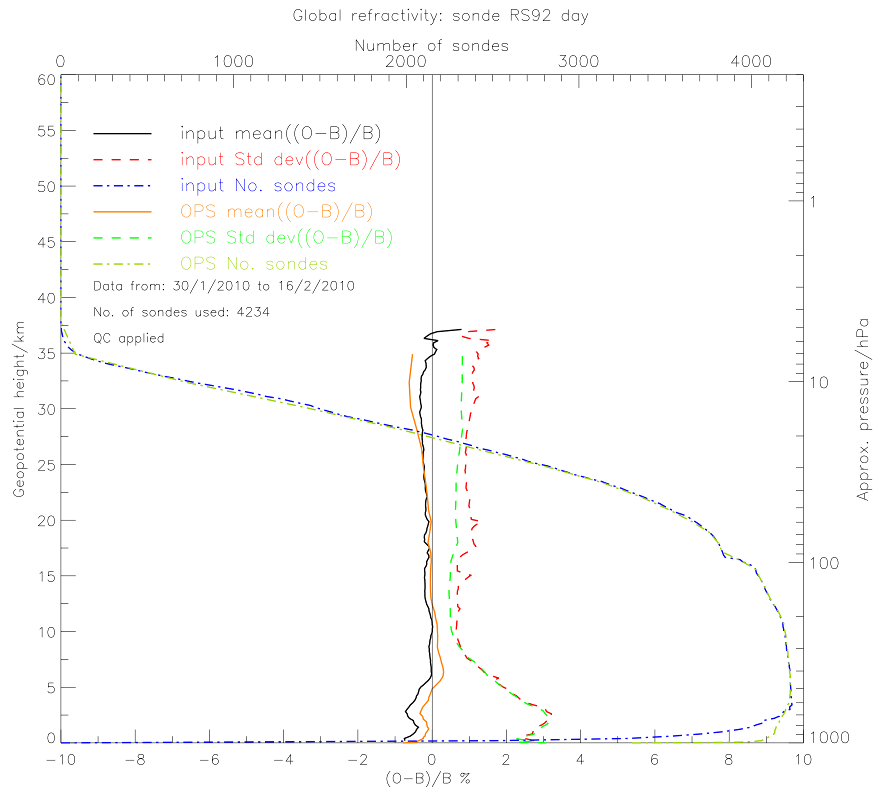


(a) Day time

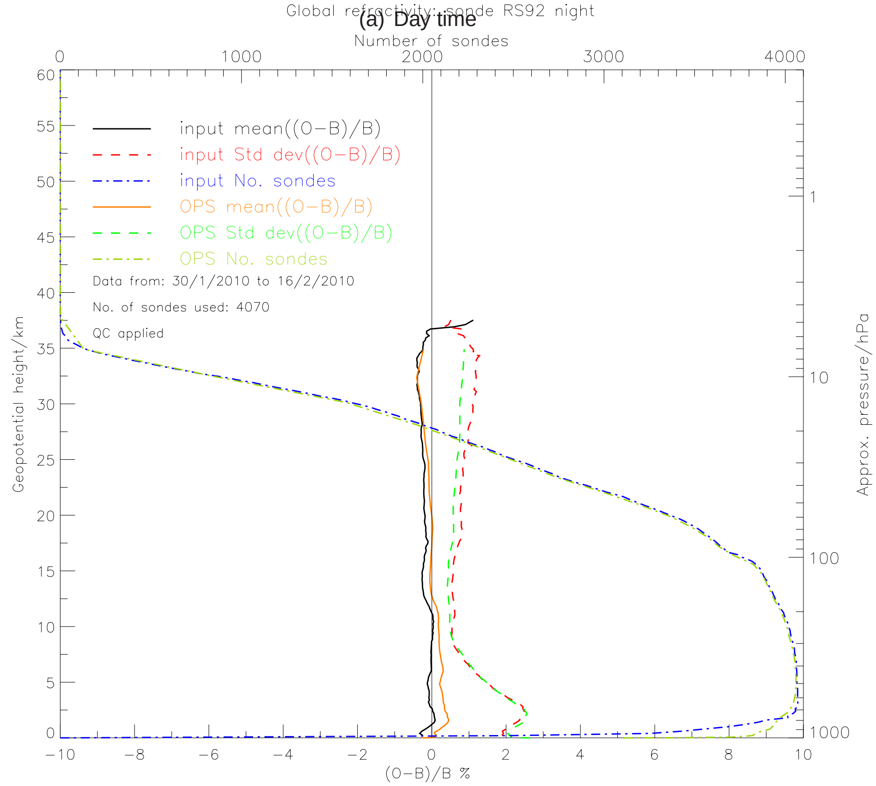


(b) Night time

Figure 10: Temperature $O - B$ statistics for French Modem radiosondes

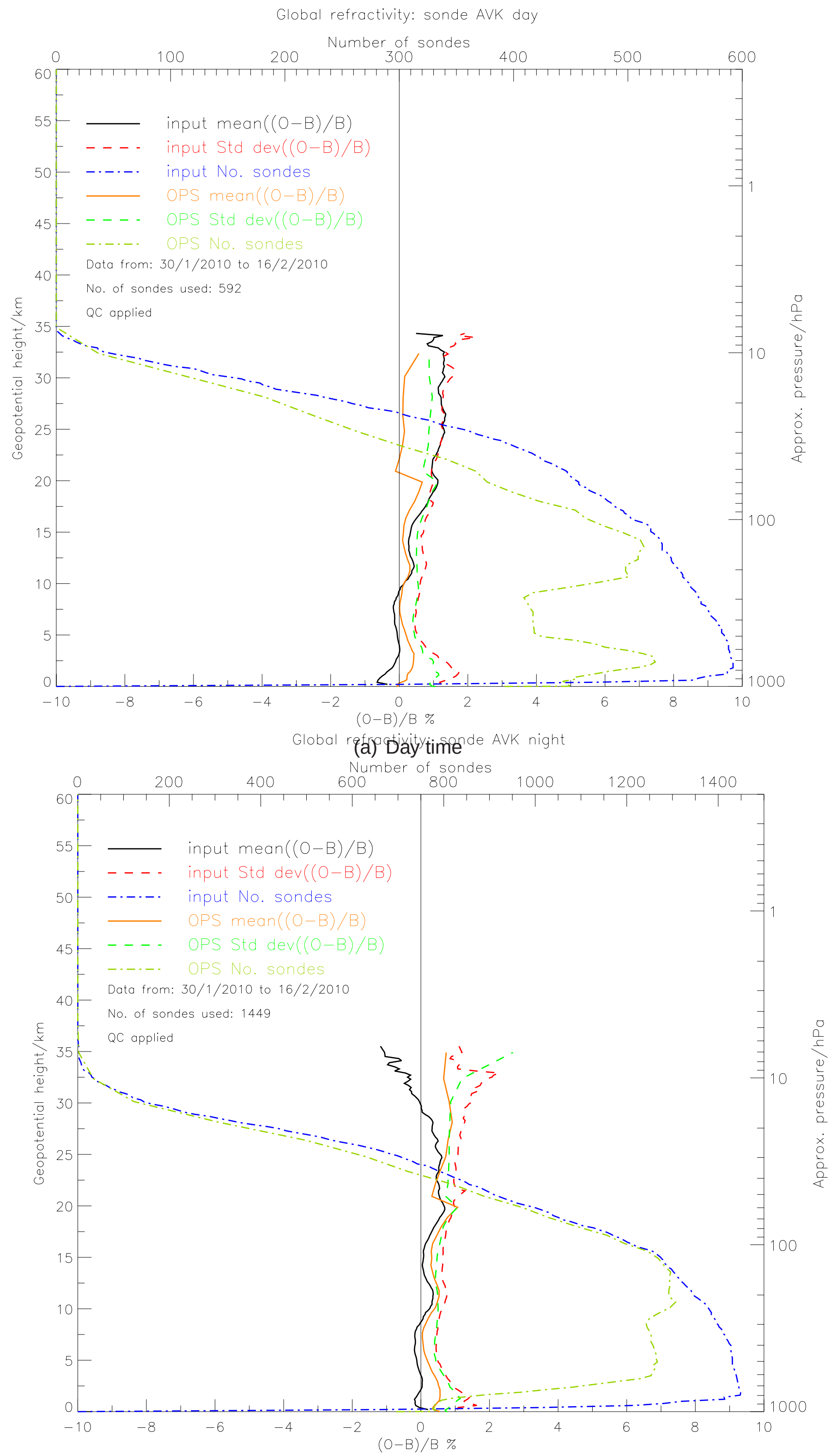


(a) Day time



(b) Night time

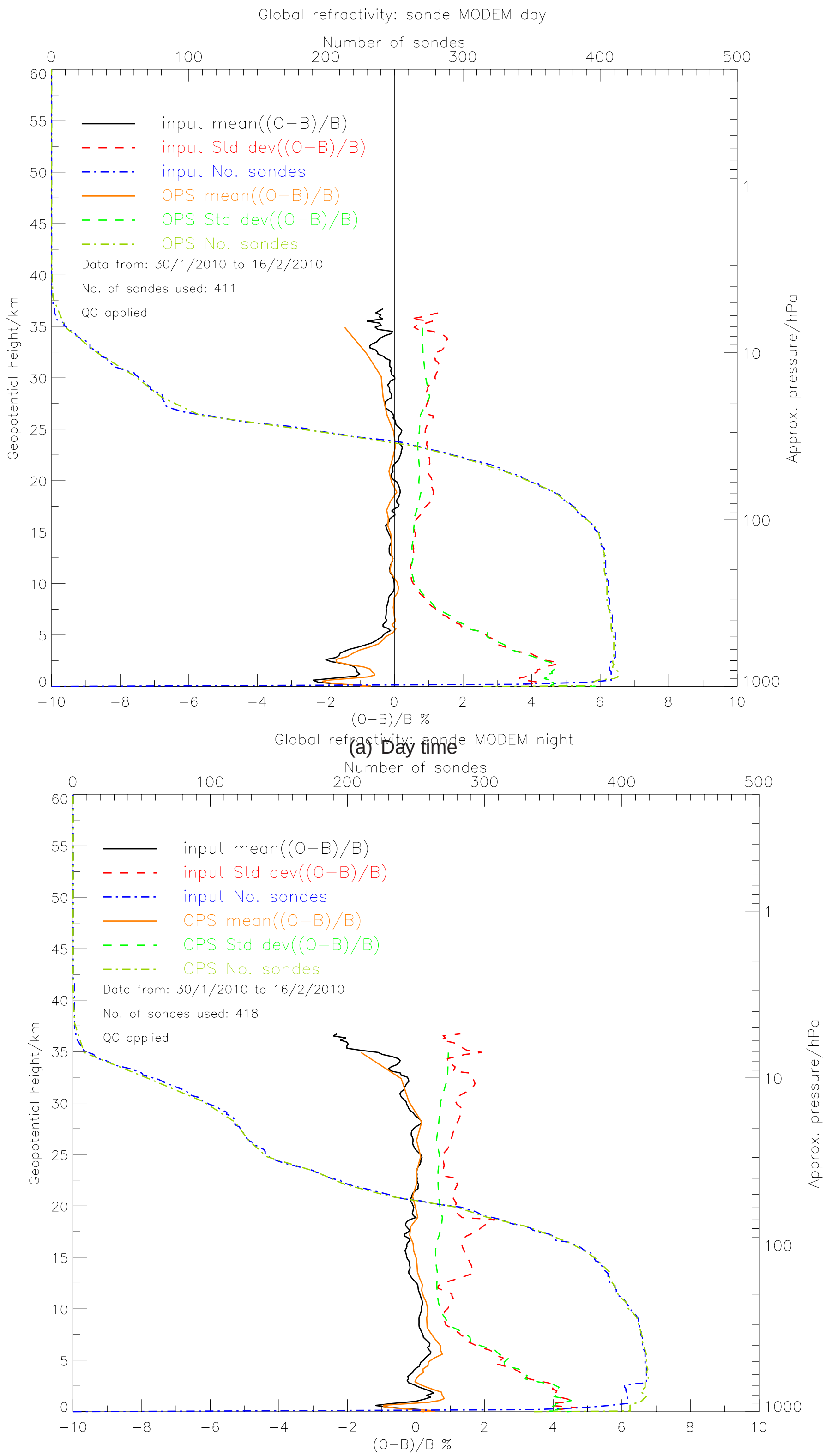
Figure 11: Refractivity $(O - B)/B$ statistics for Vaisala RS92 radiosondes



(a) Day time

(b) Night time

Figure 12: Refractivity (O-B)/B statistics for Russian AVK radiosondes



(a) Day time

(b) Night time

Figure 13: Refractivity (O-B)/B statistics for French Modem radiosondes

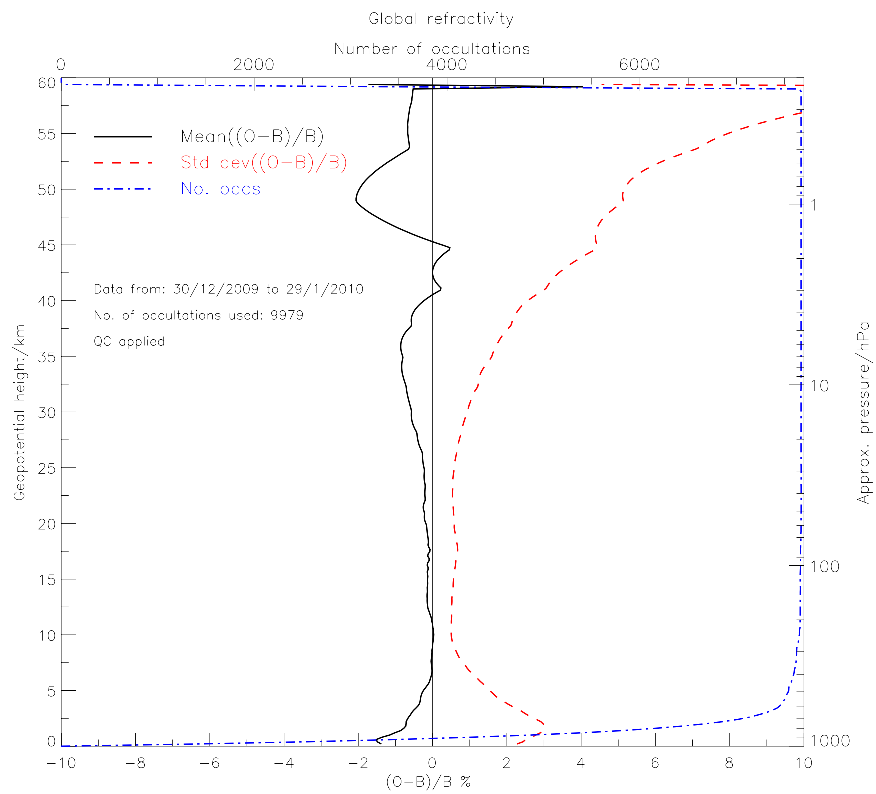


Figure 14: Global refractivity (O-B)/B for an example GPSRO instrument: COSMIC 4 (UCAR)

3.3 Coincident GPSRO-Radiosonde Statistics

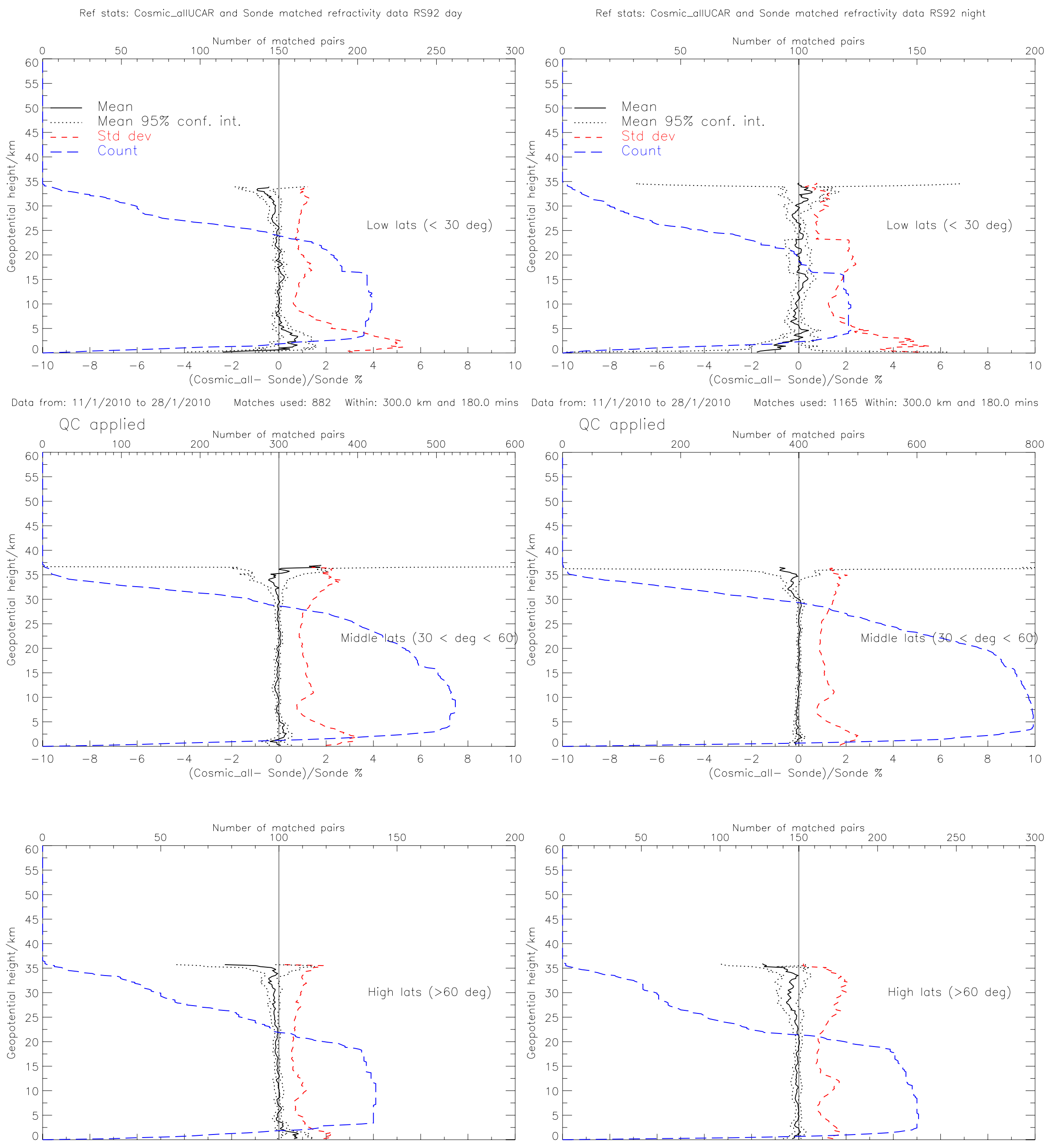
Comparisons of coincident radiosonde and GPSRO refractivity profiles are useful for determining the relative biases between the datasets. For the coincident data statistics, the chosen criteria for a match are that the separation in great-circle distance and time be less than 300 km and 3 hours respectively; this seemed a reasonable compromise, whereby the number of matches is sufficient for analysis and the difference in meteorological state is small. The plots use data covering the 11–28 January 2010 period. Note that these plots used the radiosonde data prior to any OPS processing — only IDL mapping of radiosonde data to refractivity space.

Figure 15 shows statistics of refractivity $(GPSRO - Radiosonde)/Radiosonde$ for coincident COSMIC and RS92, split by latitude bands and into day-time and night-time. The most obvious feature from the plots is that the biases between RS92 and COSMIC are very small (black mean line surrounded by 95% confidence interval) — typically less than 0.3% between 5–30 km. Also, there are negligible changes in bias between day and night. Such small biases imply the method for calculating geopotential height from the radiosonde observations and the log-linear vertical interpolation of refractivity were done correctly.

Figure 16 is the same type of plot as Figure 15, but with GRAS data (The GNSS Receiver for Atmospheric Sounding, on-board MetOp, data provided by the GRAS SAF) instead of COSMIC. There are fewer GRAS-RS92 matches than COSMIC-RS92, therefore the mean differences have larger confidence intervals. For this period GRAS data is known to be negatively biased below $\sim$8 km, and should be ignored here (due to geometrical optics rather than wave optics processing). Between 10–30 km in mid-latitudes the GRAS data have good consistency with RS92 in terms of bias. Above 30 km there is a small positive bias, which we attribute to GRAS refractivity data since a similar feature is seen when we compare COSMIC and GRAS data to the model — there are still improvements to be made in the GRAS statistical optimisation at these altitudes (applied operationally in February 2010).

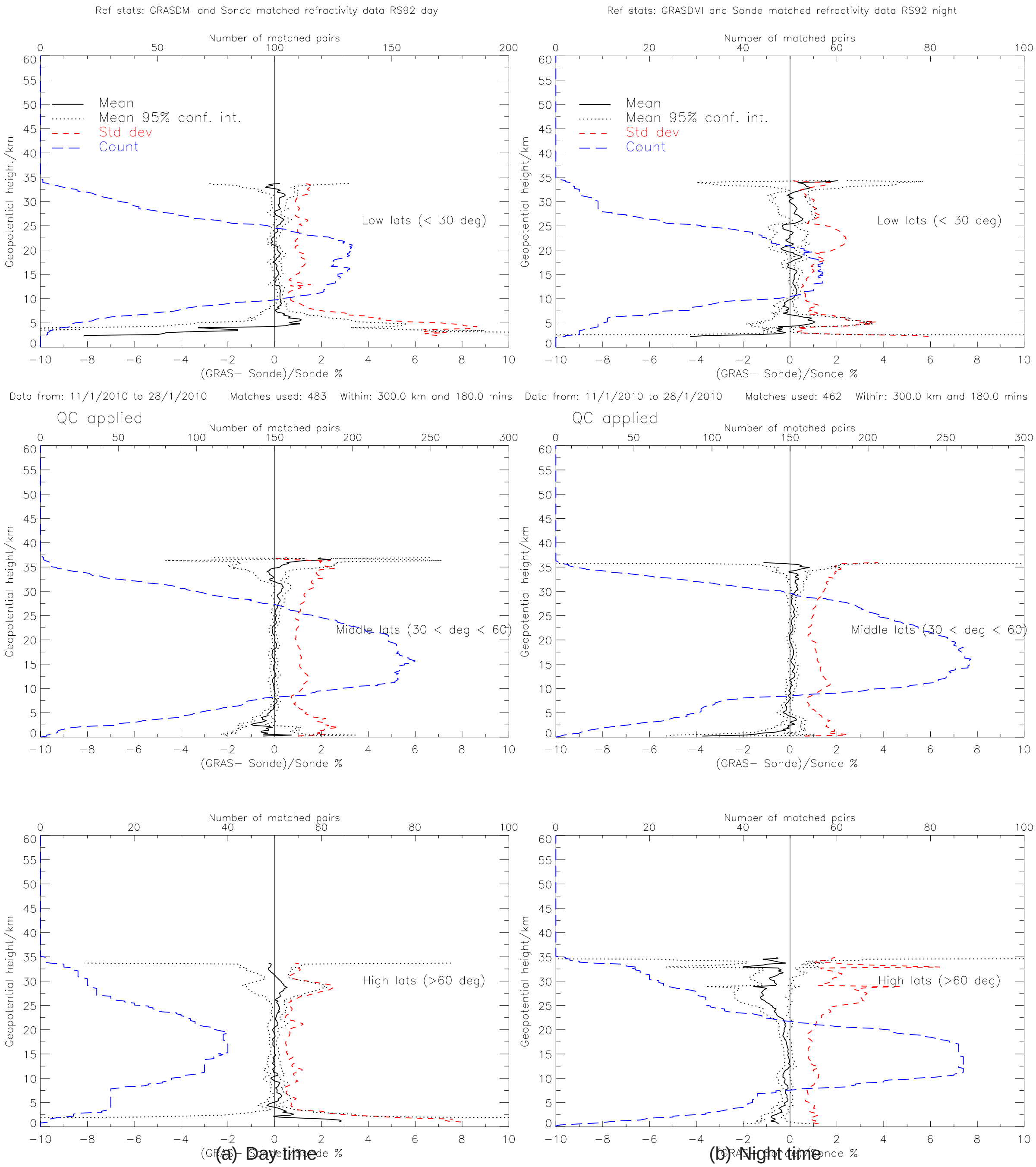
Figure 17 compares COSMIC to AVK radiosondes. There are no AVK data available at low latitudes, since they are nearly all deployed in Russia. It is clear that AVK refractivity values are too large (particularly in day-time), leading to negative bias relative to COSMIC by over 2% in places — a similar picture is seen against GRAS (not shown). This bias indicates that if the AVK temperature is assimilated then it will be ’fighting’ against GPSRO to some extent. However these plots are based on the unprocessed data, prior to OPS processing and QC, so many of the outliers that have larger bias may have been rejected. Similar biases to those in Figure 17 would be generated by AVK ascents biased as in Figure 3.

Finally, Figure 18 compares MODEM radiosondes to COSMIC. The lack of matches makes the statistics rather noisy and dominated by the occasional outlier. Even so, the statistics look reasonable, particularly at low latitudes, apart from the day/night bias variation in lowest few kilometres related to humidity.



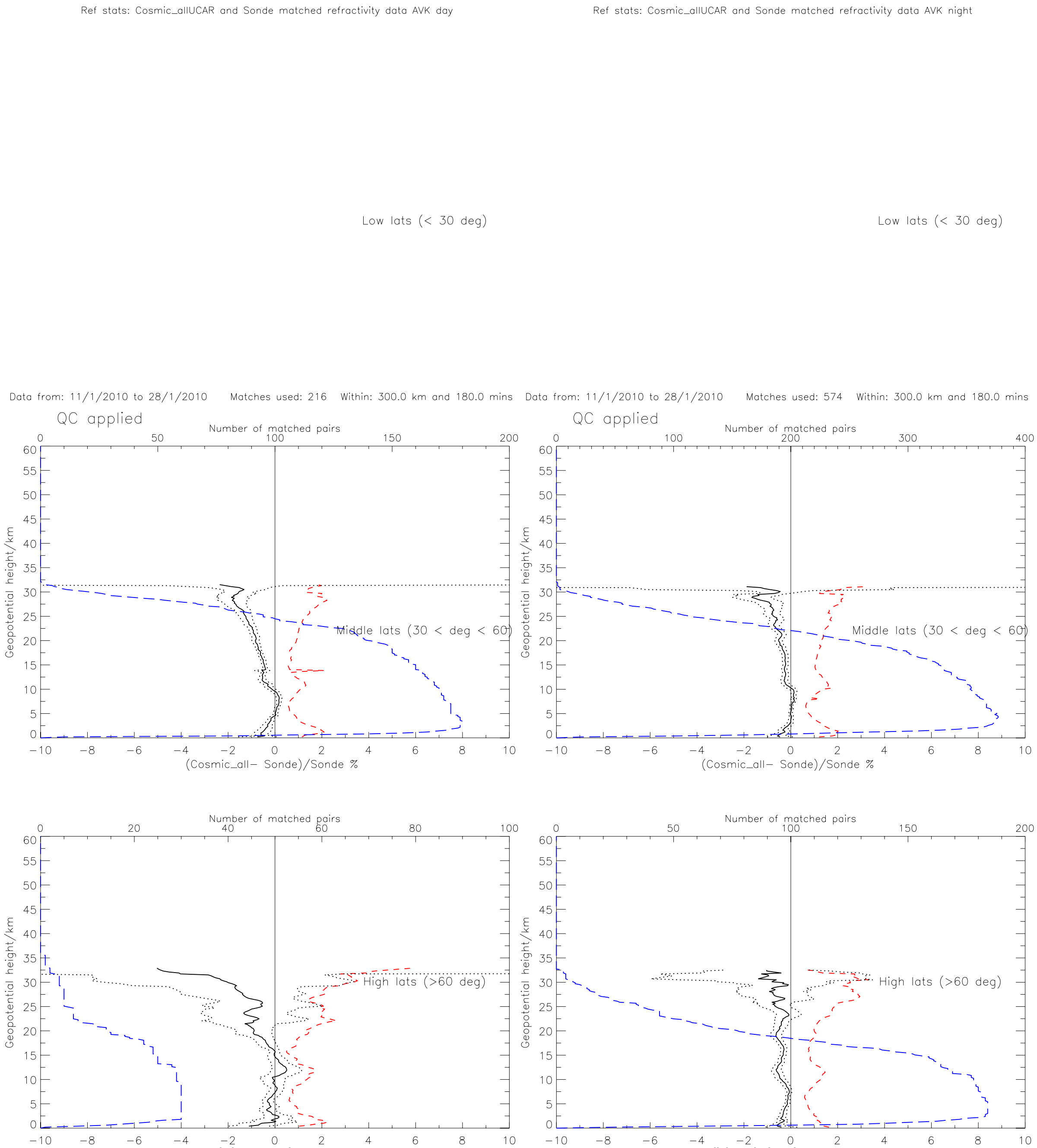
(a) Day time (b) Night time

Figure 15: Coincident GPSRO and radiosonde refractivity statistics, split into low, middle and high latitudes. Using RS92 radiosonde data and COSMIC data.



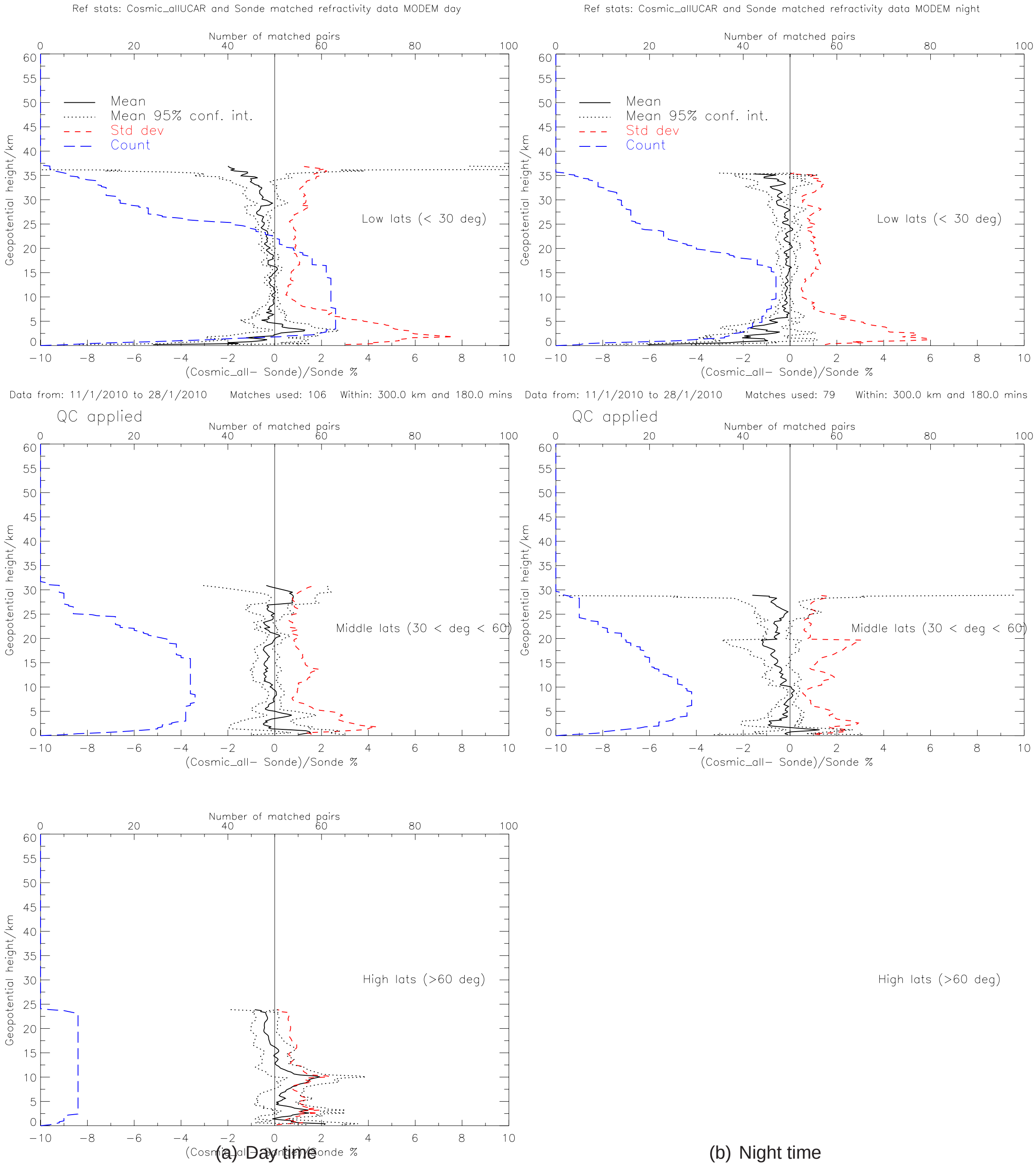
(a) Day time (b) Night time

Figure 16: Format same as Figure 15, but using GRAS data instead than COSMIC.



(a) Day time (b) Night time

Figure 17: Format same as Figure 15, but using AVK radiosonde data instead of RS92.



(a) Day time (b) Night time

Figure 18: Format same as Figure 15, but using Modem radiosonde data instead of RS92.

4 Discussion on interpolation/processing methods

The GPSRO forward model (see OSDP18 (2009)) and the OPS processing of radiosonde observation temperature onto model levels have some consistency in the sense that they both calculate a layer-mean virtual potential temperature using a formula of the form:

$$\overline{\theta_v} = \frac{g}{c_p}\frac{\Delta Z}{\Delta \Pi} \tag{6}$$

where $\Delta\Pi$ is the difference in Exner between adjacent ρ levels and ΔZ is the difference in geopotential height between adjacent ρ levels. For the GPSRO forward model ΔZ is calculated from the difference in geopotential heights of the surrounding ρ levels, whereas for the radiosonde processing it is given from the summation of thicknesses (calculated using the hypsometric equation) with the observation temperatures and pressures, and the temperatures linearly interpolated in log-pressure to the surrounding ρ levels. $\Delta\Pi$ is the same in both cases i.e. calculated from the ρ level pressures.

In section 3.2 we discovered a stratospheric bias of order 1 K in the OPS processing of radiosonde temperatures onto θ model levels. In this section we investigate how this bias is related to the difference between the layer-mean and the spot-value interpretation of temperature — it is unclear how θ values on model θ levels should be interpreted.

The plot in Figure 19 shows the differences between the layer-mean forward modelled temperatures and the θ level model temperatures ($T_\theta = \theta\Pi_\theta$), where θ is provided from the UM background forecast. The layer-mean temperature is calculated in the GPSRO forward model (for a globally distributed sample of profiles), using equation 7.

$$\overline{T} = \frac{\overline{\theta_v}\Pi_\theta}{1 + q/(\epsilon - 1)} \tag{7}$$

Π_θ is Exner on the θ level (determined by linear interpolation of ρ level Exner with geopotential height) and q is the θ level specific humidity. In Figure 19, there are differences between the methods, however the biases are much less than 1 K at 30–40 km. So there is reasonable consistency between the layer-mean and the θ level model spot values.

A similar plot using potential temperature rather than temperature showed the same bias structure and magnitude (given the scaling by Exner), so the issue is not related to the derived Exner on θ levels — besides, the same factor multiplies the layer-mean and the spot-value potential temperature. There are large differences however for the top model layer (up to $\pm$20 K) due to the model boundary conditions: $\Delta\Pi$ for the top ρ levels is set to a fixed value. Also, there are distinct positive differences on the plot for some locations over high orography (as seen below 10 km), which we can not explain. Hence, there does not appear to be a need to update the GPSRO forward model to use the background θ on θ levels, rather than the current layer-mean temperature calculation.

Another potential source of bias is from the IDL processing methods applied to the radiosonde

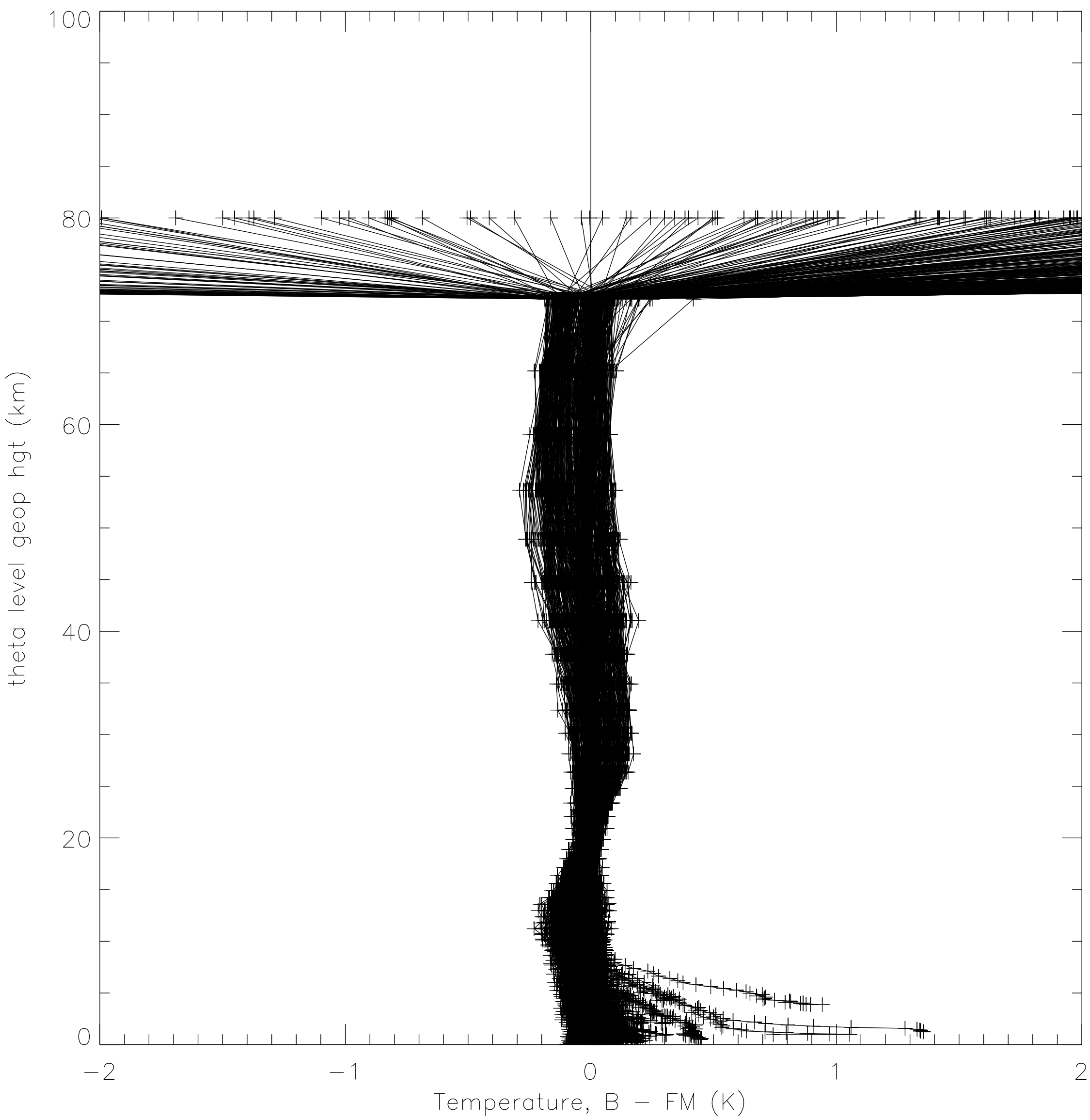


Figure 19: Plot of around 500 profiles of differences between model background temperature and layer-mean forward modelled temperature.

observation data e.g. the 1 K bias could a genuine bias between model and observations, and some artefact of the IDL processing erroneously reduces the bias. However, one should point out that without any interpolation, and judging the profiles subjectively e.g. in Figure 2, there does appear to be a positive bias in the stratosphere for the OPS processed observation temperatures relative to the data received on significant levels of the right magnitude.

Rather than interpolating onto a 200 m spaced vertical grid in IDL, we tried a vertical grid that is equal to the geopotential heights of the L70 model over zero orography. This plot showed very similar bias to the 200 m spaced grid (plot not shown), so the vertical grid is not an issue.

In the IDL code we have been interpolating the temperature linearly with geopotential height. However, the first part of the OPS processing interpolates radiosonde temperatures linearly with log-pressure onto ρ levels (prior to the thickness calculation), so we have tried a similar procedure in IDL i.e. interpolate the radiosonde temperatures linearly with log-pressure to the θ level pressures, for comparison with θ level model temperature. This method also did not alter the bias (plot not shown), hence the vertical co-ordinate for the IDL interpolation is not an issue. Finally, to test the sensitivity of the IDL interpolation we tried changing from linear to least-squares quadratic fit interpolation — this also made no discernible difference.

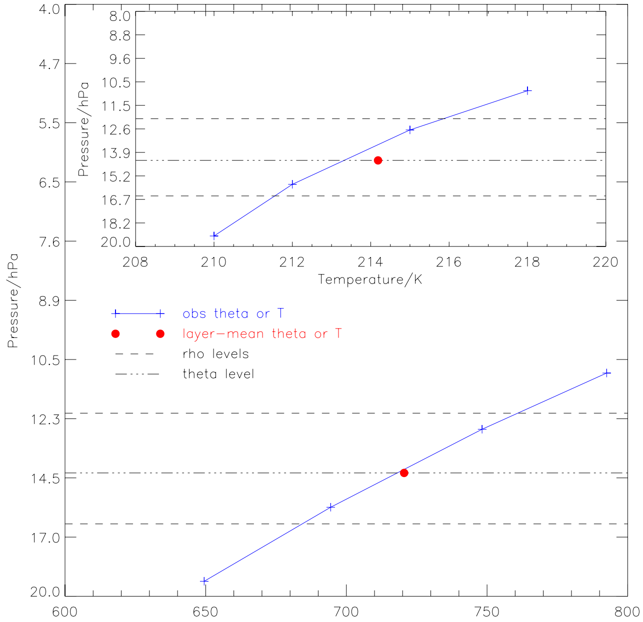
The next test was to create a simulation in IDL of the OPS radiosonde processing using the formulae of Swinbank and Wilson (1990), in particular equations 4.4 and 4.5. We chose some realistic radiosonde temperatures and a set-up of the observation levels relative to the real model levels similar to Figure 4 of the Swinbank and Wilson (1990) paper example. The data is in the stratosphere at around 28 km, a vertical location at which the positive biases are observed.

We then calculated the layer-mean potential temperature (using the Swinbank and Wilson (1990) paper formulae). However we altered the pressure at which we apply the layer-mean: at the θ level pressure (Figure 20(a)) and at half-way between the ρ level pressures in log-pressure space (Figure 20(b)). The linear interpolation with log-pressure (blue line) agrees much better with the half-way pressure level vertical placement than with the θ level pressure vertical placement.

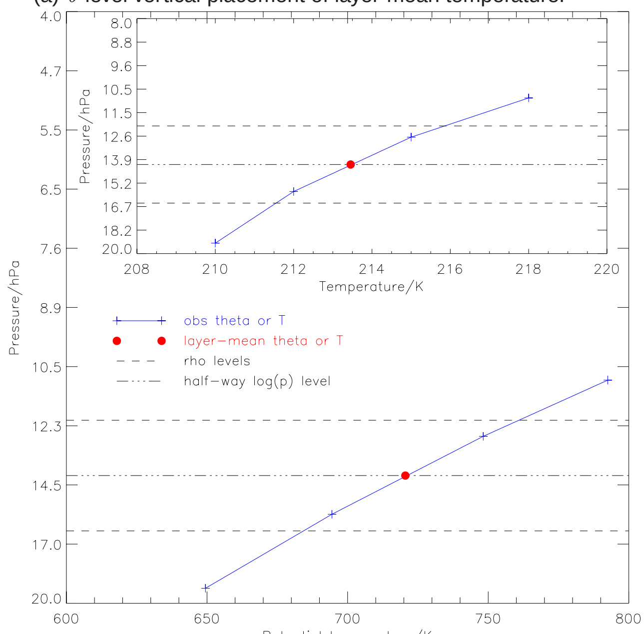
The effect comes about because the θ level is not half-way between the ρ levels in geopotential height nor in log-pressure, but is in fact slightly lower than half-way. The New Dynamics UM documentation (section 4) states: "In general the half-integral meshpoints are not equidistant from the two neighbouring integral meshpoints and neither are the integral meshpoints equidistant from their neighbouring half-integral meshpoints". In the θ level pressure case, the layer-mean potential temperature of 720.538 K is applied at a higher pressure (lower geopotential height) than the half-way pressure case, which leads to a positive bias. This bias looks relatively larger in temperature space than potential temperature space given the scaling by Exner, creating a bias positive bias of around 1 K, whereas the half-way method lies almost exactly on the blue line.

This type of bias is exacerbated by larger temperature lapse rates and larger ρ model level spacings — hence the model stratosphere has the largest bias (note the L70 model has poorer vertical resolution in the stratosphere than the L50 did). Referring to the Swinbank and Wilson (1990) paper, the layer-mean method was developed when the model had hybrid pressure levels which had half-levels that where exactly half-way between the surrounding levels in geopotential height space (log-pressure space). The introduction of New Dynamics changed the model layout, with θ levels not being exactly half-way between ρ levels.

The negative impact of assimilating positively biased radiosonde potential temperature values in the stratosphere of several Kelvins in potential temperature space is yet to be understood. Also, the reason why the forward modelled layer-mean temperatures (as in GPSRO forward model) do not show biases against the model θ level temperatures of the same magnitude as the radiosonde OPS processing biases is not understood — perhaps consistency is enforced through the UM dynamics. A potential solution to avoid the bias would be to just assimilate temperature data that has been linearly interpolated with log-pressure to the θ levels, which would be in agreement with the definition of significant levels.



(a) θ level vertical placement of layer-mean temperature.



(b) Half-way in log-pressure vertical placement of layer-mean temperature.

Figure 20: Assessing the fit of the layer-mean value to the linear interpolation with log-pressure method, via differing vertical positioning of the layer-mean.

5 Conclusions

From the analysis of individual radiosonde profiles by comparison to the global model background forecasts, the majority of profiles show plausible, well behaved temperature and dew-point temperature data, as we would expect. Subjectively, Vaisala RS92 and Sippican radiosondes show the best quality data. However, occasionally data showing signs of systematic errors are apparent — most commonly with radiosondes launched over Russia, India (although now mostly switched to MODEM radiosondes) and to some extent the Far East. Erroneous data regularly fails to be rejected by the

OPS QC, as shown by the examples of section 3.1, implying there is a need to tighten up the QC, particularly in checks comparing the observation to background in the stratosphere. The biased (as provided, before OPS processing) radiosonde data which is assimilated may well be limiting GPSRO's ability to control stratospheric biases (and also through feedback via the verification methods for NWP trials).

Compiling radiosonde $O-B$ statistics and comparisons against coincident GPSRO data, split by radiosonde type, helped in understanding the differing biases of radiosonde models. Generally Vaisala RS92 and Sippican have the smallest biases when compared to the global model and to coincident GPSRO profiles (shown in refractivity space) — RS92 has biases against COSMIC GPSRO refractivity data less than 0.3% in the 5–30 km vertical range in mid-latitudes. The Russian AVK radiosondes show significant diurnal changes in bias with mean refractivity differences of 1–2% in the day-time stratosphere when compared with GPSRO.

There is currently work underway in DAE to improve the bias correction for radiosonde data, which may mitigate some of the biases seen — however the biased cases often appear to be gross errors probably unrelated to radiation temperature corrections, since radiosondes of the same type in close proximity often do not suffer the same bias. Also, improved estimates of radiosonde observation errors are being tested in PS23 at the time of writing — these give AVK radiosondes a higher observation error than previously used. The methods outlined in this report could be used to verify the proposed future DAE changes.

Radiosonde temperature $O-B$ statistics showed that the OPS processed temperature data is positively biased (by up to 1 K) compared to the unprocessed data higher than 15 km (where model level vertical spacing becomes significantly larger) — note this is for radiosondes which are not bias corrected within the OPS processing. This bias is believed to be due to where the layer-mean processed temperature (potential temperature) is applied in the vertical. The layer-means are currently applied at the θ level heights/pressures, which are not at the centre of the ρ level layers in terms of geopotential height or log-pressure — they are slightly lower. This vertical misplacement leads to positive biases in the stratosphere, where temperature usually increases with altitude and the model level spacing is sufficiently large. The magnitude of the bias is not very large, however if it is present for all stratospheric radiosonde data then it could well be fighting against the GPSRO observations.

The OPS processing of the data onto θ levels also showed some unsatisfactory behaviour when the model is at a higher vertical resolution than the provided radiosonde significant level data. In such scenarios increasing the model vertical resolution would create more 'model obs' between the observation levels, and in effect give more weight to the radiosonde profile. This may well be reasonable if the significant levels are realistic, but it seems some radiosonde types have intrinsically lower vertical resolution than others (e.g. AVK compared to RS92) or have different accuracy criteria for deriving significant levels. Hence giving all radiosonde observations the same observation errors (as done prior to PS23) is probably over-weighting the low vertical resolution radiosondes. This could be addressed by interpolating the model fields to the observation location in the vertical, as is done in the GPSRO forward model, but that would be misinterpreting the meaning of the significant levels — it would be best to have the raw radiosonde observed temperatures for this method.

To make radiosonde data assimilation methods more consistent with GPSRO one should: change to a forward model which interpolates the model to the observation (data at intrinsic instrument resolution, not significant levels) in the vertical domain i.e. treat the unprocessed radiosonde data as independent point observations, use the GPS-derived height (available in future in BUFR format) as the vertical co-ordinate of the radiosonde and include a stricter quality control — a whole profile $2J/m$ type check e.g. see OSDP18 (2009) may be of benefit. If this could not be done, then to reduce the stratospheric bias from processing temperatures to layer-mean θ level values, one should move to a simple linear interpolation of temperature with log-pressure to surrounding θ levels.

To summarise, the final recommended actions following this report are:

1. The OPS temperature QC checks against background should be tightened to minimise the impact of biased data currently being assimilated — particularly in the stratosphere over Russia.

2. The WMO definition of significant levels perhaps needs to be tightened so that there is more consistency between independent radiosonde types.

3. If the observation temperature data continues to be disseminated in its current format, then a change from the OPS layer-mean processing of temperatures, to a simple linear interpolation onto θ level model pressures (in log-pressure) should be developed.

4. If possible radiosonde data should be disseminated in a high vertical resolution format. Ideally this would be at the resolution deemed representative for the instrument — rather than the raw measurement. This would then allow for the use of forward models that interpolate model data to the observation vertical location (pressure or geopotential). This would be more consistent with GPSRO assimilation. Such forward modelling would also eliminate the stratospheric bias currently introduced by the OPS processing and the potential over-weighting of the data with increases in model vertical resolution.

5. The analysis methods of this report should be reproduced following DAE updates to radiosonde pre-processing methods.

6. If independent geopotential height measurements (i.e. GPS-derived) become available in future BUFR format (not derived from pressure, temperature, humidity), then the temperature, pressure and humidity information should be assimilated as a function of geopotential height — which would be more consistent with GPSRO.

Acknowledgements

I would like to thank members of DAE (Colin Parrett, Rick Rawlins and Bruce Ingleby) for answering my numerous questions about radiosonde data assimilation.

References

He, W., Ho, S., Chen, H., Zhou, X., Hunt, D., and Kuo, Y., Assessment of radiosonde temperature measurements in the upper troposphere and lower stratosphere using COSMIC radio occultation data, *Geophys. Res. Lett.*, *36*, doi:10.1029/2009GL038 712, 2009.

Kursinski, E., Hajj, G., Schofield, J., Linfield, R., and Hardy, K., Observing earth's atmosphere with radio occultation measurements using the Global Positioning System, *J. Geophys. Res.*, *102*, 23.429–23.465, 1997.

OSDP18, OSDP 18 — GPSRO Processing, Met Office internal document, http://www-nwp/∼opsrc/OPS/view/dev/doc/OSDP18.html, 2009.

OSDP5, OSDP 5 — Radiosonde Processing, Met Office internal document, http://www-nwp/∼opsrc/OPS/view/dev/doc/OSDP5.html, 2008.

Rennie, M., The Impact of GPS Radio Occultation Assimilation at the Met Office, *Q. J. R. Meteorol. Soc.*, *accepted*, 2010.

Schreiner, W., Rocken, C., Sokolovskiy, S., Syndergaard, S., and Hunt, D., Estimates of the precision of GPS radio occultations from the COSMIC/FORMOSAT-3 mission, *Geophys. Res. Lett.*, *34*, L04 808, doi:10.1029/2006GL027 557, 2007.

Smith, E. and Weintraub, S., The constants in the equation for atmospheric refractivity index at radio frequencies, in *Proc. IRE*, vol. 41, pp. 1035–1037, 1953.

Steinbrecht, W., Claude, H., Schnenborn, F., Leiterer, U., Dier, H., and Lanzinger, E., Pressure and Temperature Differences between Vaisala RS80 and RS92 Radiosonde Systems, *J. Atmos. Ocean. Tech.*, *25*, 909–927, 2008.

Swinbank, R. and Wilson, C., Vertical Interpolation of Temperature Observations and Model Data, Short-Range Forecasting Research Technical Note 48, Met Office, 1990.

von Engeln, A., Healy, S., Marquardt, C., Andres, Y., and Sancho, F., Validation of operational GRAS radio occultation data, *Geophys. Res. Lett.*, *36*, doi:10.1029/2009GL039 968, 2009.

World Meteorological Organisation (WMO), Publication No. 306 - Manual on Codes, On-line, http://www.wmo.int/pages/prog/www/WMOCodes/VolumeI1.html, 2009.

Met Office
FitzRoy Road, Exeter
Devon, EX1 3PB
UK

Tel: 0870 900 0100
Fax: 0870 900 5050
enquiries@metoffice.gov.uk
www.metoffice.gov.uk