Q. J. R. Meteorol. Soc. (2004), **130**, pp. 607–626 doi: 10.1256/qj.03.76

The accuracy of stratospheric analyses in the northern hemisphere inferred from long-duration balloon flights

By A. HERTZOG[1]*, C. BASDEVANT[1], F. VIAL[1] and C. R. MECHOSO[2]

[1]*Laboratoire de Météorologie Dynamique, IPSL, Palaiseau, France*

[2]*Department of Atmospheric Sciences, UCLA, USA*

(Received 7 May 2003; revised 10 October 2003)

SUMMARY

In January and February 2002, six superpressure balloons (SPBs) were launched from Kiruna, Sweden (69° N, 21° E). The balloons drifted in the polar lower stratosphere for up to 45 days. Temperature, wind and altitude observations were collected every 15 minutes during the flights. This dataset serves as a reference to assess the accuracy of the European Centre for Medium-Range Weather Forecasts (ECMWF) analyses and the National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) reanalyses in the lower stratosphere during the 2002 winter. In particular, it is found that both models succeeded in simulating the large temperature fluctuations observed during the SPB flights, which were induced by a geographical shift between the stratospheric vortex and the low temperature area. Furthermore, the model biases are found to be small for all variables (temperature, wind, height). Nevertheless, it is found that NCEP/NCAR reanalyses tend to be slightly warmer (0.8 K) than the observations, while the converse is true for ECMWF analyses (−0.3 K). The observations also exhibit small-scale fluctuations that are presumably produced by (inertia-)gravity waves and which induce some scatter in the analysis/observation comparisons. Finally, trajectory comparisons are performed. It is found that trajectories built with ECMWF winds are more accurate than those built with NCEP/NCAR winds. With both models, the trajectory errors increase with time upto about 15 days and then tend to fluctuate. Typical errors after 15 days are 1000 ± 1200 km for ECMWF and 2300 ± 1300 km for NCEP trajectories.

KEYWORDS: ECMWF analyses NCEP/NCAR reanalyses Polar vortex Superpressure balloon

1. INTRODUCTION

Operational analyses and historical reanalyses compiled by various meteorological centres have proved to be very useful tools for a wide range of atmospheric studies. For example, in the stratospheric context which is the focus of this work, these (re)analyses provide the wind velocities that are used in chemistry-transport models (CTMs) to advect chemical species and in contour-advection or high-resolution models to study transport of tracers. Analysed temperatures are also essential for the modelling of stratospheric chemistry, as most reaction rates are temperature-dependent. Moreover, simulations of heterogeneous chemistry processes, of primary importance in ozone depletion, rely on microphysical schemes which are also highly dependent on temperature.

While large discrepancies between observations and analysed fields were reported some years ago (Knudsen and Carver 1994; Knudsen *et al.* 1996; Manney *et al.* 1996; Pullen and Jones 1997; Keil *et al.* 2001; Knudsen *et al.* 2001), Knudsen *et al.* (2002) showed that the accuracy of stratospheric analyses has considerably improved in the past decade, as a consequence of finer model resolutions and assimilation of new satellite-borne observations. However, recent studies (Pawson *et al.* 1999; Manney *et al.* 2003) found significant discrepancies between products from different centres, specifically at low temperatures. Additionally, the limited resolution of the analyses prevents adequate resolution of small to mesoscale stratospheric temperature disturbances induced by mountain waves (Hertzog *et al.* 2002a), for example.

These remaining shortcomings may have a significant impact in ozone loss computations. For instance, the formation of polar stratospheric clouds (PSCs) that convert inert halocarbons to ozone-depleting species (Peter 1997) requires low temperatures.

* Corresponding author: Laboratoire de Météorologie Dynamique, École Polytechnique, F-91128 Palaiseau Cedex, France. e-mail: hertzog@lmd.polytechnique.fr

TABLE 1. KIRUNA 2002 SUPERPRESSURE BALLOON FLIGHTS

Flight	Diameter (m)	Launch date	End date	Duration (days)	Density (kg m^{-3})	Mean altitude (km)	Mean Pressure (hPa)
SPB1	10	22/1	20/2	29	0.099	18.5	58.6
SPB2	10	22/1	6/2	15	0.099	18.6	58.8
SPB4	8.5	26/1	3/2	8	0.136	16.5	82.8
SPB5	8.5	1/2	11/2	10	0.138	16.4	85.1
SPB6	10	4/2	21/3	45	0.109	17.9	64.5
SPB7	10	5/2	11/2	8	0.110	17.8	64.7

Moreover, even though most PSCs may be generated in synoptic- or planetary-scale temperature disturbances (Teitelbaum and Sadourny 1998), mountain waves can also locally produce favourable conditions for PSC formation (Murphy and Gary 1995; Dörnbrack *et al.* 2001; Dörnbrack *et al.* 2002).

The reliability of CTM simulations is therefore highly dependent on the quality of the analysed thermal and dynamical fields released by meteorological centres. This paper, whose goal is to estimate the accuracy of recent analyses, compares observational data with the analyses produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) and the reanalyses produced by the National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR). These data were collected during long-duration superpressure-balloon (SPB) flights in the Arctic lower stratosphere in January and February 2002. Those observations were not assimilated in the ECMWF and NCEP/NCAR models and therefore provide an independent dataset that can be used to test the quality of both analyses.

The structure of the paper is as follows. The balloon flights and the observations collected during the campaign, as well as the ECMWF analyses and NCEP/NCAR reanalyses, are described in section 2. A short description of the Arctic vortex behaviour during winter 2001/02 is also given in this section. Section 3 is devoted to the comparison of the meteorological observations with the analyses at the balloon positions. In section 4, we estimate the accuracy of trajectories computed with analysed winds. This is achieved through comparisons between these simulated trajectories and the real balloon trajectories. A summary and final conclusions form the last section of this paper.

2. DATASETS

(*a*) *Long-duration balloon flights*

Six SPBs were launched from Kiruna, Sweden (69°N, 21°E) in January and February 2002. This campaign was part of the final stage in the assessment of the observing system designed for the Stratéole/Vorcore experiment (Vial *et al.* 1995).

The Vorcore system has two components: a SPB developed by the Centre National d'Etudes Spatiales (CNES, the French Space Agency), and a gondola developed by Laboratoire de Météorologie Dynamique. The SPBs are designed to drift in the lower stratosphere for several weeks carrying a scientific payload of up to 15 kg. By design, SPBs move on constant-density surfaces and thus behave as quasi-Lagrangian atmospheric tracers. Further details on the CNES SPBs and their behaviour in the atmosphere can be found in Vial *et al.* (2001), Hertzog and Vial (2001) and Cocquerez *et al.* (2001). During the Kiruna 2002 campaign, two types of SPB were used, of diameter 10 m and 8.5 m. As reported in Table 1, the smaller SPBs fly at a lower altitude than the

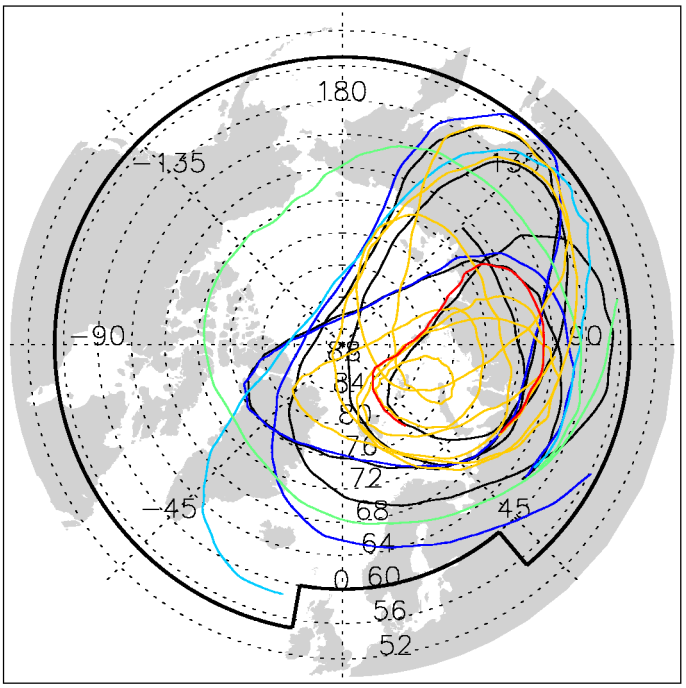

Figure 1. Superpressure-balloon trajectories during the Kiruna 2002 campaign: SBP1 black, SPB2 dark blue, SPB4 light blue, SPB5 green, SPB6 orange and SPB7 red. The authorized flight zone is located poleward of the thick black line. The first 24 hours of each flight are not shown.

larger ones. The flight-level differences within each SPB type are due to variations in the weights carried by the balloons.

The gondola itself—named Rumba—carries several sensors able to measure air temperature and pressure, uses the Global Positioning System (GPS) for location-finding and the ARGOS system for transmission of the collected data to ground stations. The accuracy of the temperature measurements is 0.3 K. However, a close inspection of temperature records during the Kiruna 2002 campaign revealed that daytime temperature measurements were slightly positively biased. The method used to correct this bias is presented in an appendix to this paper. The accuracy of the pressure sensor is 0.01% (approximately 0.6 Pa). The typical error in GPS position is 10 m in the horizontal and 20 m in the vertical. We estimate zonal and meridional wind speeds along the flight by finite differences of data from successive GPS points. The sampling rate is one measurement every 15 minutes, which yields a $\sim$0.02 m s^{-1} accuracy in the horizontal-wind speeds. Further details on the Rumba gondola can be found in Pommereau *et al.* (2002).

The SPB trajectories are shown in Fig. 1. Due to safety considerations, SPBs are not authorized to fly equatorward of 55°N (60°N above densely-populated Europe). This limit caused the automatic end of the SPB2, SPB4, and SPB5 flights. The SPB1 and SPB6 flights terminated because the gondolas were running out of power. A problem due to very low temperatures inside the gondola caused the end of the SPB7 flight. All SPBs except SPB5 flew inside the stratospheric vortex at all times.

(*b*) *Analyses*

In the following sections, we compare the balloon observations with the stratospheric analyses and reanalyses released by ECMWF and NCEP/NCAR, respectively. We present here briefly the main features of each dataset.

(i) *ECMWF analyses.* The ECMWF uses a spectral model with a T511 truncation and a four-dimensional variational assimilation scheme to issue its atmospheric analyses every 6 hours (Untch and Simmons 1999; Rabier *et al.* 2000). The analyses are available on a $0.5^\circ \times 0.5^\circ$ grid. The model has 60 levels in the vertical and extends from the ground, where the vertical coordinate is terrain-following, to 0.1 hPa, where the vertical coordinate is pressure. In the lower stratosphere, the vertical spacing between levels is 1.5 km. The assimilation scheme has used the TIROS Operational Vertical Sounder (TOVS/ATOVS) radiances and the Advanced Microwave Sounding Unit radiances since 1998 (McNally *et al.* 1999), which have contributed to significant improvements in the analysed temperatures within the polar vortex (Knudsen *et al.* 2002).

(ii) *NCEP/NCAR reanalyses.* The NCEP/NCAR reanalysis project uses a frozen state-of-the-art analysis/forecast system, and performs data assimilation from 1957 onwards. The model is identical to the NCEP global model implemented in January 1995, but with a smaller horizontal resolution (T62) and with 28 layers in the vertical. Observational data are assimilated using spectral statistical interpolation or three-dimensional variational assimilation, with no need for nonlinear normal mode initialization. Meteorological variables are available every 6 hours on a $2.5^\circ \times 2.5^\circ$ grid at 17 levels from 1000 hPa to 10 hPa. Further details on the NCEP/NCAR reanalyses can be found in Kalnay *et al.* (1996).

Recently, Trenberth and Stepaniak (2002) found a pathological problem in the NCEP/NCAR reanalyses that was attributed to difficulties with the vertical coordinate used in the model. The reanalyses tend to produce spurious large-amplitude waves in the uppermost levels (pressure <50 hPa) above steep topography. Since SPBs were flying at lower altitudes during the campaign and away from major topography, we think that this problem does not influence our results significantly.

(*c*) *Vortex evolution in the Arctic winter of 2001/02*

Maps of potential vorticity and temperature on the 475 K isentrope (which is close to the balloon flight levels) are shown in Fig. 2. In December 2001 and early January 2002, the vortex in the lower stratosphere was frequently disturbed by planetary-wave activity. A minor warming occurred in mid-January and led to the detachment of a high potential vorticity (PV) cell out of the main vortex body. This structure can be seen on the map for 22 January, i.e. the day of the first SPB launch. The main vortex was centered over northern Siberia, and there was a high PV anomaly located above northeastern Canada. Minimum temperatures on this day were found above Finland and western Russia, while maximum temperatures covered a belt extending from eastern Siberia to northeastern Canada.

The main vortex hardly moved during February and early March. On 15 February, a small shift to the east was seen on the PV and temperature fields. The vortex was then centered above northern Siberia and the minimum temperatures were found above northwestern Siberia. The high-PV anomaly that resulted from the warming in mid-January had disappeared. The temperature inside the vortex decreased slightly, as can be seen by the change to the 200 K contour.

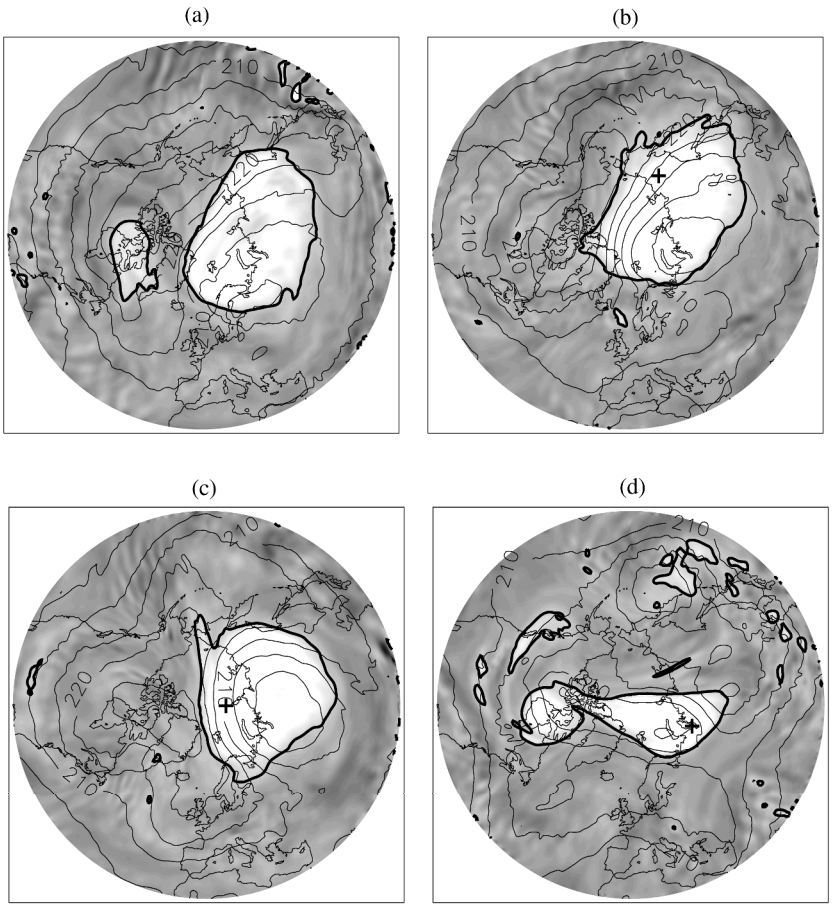

Figure 2. Potential vorticity (PV) and temperature on the 475 K isentrope at 00UTC on four dates covering the balloon flights: (a) 22 January, (b) 15 February, (c) 1 March and (d) 21 March 2002. PV is shown as a grey scale, with light (dark) grey indicating high (low) values. The bold PV contour, corresponding to 30 PV units, is located in the PV gradient that determines the vortex edge. Temperatures are shown as thin contours at 5 K intervals. On (b), (c) and (d), a cross indicates the position of SPB6 on that day.

On 1 March, the vortex was still centred over northern Siberia. The temperature inside the vortex had decreased further and the minimum temperature was now centred above western Siberia.

A major stratospheric warming occurred in the second half of March and the lower stratospheric vortex split into two parts. On 21 March (the last day of the longest flight), the vortex was almost broken with one part centred above Novaya Zemlya and the other above Baffin Bay, and the temperature had significantly increased in the polar region to be >200 K everywhere.

Thus, except for the end of SPB6 flight, the polar stratospheric vortex during the Kiruna 2002 balloon campaign was relatively stable, and displaced off the pole by a wave-1 pattern. During most of the campaign, the temperature and PV structures were not concentric, the temperature minima being predominantly located in the southwestern flank of the vortex. Such a large-scale shift between temperature and PV structures is actually of great interest for ozone-loss computations, since it prevents PSC particles from reaching very large sizes and therefore limits vortex denitrification (e.g. Mann

et al. 2002). Comparisons between observations and analysed fields will therefore provide interesting information on how well the observations agree with the stratospheric structure depicted by the analyses during the campaign period.

3. Observation/analysis comparisons

(*a*) *Method*

The analysed fields (temperature, wind, geopotential height) were interpolated to the balloon positions by means of cubic splines in order to facilitate comparison with observational data. The three nearest time steps on each side of the observation time were used for the temporal interpolation, and the closest 8×8 grid points were used for the horizontal interpolation. In the vertical, the observed pressure, which is more accurate than the GPS altitude, was taken to determine the SPB position. We used the logarithm of pressure as the coordinate for the vertical interpolation. The same levels in both models (500, 400, 300, 250, 200, 150, 100, 70, 50, 30, 20, and 10 hPa) were used for this log-linear vertical interpolation.

Separate temperature comparisons for daytime and nighttime will be shown, with the daytime observations corrected as described in the appendix.

(*b*) *Results*

(i) *Illustration of flight SPB6.* The comparisons between SPB6 observations and ECMWF and NCEP analyses are shown in Fig. 3. The SPB6 flight is the longest of the Kiruna 2002 campaign and allows us to illustrate typical features of this campaign.

Analysed and observed temperatures are presented in Fig. 3(a). The observed values show temperature fluctuations of up to 15 K over time-scales of 5–6 days. These fluctuations, which are a direct consequence of the non-concentricity of the PV structure (vortex) and the temperature field, confirm the vortex picture discussed in the previous section. Thus the ECMWF analyses and NCEP reanalyses were able to capture the geographical pattern associated with the lower stratospheric vortex during the 2001/02 winter. Furthermore, the analysed temperatures are in very good agreement with the measurements; ECMWF (NCEP) temperatures are within 1 K of the observations for more than 75% (55%) of the time.

However, there are some small discrepancies between analyses and observed data. The NCEP reanalyses tend to slightly overestimate temperatures (warm bias), while the converse is true for ECMWF analyses (cold bias). The warm NCEP bias can be up to 3 K (for instance on day 47). The cold ECMWF bias is generally smaller, but can occasionally reach 2 K (on day 54, for instance). Another difference between the observations and the analyses is associated with the short-timescale variance that is present in the temperature observations but which is not resolved by the 6-hourly analyses. These ubiquitous temperature fluctuations are induced mainly by gravity waves and have typical amplitudes of 1–2 K (Hertzog *et al.* 2002a).

The largest fluctuations of SBP6 altitude (Fig. 3(b)) are inversely correlated with those of temperature. This simply reflects the fact that upward (downward) vertical displacements of isopycnic surfaces are associated with adiabatic cooling (heating) of air parcels. (The vertical displacements of isopycnic surfaces that are induced by Rossby or gravity waves are actually in phase with those of the isentropes, even though they are generally smaller (Hertzog and Vial 2001).) The balloon altitudes are very closely reproduced by both analyses. Typical differences are less than 50 m, which is very close to the error of the altitude measurements. However, at the end of the flight, the analyses

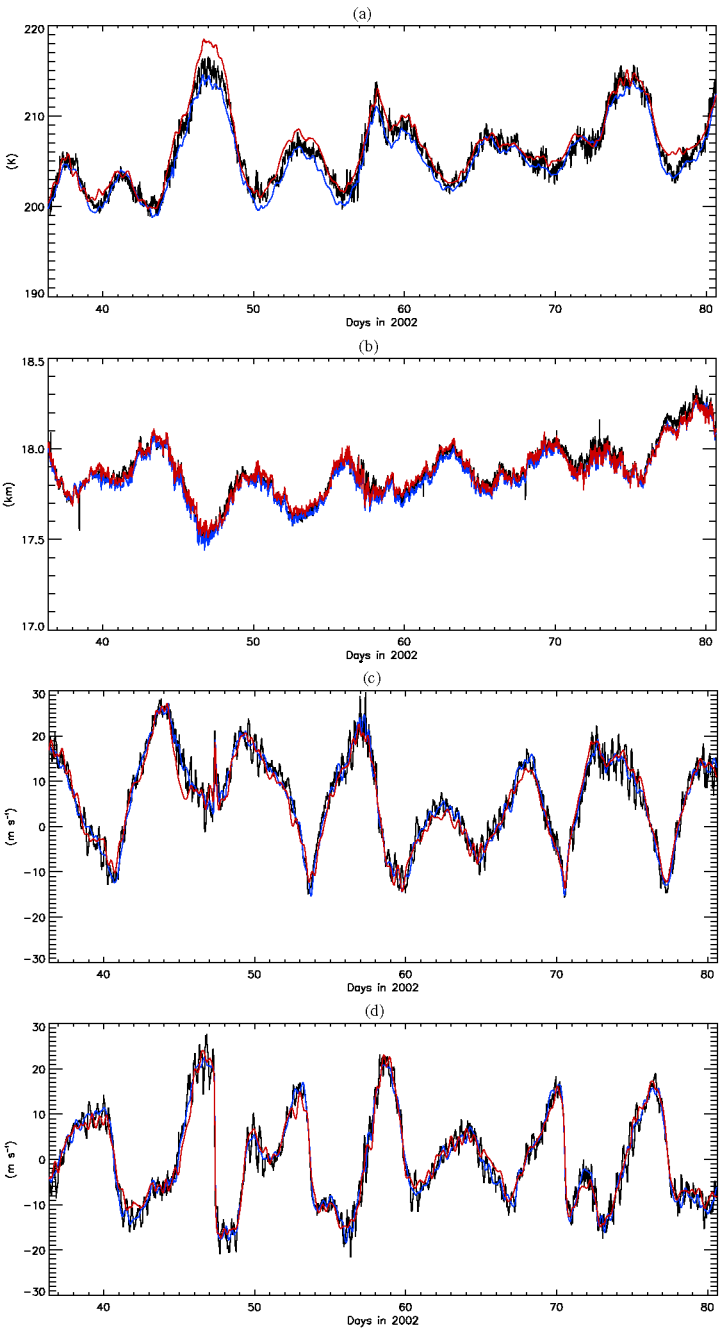

Figure 3. Comparison between SBP6 observations (black), ECMWF (blue) and NCEP (red) analyses: (a) temperature, (b) altitude, (c) zonal speed and (d) meridional speed.

tended to underestimate the SPB6 altitude, but the meteorological situation was quite disturbed at that time (see Fig. 2).

Zonal and meridional wind speeds are shown in Figs. 3(c) and (d). Once again, both analyses are in close agreement with the observations. The largest fluctuations in the observed speeds are mainly due to the planetary wave-1 pattern that displaced the vortex off the geographical pole. The agreement between observations and analyses of this large-scale stratospheric structure is excellent.

The observed speeds also exhibit fluctuations with periods close to 12 hours and typical amplitudes around 3–4 m s^{-1}. These fluctuations are induced by inertia–gravity waves and are ubiquitous in the polar lower stratosphere in winter (Hertzog *et al.* 2002b). As for the short-timescale temperature variance, the 6-hourly analyses do not adequately capture these waves, even though some hint of them can be seen on the analysed speeds. However, the fluctuations in the interpolated time series are not related to the presence of gravity waves in the analyses but rather to their presence in the real atmosphere; inertia–gravity waves induce horizontal displacements of the balloons which can be as large as 50 km. The balloon positions are then used in the interpolation of the analysed fields. These displacements, with typical periods of 12 hours, can therefore produce fluctuations with the same time period in the interpolated analysed fields, provided that the large-scale gradient in the analysed field is not too weak, which is obviously the case for the wind speed near the vortex edge.

(ii) *Histograms of differences.* In this section, a more quantitative comparison between balloon observations and analyses is presented as histograms of differences (analysis minus observation). These are shown in Fig. 4 for temperature, altitude, zonal and meridional speeds for all the 2002 balloon flights. The moments of these distributions are shown in Table 2.

The small ECMWF cold bias and NCEP warm bias found in SPB6 observations are confirmed over all cases (−0.3 K and 0.8 K, respectively, with both figures significant at the 99% confidence level). The standard deviations of the corresponding temperature histograms are 0.8 K and 1.0 K, respectively. Furthermore, daytime and nighttime distributions are found to be identical; the hypothesis that they represent the same distribution cannot be rejected at the 95% confidence level. This is illustrated in Fig. 5, which represents the separate day and night histograms of temperature difference between ECMWF analysis and observations. (The corresponding NCEP histograms look very similar.) When the observations are low-pass filtered to exclude short-period gravity waves (with a cutoff period at 18 hours, for instance), the standard deviations are reduced to 0.6 K (ECMWF) and 0.8 K (NCEP). Thus, gravity waves are partially responsible for the discrepancies between the models and the observations. Nevertheless, other sources of disagreement still exist. Since SPB6 (and the other balloons) mostly flew close to strong horizontal temperature gradients, a small shift in the analysed temperature (or PV) field with respect to the real stratosphere may contribute to the standard deviation.

Altitude histograms (Fig. 4(b)) indicate that both analyses exhibit a small positive bias with respect to the observations. The values of these biases are 20 m for ECMWF and 50 m for NCEP. To obtain these values, the analysed geopotential height, z^*, was converted to height above mean sea level, z, according to

$$z = \frac{Rz^*}{R - z^*}, \tag{1}$$

where R is the mean Earth radius. The standard deviations of the altitude histograms are 24 m (ECMWF) and 29 m (NCEP). However, the GPS model used during this campaign

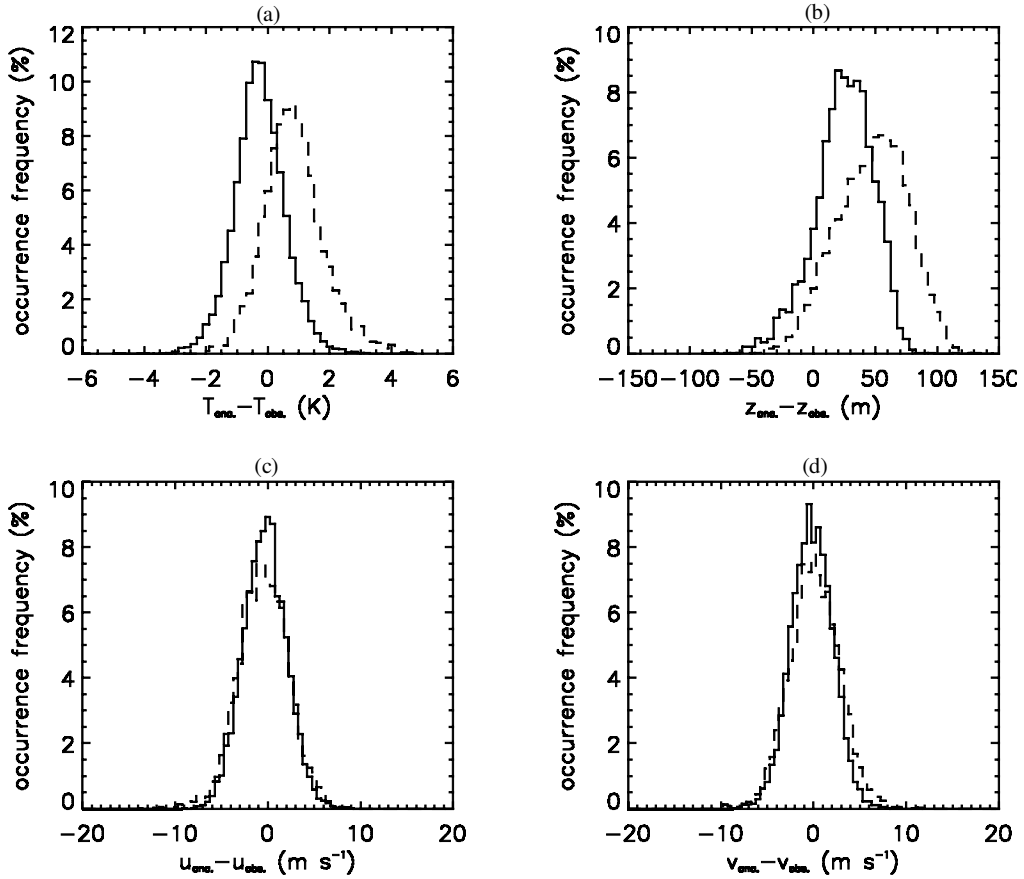

Figure 4. Histograms of differences—analyses (ECMWF solid, NCEP dashed) minus balloon observations: (a) temperature, (b) altitude, (c) zonal speed and (d) meridional speed.

TABLE 2. STATISTICS OF ECMWF/NCEP ANALYSIS MINUS BALLOON OBSERVATION

	Temperature (K)		Altitude (m)		Zonal speed (m s^{-1})		Meridional speed (m s^{-1})	
	ECMWF	NCEP	ECMWF	NCEP	ECMWF	NCEP	ECMWF	NCEP
Bias	−0.3	0.8	20.0	50.0	−0.1	−0.3	0.1	0.3
Standard deviation	0.8	1.0	20.0	30.0	2.3	2.7	2.2	2.7
Skewness	0.1	0.4	−0.5	−0.2	0.1	−0.2	0.0	0.0
Excess kurtosis	0.9	1.1	0.4	−0.4	0.0	0.6	0.1	0.4

Skewness and excess kurtosis are without dimension. The excess kurtosis is defined as (kurtosis − 3).

only provides altitudes above a reference ellipsoid (WGS-84), while the actual geoid, i.e. the iso-geopotential surface corresponding to the mean sea level, differs from the reference ellipsoid by values that are typically a few tens of metres and that locally can reach up to 100 m. Therefore, it is difficult to go any further in the interpretation of the analysis biases, and particularly to determine the real accuracy of the analysed height.

Figures 4(c) and (d) present difference histograms for the zonal and meridional wind speeds. They primarily show that analyses produce unbiased horizontal speeds. However, the standard deviations of these histograms are relatively large: $\sim$2.2 m s^{-1}

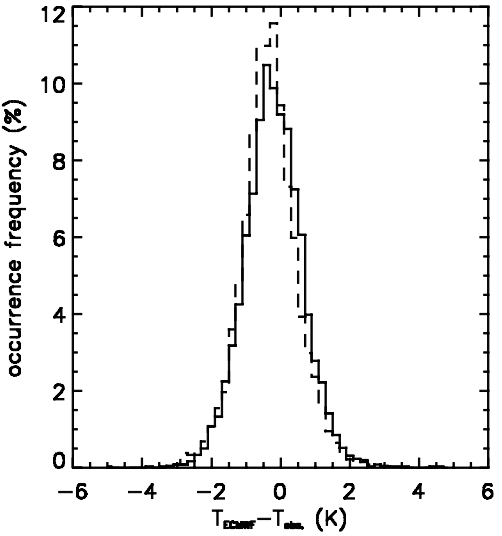

Figure 5. Histogram of temperature differences (ECMWF analysis minus observation); day—solid and night—dashed.

(ECMWF) and ∼2.7 m s^{-1} (NCEP). (The standard deviations are almost the same on both components for each analysis). A large part of this variance is caused by inertia-gravity waves; values become 1.3 m s^{-1} and 1.9 m s^{-1}, respectively, when the observations are low-pass filtered. Note also that the NCEP histograms tend to have a larger spread than those for ECMWF. This is seen in the kurtosis figures reported in Table 2.

(*c*) *Discussion*

The comparisons between *in situ* observations and recent analyses performed at ECMWF and NCEP show that the analyses provide a very fair picture of the lower stratosphere meteorology. The large-scale patterns are well reproduced by the analyses and the biases, when they exist, are small. Nevertheless, in agreement with Knudsen *et al.* (2002) and Manney *et al.* (2003), we found that NCEP reanalyses tend to produce higher temperatures than ECMWF analyses. Knudsen *et al.* (2002), who compared various analyses with balloon-borne temperature measurements performed in recent (1997, 1999, 2000) Arctic winters, found that ECMWF analyses have a -0.5 ± 0.9 K (cold) bias in the lower stratosphere in 2000. Our results for the biases (-0.3 ± 0.8 K) are very close to their estimates, and even suggest that ECMWF analyses have improved since 2000.

There was a problem with the TOVS-data processing in the 1999 and 2000 NCEP reanalyses used in Knudsen *et al.* (2002). (The description of this problem can be found at http://www.cdc.noaa.gov/cdc/reanalysis/problems.shtml) It is therefore not very meaningful to compare our result (0.8 ± 1.0 K) with theirs for 1999 and 2000. Nevertheless, our results are close to theirs for the 1997 winter (0.5 ± 1.22 K).

We have also tried to identify trends in the differences between the observations and analyses (e.g. systematically larger biases in lower temperatures, as reported in Manney *et al.* (2003)). However, no significant trend in the temperature or in any other variable was detected. However, the vortex in 2001/02 was warmer than those studied by Manney

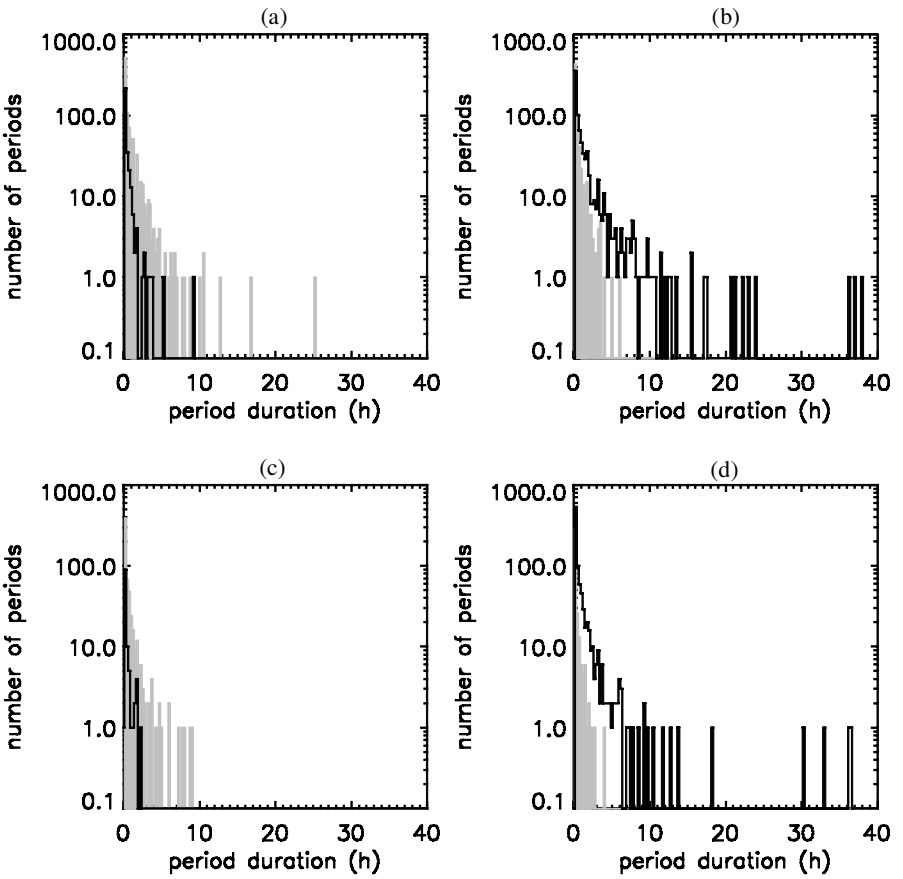

Figure 6. Duration of time periods during which the analysed temperatures (ECMWF shaded and NCEP solid lines) are (a) lower and (b) higher than the observed temperatures by over 0.5 K. (c) and (d) are similar, but using a threshold of 1.0 K.

et al. (2003), who found significant trends only at temperatures lower than ~195 K. Such low temperatures were not observed during the Kiruna 2002 balloon campaign.

The histograms show a significant scatter in the observation/analysis differences. Part of this scatter was attributed to the effect of inertia-gravity waves, which are obviously present in the observations but not in the time series extracted from the analyses. As an apparent paradox, many authors (e.g. Moldovan *et al.* 2002; Hertzog *et al.* 2002a) actually found inertia-gravity waves in global-scale analyses. These waves were identified on instantaneous maps of wind velocity or divergence. Thus, the models used to produce the analyses are intrinsically able to resolve at least part of the wind variance induced by inertia-gravity waves. However, the fact that analyses are available only every 6 hours causes those waves to disappear during the temporal interpolation that we performed for the comparison with the observations; inertia-gravity waves have frequencies that are too close to the Nyquist frequency. More generally, any temporal interpolation (e.g. when computing air-parcel trajectories) will severely damp the inertia-gravity waves present in the analyses.

The temperature biases of the analyses and the associated scatters are particularly relevant for microphysical processes in the polar stratosphere. Furthermore, the duration

of time periods for which the analysed temperatures significantly differ from the observed temperatures is important. During such periods, microphysical models may seriously either over- or underestimate particle formation or destruction. Figure 6 shows the duration of continuous periods during which the absolute values of temperature difference exceeded 0.5 K and 1 K. All the SPB flights during the campaign are included in these statistics. In agreement with the biases reported above, the periods during which ECMWF temperatures are significantly lower than the observations tend to be longer than those during which they are higher. The converse is true for NCEP analyses. However, the majority of these periods are found to be of relatively short duration (less than 5 hours). Nevertheless, a 1-day period was found for which the ECMWF temperatures were continuously lower than observations by 0.5 K, and several periods longer than 1 day were observed with NCEP temperatures higher than observations by more than 1 K.

Finally, no strong mountain wave event (as indicated by temperature fluctuations of several degrees) was observed during the Kiruna 2002 campaign. The vortex configuration actually prevented the balloons from flying over Scandinavia and Greenland, which are favourable regions for the occurrence of mountain waves in the Arctic. In such mountain waves, general-circulation models generally significantly underestimate temperature perturbations (e.g. Hertzog *et al.* 2002a).

4. Trajectories

(*a*) *Method*

In this section, trajectories computed from the ECMWF and NCEP winds are compared to the real balloon trajectories. More precisely, we have used the analyses to compute isopycnic trajectories in order to mimic the SPB behaviour in the atmosphere. (Actually, typical relative variations of air density recorded during the SPB flights in 2002 are less than 1%, corresponding to vertical displacements of ± 50 m about the mean density level.) The procedure used to interpolate the horizontal-wind fields onto an isopycnic surface is as follows:

- At each analysis time and horizontal-grid point, the analysed temperature is used to compute the density at each model pressure level.
- This new density coordinate is then used to interpolate the horizontal wind on the desired density level (which differs from one SPB to another, as seen in Table 1).

Simulated balloons were launched every 24 h along the real SPB trajectories. The simulated balloons were advected until the end of the real SPB flight via a second-order Runge-Kutta scheme (The use of a fourth-order Runge-Kutta scheme has almost no impact on the results shown below). We have used a time step of 10 minutes for the temporal integration. Special attention has been paid when the simulated trajectories came close to the North Pole. In particular, the time step was decreased in such cases, in order to respect the Courant–Friedrichs–Lewy criterion. As an example, the simulated trajectories for flight SPB4 with ECMWF winds are shown in Fig. 7.

The distance on the globe between the simulated and real SPB flights was recorded every 15 minutes. The next section presents statistics of this distance obtained from the longest SPB flights (SPB1, SPB2 and SPB6).

(*b*) *Results and discussion*

Figure 8 shows the mean separation between the real and simulated SPB flights as a function of the elapsed time since the simulated SPB launches. The mean separation

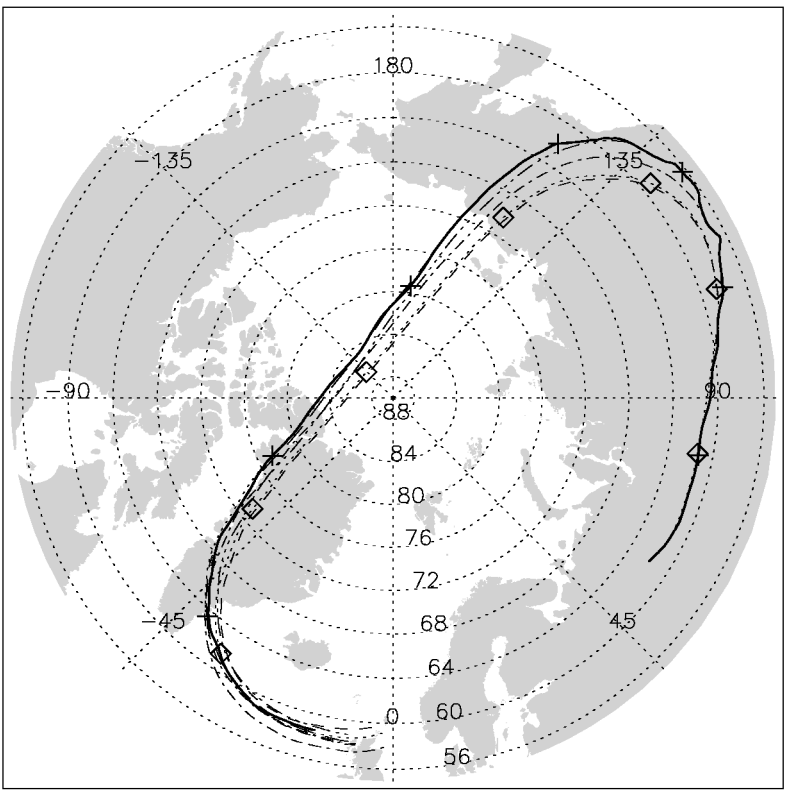

Figure 7. SPB4 trajectory (solid) and the 8 isopycnic trajectories computed with the ECMWF analysed winds (other line styles). Symbols are drawn every day at 00 UTC: + on the real SPB4 trajectory, and ◇ on one of the simulated trajectories.

and the standard deviation increase almost linearly during the first days and this increase is much more rapid for the NCEP reanalyses than for the ECMWF analyses. After 5 days of 'flight', the separation between the real SPBs and the NCEP-simulated SPBs is 1700 ± 1400 km, while it is only 270 ± 230 km with the ECMWF-simulated SPBs. The separation continues to increase upto around 10 days, where it begins to level off. So, for flight duration between 10 and 30 days, the separation between real and simulated SBPs tends to oscillate typically about 1000 ± 1200 km for ECMWF, and 2300 ± 1300 km for NCEP. For comparison purposes, the maximum distance for two air parcels lying on the same PV contour as SPB6 is typically 5000 km in February and early March (i.e. before the vortex split, see Fig. 2). The reported smaller trajectory errors obtained with ECMWF-simulated SPBs are probably linked to the fact that the occurrence of large wind errors is more likely with NCEP winds, as revealed by the kurtosis values in Table 2. A possible reason for this discrepancy is the better vertical and horizontal resolution of the ECMWF model in the lower stratosphere. We have also found that the typical decorrelation time for the wind differences is larger with the NCEP reanalyses (20 hours) than with the ECMWF analyses (10 hours). In other words, the ECMWF wind errors seems to dissipate more quickly than the NCEP wind errors, which can also partly explain the better ECMWF trajectories.

After 30 days, our statistics rely entirely on SPB6 and the separation between the real and the ECMWF-simulated SPB trajectories rapidly increases, whereas it decreases

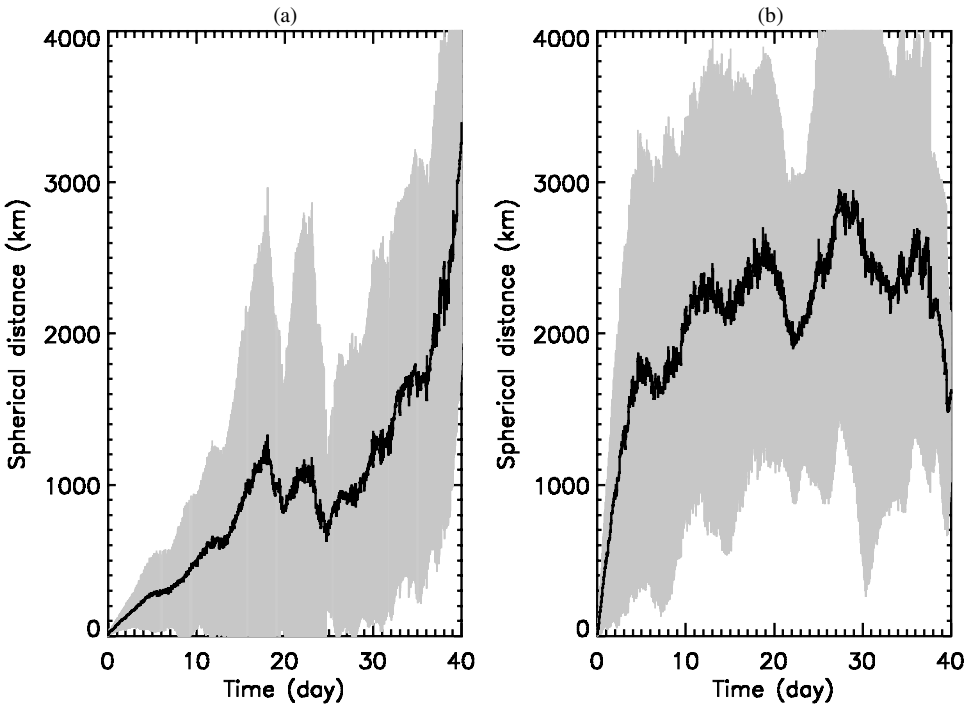

Figure 8. Mean (solid line) and standard deviation ($\pm 1\sigma$) (grey shading) of separation distance between real and simulated SPB trajectories as a function of the elapsed time since the simulated SPB launch. Simulations were computed with (a) ECMWF and (b) NCEP winds.

slightly for the NCEP-simulated trajectory. This significant difference in behaviour of the long-duration simulations may be explained by the splitting of the polar vortex (see Fig. 2), after which more NCEP-simulated SPBs were trapped in the same part of the vortex as the real balloon. Further interpretation of the statistics on long-duration flights is inappropriate, except to note that following an air parcel is not an easy task in very dispersive meteorological situations, e.g. during stratospheric warmings. In particular, we do not believe that such statistics reveal different skills of the analyses used in this paper.

Therefore we will now concentrate on simulated SPB flights shorter than one month; our statistics are more robust for such periods, having more trajectories and, more significantly, the use of longer air-parcel trajectories would require careful consideration of diabatic processes, whose representation within GCMs is still problematic. A striking feature of Fig. 8 is that both the mean separation and standard deviation exhibit oscillations after their initial increases (i.e. between 10 and 30 days). These oscillations are linked to the ellipsoidal shape of the stratospheric vortex in February 2002 and are well illustrated by Fig. 7. It can been seen that the distance between SPB4 and the simulated trajectory marked by diamonds increases when the flow is accelerated, while it decreases when the flow is decelerated in large-curvature areas. This feature is reproduced by numerous simulated SPBs and is the reason for the oscillations shown on Fig. 8. It can also be seen on Fig. 7 that the along-flow separation between SBP4 and the simulations tends to be far larger than the cross-flow separation. This is a general behaviour in our experiments, as has already been noticed by Knudsen and Carver (1994).

Compared to the results of Knudsen *et al.* (2001), our results on trajectory separation seem to indicate that the ECMWF analyses have improved since 1999 and are able now to compute more accurate trajectories. However, it should be kept in mind that their most recent results on trajectory errors were obtained from infrared Montgolfier flights. These balloons perform large vertical excursions at sunset and sunrise and, consequently, trajectory simulations with such balloons are much more difficult than with SPBs. This difference may explain part of the observed improvement.

The last question for discussion is whether our results on isopycnic trajectories can be applied to air-parcel trajectories that are generally assumed to be isentropic. Neither of the two models is built on isentropic or isopycnic surfaces; the ECMWF model uses almost constant-pressure levels in the stratosphere, while the NCEP model levels are hybrid up to the top of the domain. Therefore, for both analyses, a vertical interpolation has to be performed from the model levels to either isopycnic or isentropic surfaces and, consequently, there is no obvious reason for the ECMWF and NCEP fields to be more accurate on isentropes than on isopycnic surfaces. Thus, we believe that our results on isopycnic trajectories (which are a very good approximation of the real SPB behaviour) can give a fair estimate of the accuracy of air-parcel trajectories. However, it should be stressed that the statistics we obtained on trajectory accuracies are most probably valid only for the interior of the (Arctic) stratospheric vortex. In particular, the stratospheric flow is known to be much more dispersive in the midlatitude surf zone (e.g. Haynes and Shuckburgh 2000), and consequently simulated trajectories may be less accurate there.

5. Conclusions

We have used meteorological observations collected during long-duration SPB flights in the 2002 Arctic lower stratosphere to assess the accuracy of (re)analyses operationally released by ECMWF and NCEP/NCAR. The analyses agree very closely with the observations on the large-scale meteorological conditions that prevailed in the lower stratosphere at that time. In particular, the shift between the vortex core and the region of lowest temperature depicted in the analyses induces the large-amplitude temperature fluctuations observed during the SPB flights. More generally, the histograms of the analysis/observation differences reveal that the biases of ECMWF and NCEP analyses are small. Nevertheless, NCEP reanalyses tend to produce slightly higher temperatures (+0.8 K) than observed, while the converse is true for ECMWF (−0.3 K). These histograms also underline the significant scatter found in the comparisons between observed and analysed variables (temperature and winds primarily). Part of this scatter is attributed to mesoscale inertia-gravity waves, which are obviously present in the observations but are either misrepresented in the analyses or damped during temporal interpolations (since the analyses are released only every 6 hours). A possible (small) shift of the vortex structure in the analyses can also contribute to this scatter.

We have also compared the real SPB trajectories with isopycnic trajectories computed by using ECMWF and NCEP analysed winds. Our results show that ECMWF trajectories tend to be more accurate than NCEP trajectories in the polar stratosphere during the selected period of study. Nevertheless, the separation between real and simulated SPBs is significant whatever the analyses. The mean separation for a 15-day trajectory is typically 1000 ± 1200 km with ECMWF analyses, and 2300 ± 1300 km with NCEP reanalyses.

Our results are based on a relatively small dataset of six SPB flights. Nevertheless, the *in situ* information has provided what is arguably one of the most significant and accurate non-assimilated datasets in the lower-stratospheric vortex during the northern

winter of 2002. Consequently, our findings provide useful information on the real accuracy of the ECMWF and NCEP analyses. Furthermore, operational analyses and assimilation processes keep evolving, as does their accuracy. Therefore, larger flight campaigns with longer flight durations are certainly needed in the future to confirm or modify the figures given in this paper.

Acknowledgements

The authors are very grateful to all the CNES people involved in the successful Kiruna 2002 superpressure balloon campaign, and in particular to Philippe Cocquerez for managing the SPB Vorcore project. The work of Bernard Brioit and his team at Laboratoire de Météorologie Dynamique on the Rumba gondola is also gratefully appreciated. ECMWF is acknowledged for providing the operational analyses used in this article. NCEP/NCAR reanalysis data were provided by the NOAA-CIRES Climate Diagnostics Center, Boulder, Colorado, USA, from the mirror web site at http://climserv.lmd.polytechnique.fr/. We are also grateful to Martin Juckes and an anonymous reviewer for their numerous suggestions that have helped to improve this paper.

Appendix

Correction of temperature measurements

A common problem in daytime temperature measurements in the stratosphere is that the temperature sensors (thermistors) are themselves heated by solar radiation and consequently tend to overestimate the actual atmospheric-gas temperature. Therefore, it would be helpful to be able to correct the temperature measurements for this effect, as is done in meteorological soundings (e.g. Luers and Eskridge 1995). The daytime-heating problem is even more severe for SBPs, since these are advected by the wind and, consequently, the thermistor ventilation is low (in contrast with radiosoundings for which the balloon ascent ensures a good sensor ventilation). This appendix describes how we have empirically estimated the daytime temperature bias and shows the effect of the proposed correction.

Our first assumption is that the thermistors give an unbiased estimate of the atmospheric temperature at night. This is in agreement with several papers reporting on the same method for measuring temperature in the atmosphere (Fourrier *et al.* 1970; Cadet and Ovarlez 1974). Our second assumption is that the daytime temperature bias depends only on the solar zenith angle (SZA), which is valid at first order. Using these assumptions, our goal is thus to determine the functional dependence $\Delta T\,(\mathrm{SZA})$, where ΔT is the measurement bias. Since temperatures exhibit large fluctuations during the balloon flights, a mere regression of the measured temperatures on the SZAs will provide a rather imprecise estimate of that functional dependence. However, we can obtain a good estimate of the derivative $\mathrm{d}(\Delta T)/\mathrm{d}(\mathrm{SZA})$ by finite differences between two measurement points, since the largest temperature fluctuations are associated with timescales much longer than 15 minutes. We can then integrate this derivative, taking into account that the measurement bias vanishes at large SZAs (nighttime). The resulting estimate of the $\Delta T\,(\mathrm{SZA})$ law is shown in Fig. A.1.

In agreement with our assumption, the bias is almost zero for SZA above 94.5°. This is the angle at which the sun rises or sets for an observer at 20 km above the ground. The bias increases rapidly for SZAs between 95° and 80°, and reaches a limit of roughly 1.2 K at lower SZAs. We have fitted the empirical behaviour shown in Fig. (A.1)

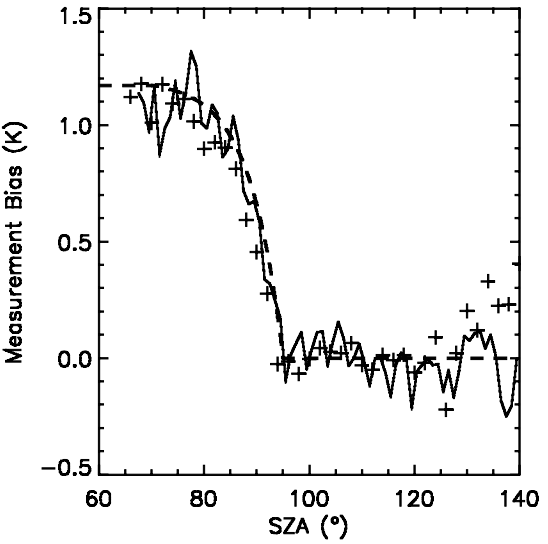

Figure A.1. Determination of the daytime bias in temperature measurements; empirical result from observations (solid line), the modelled correction scheme (heavy dashed) and ECMWF–observation differences when the heating correction is not applied (crosses).

by the following expression:

$$\Delta T = \begin{cases} 0 & \text{SZA} \geq 95^\circ, \\ 1.19 \left\{ 1 - \exp\left(\dfrac{95 - \text{SZA}}{6.21} \right) \right\} & 70^\circ < \text{SZA} < 95^\circ, \\ 1.17 & \text{SZA} \leq 70^\circ. \end{cases} \tag{A.1}$$

The modelled values are shown by the heavy dashed line on Fig. A.1. Also shown are the differences between the raw temperature observations and the ECMWF temperatures offset by their bias estimated at SZA $> 95^\circ$ (crosses). These differences nicely follow the empirically determined law, which gives confidence in our heating estimation.

One way of assessing the effect of this correction is to compare the frequency spectra of corrected and uncorrected temperature fluctuations. The spectra derived from the SPB6 flight are shown in Fig. A.2. The only significant difference between the spectra is at frequencies close to the diurnal frequency. The energy peak observed on the raw-temperature spectrum disappears on the corrected-temperature spectrum. This is a desirable effect, since there are many reasons to believe that the diurnal peak on the raw-temperature spectrum, corresponding to fluctuations with an amplitude of $\sim$0.6 K, is spurious. Temperature fluctuations induced by the diurnal tide at the balloon float level and latitude are believed to be $\sim$0.1 K, i.e. they are associated with an energy 36 times lower than that shown on the raw-temperature spectrum at the diurnal frequency. On the other hand, if the diurnal peak of temperature were caused by diurnal vertical excursions (20 m on average) of SPB6, this would imply a vertical temperature gradient of 30 K/km, which is unrealistically large.

All the temperature observations shown in this paper have been corrected for solar-induced heating according to Eq. (A.1).

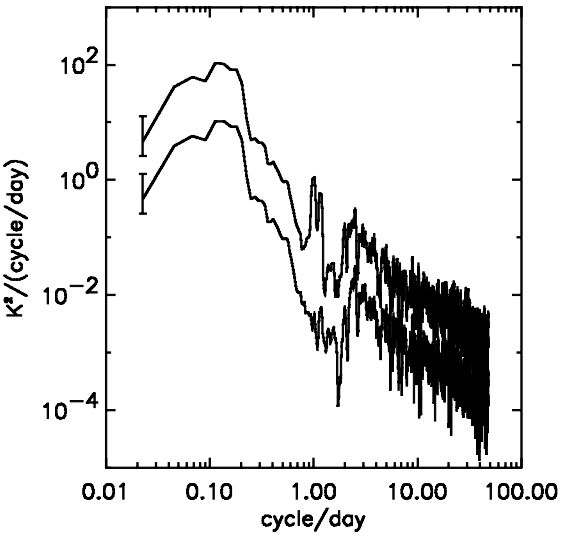

Figure A.2. Frequency spectra of temperature fluctuations during the SPB6 flight, computed from raw temperature measurements and after the heating corrections have been applied. For clarity, this latter spectrum has been shifted downwards by one decade. The error bars show the 90% confidence levels.

References

Cadet, D. and Ovarlez, J. 1974 'La mesure précise de la température dans la troposphère et la basse stratosphère: équilibre thermodynamique du senseur.' Technical Note 62, Laboratoire de Météorologie Dynamique, France

Cocquerez, P., Guigue, P., Guilbon, R., Phulpin, T., Durand, M., Lefevre, J.-P., Eymard, M. and Lafourcade, M. 2001 'Test flights of CNES stratospheric superpressure balloons in experimental Arctic campaigns 2000-2001: Objectives and results'. Pp 597–601 in Proceedings of 15th ESA Symposium on European Rocket and Balloon Programmes and Related Research ESA SP-471. Eur. Space Agency, Noordwijk, The Netherlands

Dörnbrack, A., Leutbecher, M., Reichardt, J., Behrendt, A., Müller, K. P. and Baumgarten, G. 2001 Relevance of mountain wave cooling for the formation of polar stratospheric clouds over Scandinavia: Mesoscale dynamics and observations for January 1997. *J. Geophys. Res.*, **106,** 1569–1581

Dörnbrack, A., Birner, T., Fix, A., Flentje, H., Meister, A., Schmid, H., Browell, E. V. and Mahoney, M. J. 2002 Evidence for inertia-gravity waves forming polar stratospheric clouds over Scandinavia. *J. Geophys. Res.*, **107**(D20)**,** 8287, 10.1029/2001JD000452

Fourrier, J., Morel, P. and Sitbon, P. 1970 Ambient temperature measurements from constant-level balloons. *J. App. Meteorol.*, **9,** 154–157

Haynes, P. and Shuckburgh, E. 2000 Effective diffusivity as a diagnostic of atmospheric transport. 1. Stratosphere. *J. Geophys. Res.*, **105,** 22777–22794

Hertzog, A. and Vial, F. 2001 A study of the dynamics of the equatorial lower stratosphere by use of ultra-long-duration balloons, 2. Gravity waves. *J. Geophys. Res.*, **106,** 22745–22761

Hertzog, A., Vial, F., Dörnbrack, A., Eckermann, S. D., Knudsen, B. M. and Pommereau, J.-P. 2002a In-situ observations of gravity waves and comparisons with numerical simulations during the SOLVE/THESEO 2000 campaign. *J. Geophys. Res.,* **107**(D20)**,** 8292, 10.1029/2001JD001025

Hertzog, A., Vial, F., Mechoso, C. R., Basdevant, C. and Cocquerez, P. 2002b Quasi-Lagrangian measurements in the lower stratospere reveal an energy peak associated with near-inertial waves. *Geophys. Res. Lett.*, **29**(8)**,** 1229, 10.1029/2001GL014083

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Jenne, R. and Joseph, D. 1996 The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.*, **77,** 437–471

Keil, M., Heun, M., Austin, J., Lahoz, W., Lou, G. P. and O'Neill, A. 2001 The use of long-duration balloon data to determine the accuracy of stratospheric analyses and forecasts. *J. Geophys. Res.*, **106,** 10299–10312

Knudsen, B. M. and Carver, G. D. 1994 Accuracy of the isentropic trajectories calculated for the EASOE campaign. *Geophys. Res. Lett.*, **21,** 1199–1202

Knudsen, B. M., Rosen, J. M., Kjome, N. T. and Whitten, A. T. 1996 Comparison of analyzed stratospheric temperatures and calculated trajectories with long-duration balloon data. *J. Geophys. Res.*, **101,** 19137–19145

Knudsen, B. M., Pommereau, J.-P., Garnier, A., Nunez-Pinharanda, M., Denis, L., Letrenne, G., Durand, M. and Rosen, J. M. 2001 Comparison of stratospheric air parcel trajectories based on different meteorological analyses. *J. Geophys. Res.*, **106,** 3415–3424

Knudsen, B. M., Pommereau, J.-P., Garnier, A., Nunes-Pinharanda, M., Denis, L., Newmann, P., Letrenne, G. and Durand, M. 2002 Accuracy of analyzed stratospheric temperatures in the winter Arctic vortex from infrared Montgolfier long-duration balloon flights. 2: Results. *J. Geophys. Res.*, **107**(D20), 4316, 10.1029/2001JD001329

Luers, J. K. and Eskridge, R. E. 1995 Temperature corrections for the VIZ and Vaisala radiosondes. *J. App. Meteor.*, **34,** 1241–1253

Mann, G. W., Davies, S., Carslaw, K., Chipperfield, M. P. and Kettleborough, J. 2002 Polar vortex concentricity as a controlling factor in Arctic denitrification. *J. Geophys. Res.*, **107**(D22), 4663, 10.1029/2002JD002102

Manney, G. L., Swinbank, R., Massie, S. T., Gelman, M. E., Miller, A. J., Nagatani, R., O'Neill, A. and Zurek, R. W. 1996 Comparison of UK Meteorological Office and US National Meteorological Center stratospheric analyses during northern and southern winter. *J. Geophys. Res.*, **101,** 10311–10334

Manney, G. L., Sabutis, J. L., Pawson, S., Santee, M. L., Naujokat, B., Swinbank, R., Gelman, M. E. and Ebisuzaki, W. 2003 Lower stratospheric temperature differences between meteorological analyses in two cold Arctic winters and their impact on polar processing studies. *J. Geophys. Res.*, **108**(D5), 8328, 10.1029/2001JD001149

McNally, A. P., Andersson, E., Kelly, G. and Saunders, R. W. 1999 The use of TOVS/ATOVS radiances in the ECMWF 4D-Var assimilation system. *ECMWF Newsletter*, **83,** 2–7

Moldovan, H., Lott, F. and Teitelbaum, H. 2002 Wave breaking and critical levels for propagating inertio-gravity waves in the lower stratosphere. *Q. J. R. Meteorol. Soc.*, **128,** 713–732

Murphy, D. M. and Gary, B. L. 1995 Mesoscale temperature fluctuations and polar stratospheric clouds. *J. Atmos. Sci.*, **52,** 1753–1760

Pawson, S., Krüger, K., Swinbank, R., Bailey, M. and O'Neill, A. 1999 Intercomparison of two stratospheric analyses: Temperatures relevant to polar stratospheric cloud formation. *J. Geophys. Res.*, **104,** 2041–2050

Peter, T. 1997 Microphysics and heterogeneous chemistry of polar stratospheric clouds. *Ann. Rev. Phys. Chem.*, **48,** 785–822

Pommereau, J.-P., Garnier, A., Knudsen, B. M., Letrenne, G., Durand, M., Cseresnjes, P., Nunes-Pinharanda, M., Denis, L., Vial, F., Hertzog, A. and Cairo, F. 2002 Accuracy of stratospheric meteorological analyses in the winter Arctic vortex from infrared Montgolfier long-duration balloon flights. 1: Measurements. *J. Geophys. Res.*, **107**(D20), 8260, 10.1029/2001JD001379

Pullen, S. and Jones, R. L. 1997 Accuracy of temperatures from UKMO analyses of 1994/95 in the arctic winter stratosphere. *Geophys. Res. Lett.*, **24,** 845–848

Rabier, F., Järvinen, H., Klinker, E., Mahfouf, J.-F. and Simmons, A. 2000 The ECMWF operational implementation of four dimensional variational assimilation. I: Experimental results with simplified physics. *Q. J. R. Meteorol. Soc.*, **126,** 1143–1170

Teitelbaum, H. and Sadourny, R. 1998 The role of planetary waves in the formation of polar stratospheric clouds. *Tellus*, **50A,** 302–312

Trenberth, K. E. and Stepaniak, D. P. 2002 A pathological problem with NCEP reanalyses in the stratosphere. *J. Climate*, **15,** 690–695

Untch, A. and Simmons, A. 1999 Increased stratospheric resolution in the ECMWF forecasting system. *ECMWF Newsletter*, **82,** 2–8

Vial, F., Babiano, A., Basdevant, C., Brioit, B., Legras, B., Sadourny, R., Ovarlez, H., Teitelbaum, H., Cariolle, D., Lefèvre, F., Simon, P., Valero, F., Mechoso, C., Fararra, J., Kelder, H., Camy-Peyret, C., Hart, T. and McIntyre, M. E. 1995 STRATÉOLE: A project to study Antarctic polar vortex dynamics and its impact on ozone chemistry. *Phys. Chem. Earth*, **20,** 83–96

Vial, F., Hertzog, A., Mechoso, C. R., Basdevant, C., Cocquerez, P., Dubourg, V. and Nouel, F. 2001 A study of the dynamics of the equatorial lower stratosphere by use of ultra-long-duration balloons. 1: Planetary scales. *J. Geophys. Res.*, **106,** 22725–22743