Comparing ECMWF high resolution analyses to lidar temperature measurements in the middle atmosphere

Benedikt Ehard[a], Sylvie Malardel[b], Andreas Dörnbrack[a*], Bernd Kaifler[a], Natalie Kaifler[a] and Nils Wedi[b]

[a]*Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany*
[b]*European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom*

[*]Correspondence to: Deutsches Zentrum für Luft und Raumfahrt, Institut für Physik der Atmosphäre, Münchener Strasse 20, 82234 Wessling, Germany (andreas.doernbrack@dlr.de)

Middle atmospheric lidar temperature observations conducted above Sodankylä, Finland (67.4°N, 26.6°E), during December 2015 are compared to two estimates of the atmospheric state computed by the integrated forecast system (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF). The first set corresponds to an hourly sampling of the middle atmosphere by the high-resolution analyses and very short-range forecasts produced by the operational IFS cycle 41r1 at a horizontal resolution of 16 km. The second set is retrieved from the upgraded IFS cycle 41r2 (horizontal resolution at 9 km) which was running in parallel with cycle 41r1 during the validation before it became operational. A remarkable agreement between both IFS data sets and the lidar temperature observations above Sodankylä is found below 45 km altitude. Above 45 km altitude, within the sponge layer of the IFS, both IFS data sets depict lower temperatures than the observations, with the 9 km runs showing the coldest temperatures. Various sensitivity experiments conducted with the IFS are analyzed and compared to the lidar observations to investigate the impact of the different changes implemented in the IFS cycle 41r2. It is found that both the scientific changes and the horizontal resolution upgrade contribute to the colder mesosphere above Sodankylä. The data assimilation seems to amplify this effect even further.

Key Words: Middle atmosphere, ECMWF, Lidar, Gravity waves

1. Introduction

The integrated forecast system (IFS) of the European Centre for Medium Range Weather Forecasts (ECMWF) is widely used not only for everyday weather forecast but also for atmospheric research purposes. In middle atmospheric research, the IFS products are used especially for gravity wave studies. Amongst other purposes, such gravity wave studies utilize the IFS data to estimate the background state of the atmosphere for ray tracing studies (e.g. Preusse *et al.* 2009), characterizing ambient excitation and propagation conditions (e.g. Blum *et al.* 2004; Ehard *et al.* 2017) or for the initialization of mesoscale simulations (e.g. Leutbecher and Volkert 2000; Dörnbrack *et al.* 2002; Plougonven *et al.* 2015; Wagner *et al.* 2017). For all these purposes, however, a thorough validation of the middle atmospheric thermal and dynamical state represented by the IFS products is necessary.

Such a validation was recently undertaken by Le Pichon *et al.* (2015), who compared wintertime ground-based lidar temperature and wind radiometer measurements conducted at the Haute-Provence Observatory (43.9°N, 5.7°E) to two operational ECMWF high-resolution (HRES) atmospheric model analyses defined by the IFS cycles 38r1 and 38r2 at 16 km horizontal resolution and with 91 and 137 levels, respectively. They found good agreement in the stratosphere between the HRES analyses and the observations. However, above 60 km altitude, IFS temperatures were lower than the observations (10 K at 65 km altitude), while the IFS zonal winds were larger than the observations (e.g., 20 m s^{-1} at 65 km altitude).

Other studies examining IFS data in the middle atmosphere focused on the resolved gravity waves directly. One of the first studies comparing stratospheric gravity waves from operational ECMWF analyses with radiosonde observations over the North Atlantic was conducted by Plougonven and Teitelbaum (2003). Their radiosonde observations showed large-scale inertia-gravity waves excited at a jet exit region, which were in agreement with alternating divergence patterns present in the IFS data set. However, they noted that the wave amplitude and horizontal wavelength were not well represented by the model, which they attributed to the coarse vertical and horizontal resolution of their IFS data set. Since then, the horizontal and vertical resolution of the IFS has steadily increased over the years (Bauer *et al.* 2015)

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/qj.3206

and a number of studies have been conducted to compare middle atmosphere observations to the ECMWF HRES analysis.

Wu and Eckermann (2008) and Schroeder *et al.* (2009) compared satellite measurements of gravity wave activity to the ECMWF HRES analysis with 25 km horizontal resolution and 91 vertical levels (top at 0.01 hPa). Both studies found that the satellite measurements and the model show similar global patterns of gravity wave activity in the stratosphere. However, Schroeder *et al.* (2009) noted that gravity wave amplitudes are too low by a factor of 2 as a Rayleigh damping layer started to damp the gravity waves in the IFS above 40 km altitude, see Jablonowski and Williamson (2011).

An examination of gravity waves resolved by the ECMWF HRES analysis at 25 km horizontal resolution and 91 levels was also carried out by Preusse *et al.* (2014), who applied backward ray tracing from 25 km altitude to infer properties and sources of the resolved gravity waves. One of their findings was that the IFS underestimates the short horizontal wavelength gravity waves from convective regions, while it overestimates the long horizontal wavelength gravity waves. Furthermore, they noted that tropical gravity waves resolved by the HRES analysis have, on average, too slow horizontal phase speeds.

Jewtoukoff *et al.* (2015) compared gravity wave momentum fluxes in the lower stratosphere (at 70 hPa) derived from long duration superpressure balloon flights around Antarctica to the operational IFS analysis with a 16 km horizontal resolution and 91 levels. They found that the horizontal structure and intermittency of the momentum fluxes in the IFS agree well with the balloon observations. However, the IFS underestimates the momentum flux values by a factor of 5.

Dörnbrack *et al.* (2017) documented a space-borne lidar observation of mountain-wave induced polar stratospheric clouds over Svalbard. They found a remarkable agreement between the IFS mesoscale temperature anomalies and the observed polar stratospheric clouds at about 25 km altitude.

In a major update of the operational system at the ECMWF on 8 March 2016, the horizontal resolution of the HRES forecast model was upgraded from 16 to 9 km (Malardel and Wedi 2016). In concert with the increase of horizontal resolution from the former cycle 41r1 to new operational IFS cycle 41r2, the resolution of the incremental inner loops in the 4DVAR data assimilation system used for both the ensemble data assimilation (EDA) and the HRES assimilation were upgraded too (Hólm *et al.* 2016). The gain in effective resolution in the operational IFS cycle 41r2 is also attributed to the significant reduction of the numerical filters used in IFS cycle 41r2 and in the preparation of the orography. This was possible thanks to the new "cubic" grid which eliminates the aliasing generated by non-linear operations (Malardel and Wedi 2016). The increase of horizontal effective resolution results in a better representation of small-scale motions, such as gravity waves by the HRES operational system, and an overall improvement of the simulations within the troposphere (Hólm *et al.* 2016).

With the increase of effective resolution of both the model and the orography and the reduction of the numerical filters, the gravity wave activity in the stratosphere is larger in the IFS cycle 41r2 analyses compared to cycle 41r1. However, close to the model top, the gravity waves are still strongly dampened by a sponge layer, mainly to avoid reflection from the model top. Above 1 hPa (about 42 km altitude), gravity wave momentum deposition plays an essential role in the ageostrophic meridional circulation, commonly referred to as the Brewer-Dobson circulation, and is mainly produced by the orographic and non-orographic gravity wave drag parametrizations. However, even the large-scale circulations in the mesosphere are not properly constrained by any of the observations entering the data assimilations system. Therefore, it is of particular interest to compare and validate the state of the ECMWF HRES analyses with additional observation sets covering the upper stratosphere and the mesosphere.

In this study, Rayleigh lidar temperature observations conducted above Sodankylä, Finland (67.4°N, 26.6°E), during December 2015 are compared to the middle atmospheric thermal structure of the two IFS cycles 41r1 and 41r2. From December 2015 until March 2016, both IFS cylces run in parallel, the cycle 41r1 in the operational mode, the upgraded cycle 41r2 as so-called experimental suite (e-suite). Deterministic HRES forecasts and analyses of both cycles were archived at ECMWF.

The meteorological condition during the northern hemispheric winter 2015/16 was characterized by an exceptionally cold lower stratosphere due to reduced planetary wave activity (Matthias *et al.* 2016). Northern Scandinavia was located inside the polar vortex throughout December 2015 and January 2016. Three consecutive minor sudden stratospheric warmings (SSW) occurred on 26 January, 15, and 29 February before a final major SSW around 6 March 2016 (Manney and Lawrence 2016) led to the break-up of the polar stratospheric vortex. Thus, the data from the period analyzed in this study are characteristic of a very cold polar stratosphere, which was nearly undisturbed by planetary waves.

In addition to the mean thermal state validation of the ECMWF HRES analyses, the resolved gravity wave activity estimated from the IFS data set is compared to the lidar observations. To further investigate the effects introduced by the change of effective resolution, sensitivity experiments with the IFS forecast model were performed.

2. The ECMWF HRES system in the middle atmosphere

The IFS model is a global, hydrostatic semi-implicit, semi-Lagrangian numerical weather prediction model. The IFS cycle* 41r1 was operational from 12 May 2015 until 8 March 2016 and utilized a linear grid with a spectral truncation at wavenumber 1279 (T_L1279, see Wedi 2014; Malardel and Wedi 2016, for more explanation about linear and cubic grids) which corresponds to a horizontal resolution of approximately 16 km. In the vertical, 137 levels (L137) ranged from the model top at a pressure level of 0.01 hPa (model level 1, roughly 80 km altitude) down to the surface (model level 137, $\approx$ 10 m altitude). The vertical resolution (Fig. 1) is coarse in the mesosphere ($\approx$ 2 km) and increases with decreasing altitude ($\approx$ 400 m in the lower stratosphere).

The IFS cycle 41r1 was replaced by cycle 41r2 on 8 March 2016. The horizontal resolution of all the different operational applications using the IFS were upgraded (Hólm *et al.* 2016). The HRES analyses and forecasts are computed on a cubic octahedral grid with a resolution of approximately 9 km while the spectral truncation remained at wavenumber 1279 ($T_{Co}1279$, Malardel and Wedi 2016). A large contribution to the gain in effective resolution of the new cycle also comes from the reduced numerical filtering in the model and in the preparation of the physiographic data at the surface.

To avoid wave reflection at the model top, both IFS cycles utilize a sponge layer which starts to dampen wave activity from model level 30 (10 hPa, 28 km, see Fig. 1). The amplitude of the filtering gradually strengthens towards the model top. Between model level 30 and 16 (1 hPa, 42 km), an implicit fourth order spectral diffusion is applied on the spectral coefficients of the vorticity, divergence and temperature. From model level 15, an extra filter with a linear dependence in wavenumber is applied to

*See https://www.ecmwf.int/en/forecasts/documentation-and-support/changes-ecmwf-model for the detailed documentation of the specific IFS cycles.

the divergence field in order to remove gravity wave motions near the model top at all scales.

The sponge layer of the cubic grid ($T_{Co}1279$) differs from the sponge layer of the linear grid (T_L1279). For the cubic grid, the diffusion in the sponge layer is weaker by a factor of 16 for the fourth order diffusion and by a factor of 2 for the extra sponge on divergence compared to the diffusion applied on a linear grid with the same spectral truncation. For technical details implemented in the most recent IFS version, see Polichtchouk *et al.* (2017, p.4).

An analysis of the individual model tendencies shows that the temperature evolution in the upper stratosphere and the mesosphere in the IFS model is governed by the small imbalance between the radiation forcing and the dynamical effects of the resolved large-scale meridional circulation which is mainly controlled by the parametrized subgrid-scale orographic and non-orographic gravity wave drag (Lott and Miller 1997; Beljaars *et al.* 2004; Orr *et al.* 2010). If the sponge layer is removed, the amplitude of the gravity waves which reach the middle atmosphere grows as the density decreases. In this case, the breaking of the resolved gravity waves also contributes to the evolution of the mean state. If the sponge layer is included on the other hand, the small-scale gravity wave activity is mostly damped at the mesoscale. Attenuated gravity waves with larger horizontal wavelengths can still propagate up to the model top, mostly in the winter hemisphere at the edge of the polar vortex. This effect is amplified by the reduction of the numerical filtering introduced in IFS cycle 41r2.

Another peculiarity of the IFS cycle 41r2 concerns the modified EDA. The spectral resolution went from T_L399 (about 50 km resolution in grid-point space) to $T_{Co}639$ (18 km resolution). The spread between the 25 EDA members allows to compute a flow dependent background error variance σ_B for the HRES 4DVAR. In contrast to the cycle 41r1, the flow-dependent information is used to alter the background statistics of the EDA itself in the IFS cycle 41r2. Furthermore, the resolution increase of the analysis increments of the HRES 4DVAR was upgraded from 3 inner loops at T_L255 to 3 inner loops at incrementally higher spectral resolution from T_L255 to T_L319, and, finally, to T_L399 (Hólm *et al.* 2016). These changes, combined with the increased variability in the layer above 1 hPa (due to increased resolution and reduced numerical diffusion) contributed to an important growth of σ_B in the mesosphere. In this layer, the analysis is very weakly constrained by observations. Thus, unrealistic mesospheric flow fields developed in the first months after the start of IFS cycle 41r2. Eventually, the 4DVAR minimization did not converge and model instabilities appeared. In order to give more weight to the background state in the mesosphere analysis an ad-hoc reduction of σ_B in the mesosphere had to be introduced from autumn 2015. Thus, for December 2015, the IFS cycle 41r2 data assimilation system utilized the artificially restrained σ_B above 1 hPa which was later introduced in operation in March 2016. However, further validation of the cycle 41r2 system showed that this "tapering" of the background error statistics had some side effects, with, in particular, the appearance of an unrealistically strong zonal mesospheric jet along the equator. A re-tuning of the background error statistics in the mesosphere later introduced in more recent cycles of the IFS elimitates the mesospheric anomalous equatorial zonal jet and brought back the semi-annual oscillation in the IFS analysis, see Polichtchouk *et al.* (2017).

3. Data processing

The ECMWF IFS data is compared to Rayleigh lidar temperature measurements by the CORAL (Compact Rayleigh Autonomous Lidar) lidar conducted from October 2015 to April 2016 in Sodankylä, Finland (67.4°N, 26.6°E, Kaifler *et al.* 2017). Only

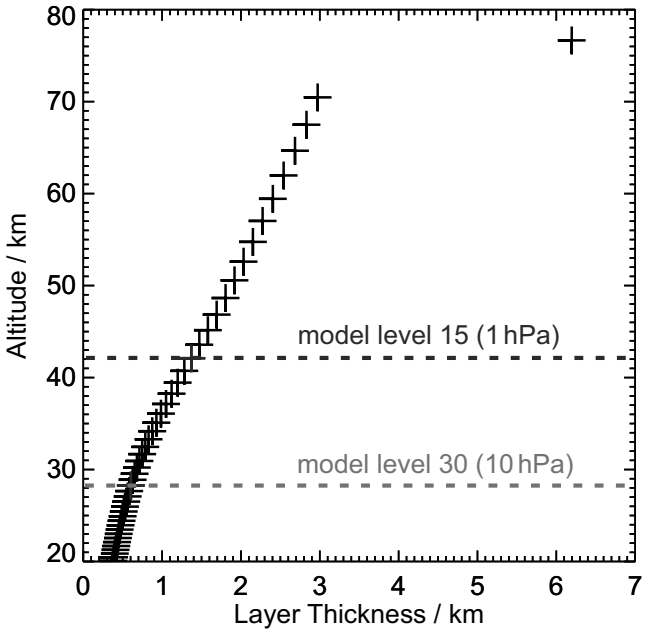

Figure 1. Mean model layer thickness as a function of the mean altitude of the 137 hybrid-pressure levels of the IFS above Sodankylä during December 2015. The black dashed line and grey dashed line mark the model levels 15 and 30 (1 hPa, 10 hPa), respectively.

data from December 2015 is used for the comparison, since the data coverage by the CORAL lidar is best during December 2015. This amounts in total to 176 h of lidar data. Temperature profiles with an effective resolution of 1 km × 1 h are retrieved by hydrostatic integration of the lidar measurements (Hauchecorne and Chanin 1980; Kaifler *et al.* 2015). The uncertainty of the lidar temperature observations is generally smaller than 1 K below 60 km altitude. Above that altitude average uncertainties gradually increase up to 5–6 K at 80 km altitude.

For the comparison, 6 hourly analysis and 1 hourly short-term forecast are combined to create a continuous hourly data set from the IFS. Thus, forecasts with lead times +1, +2, +3, +4, +5, +7, +8, +9, +10, +11 h from the 00 UTC and 12 UTC runs are taken to fill the times between the analysis times 00, 06, 12, and 18 UTC. To enable a direct comparison, the IFS temperature data is interpolated horizontally to the position of the lidar. Furthermore, the IFS data is interpolated on a regular vertical grid with a resolution of 500 m, ranging from 0.5 to 75 km altitude.

For the purposes of this study, the comparison between the IFS and the lidar is limited to the altitude range 30 to 70 km. The lower limit is chosen due to the likely presence of aerosols below this altitude, resulting in unreliable lidar temperatures. The upper limit is restricted to 70 km as the IFS resolution near the model top is very coarse with only two levels in the layer 70-80 km, see Fig. 1.

In order to extract resolved gravity waves from the lidar observations and the IFS data, we follow the suggestion of Ehard *et al.* (2015): A fifth order Butterworth high-pass filter with a cutoff wavelength of 15 km is applied to individual vertical temperature profiles. The mean gravity wave activity is characterized by the gravity wave potential energy density per unit mass

$$E_p = \frac{1}{2}\frac{g^2}{N^2}\overline{\left(\frac{T'}{T_0}\right)^2} \tag{1}$$

$$\text{with} \quad N^2 = \frac{g}{T_0}\left(\frac{dT_0}{dz} + \frac{g}{c_p}\right), \tag{2}$$

with the temperature fluctuations T', the background temperature T_0, the Brunt-Väisälä frequency N, the gravitational constant g and the heat capacity of dry air under constant pressure c_p. If

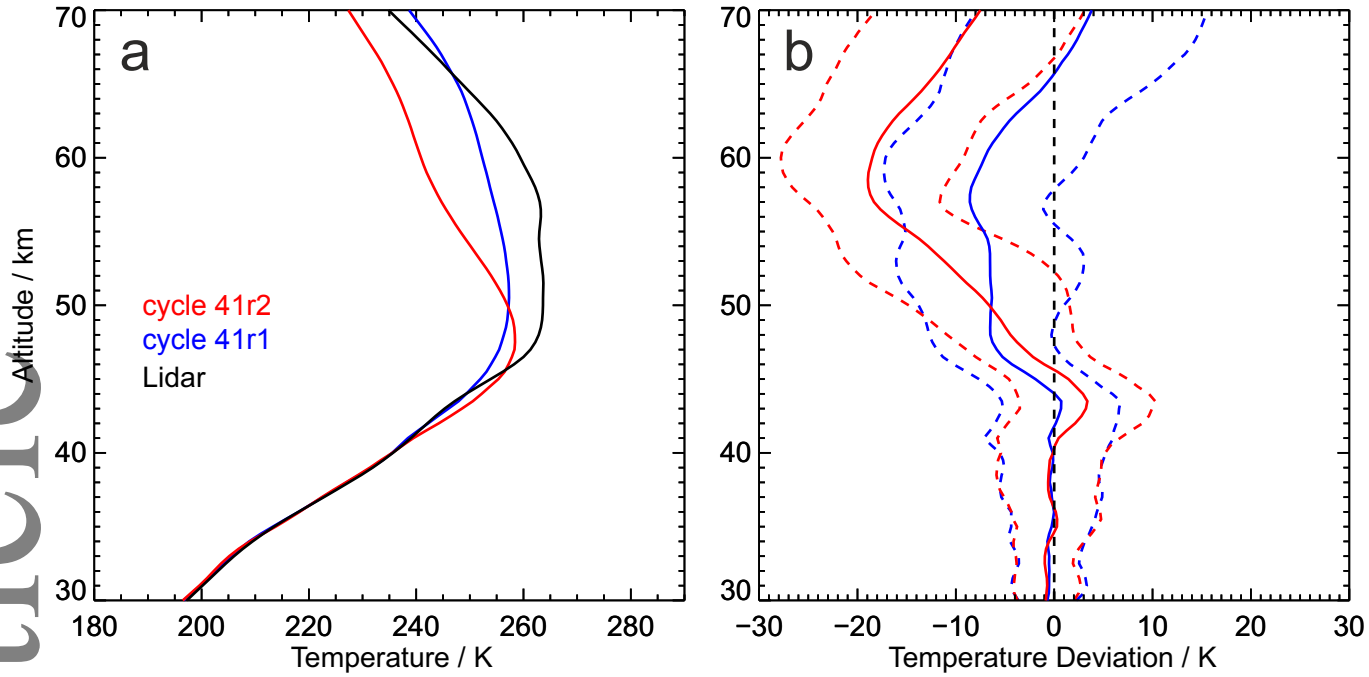

Figure 2. Left panel: Mean temperature profiles above Sodankylä during December 2015. Red line: IFS cycle 41r2 ($T_{Co}1279$); blue line: IFS cycle 41r1 (T_L1279). Black line: measurements by the CORAL lidar. Right panel: Temperature deviations (solid lines) and corresponding standard deviations (dashed lines) of the IFS cycles 41r2 (red lines) and 41r1 (blue lines) from the CORAL measurements. Only IFS profiles with simultaneous lidar measurements were used to calculate the respective profiles.

a single monochromatic wave propagates purely vertically in an atmosphere with a constant stratification and uniform background wind, E_p increases exponentially with altitude. Under realistic atmospheric conditions, the wave action $A = E/\hat{\omega}$, with the total wave energy per volume E and the intrinsic frequency $\hat{\omega}$ should be used instead, e.g., Sutherland (2010, pp. 170). In this sense, the slope of E_p-profiles can only be used with caution to investigate linear, non-dissipative vertical wave propagation, see section 2.1 of Ehard *et al.* (2017).

4. Evaluating the high-resolution IFS

Figure 2a shows the mean temperature profiles of the IFS cycles 41r1 and 41r2 in comparison to the CORAL lidar data during December 2015. Here, only IFS profiles with simultaneous lidar measurements were used. The CORAL temperature profile depicts a very broad stratopause between 46 and 58 km altitude with a mean stratopause temperature of 263 K. IFS cycle 41r1 simulates a mean stratopause at 50 km altitude with a slightly lower temperature of 257 K. The IFS cycle 41r2 simulates a very sharp stratopause at 48 km altitude with a temperature of 258 K. Above the stratopause, all three data sets exhibit a negative lapse rate in the mesosphere.

Figure 2 reveals a remarkable agreement between IFS cycle 41r1 and the lidar temperatures at Sodankylä up to an altitude of 44 km (solid blue lines). Within the stratopause region, the cycle 41r1 simulates temperatures which are too low by approximately 7 K, while above 65 km altitude the simulated temperatures are slightly too high. The cycle 41r2 (solid red lines, Fig. 2) shows a very good agreement with the lidar below 41 km altitude. At 43 km altitude the cycle 41r2 temperatures are slightly larger (3 K). Above this altitude, the negative temperature deviation increases in magnitude and the IFS cycle 41r2 reaches a maximum value of -19 K around 60 km altitude (Fig. 2b). Above 60 km altitude, this deviation decreases in magnitude, reaching -8 K at 70 km altitude. For both sets of IFS data, standard deviations of the differences are small (dashed lines Fig. 2b) in in the stratosphere (4 K) and increase in the mesosphere (10 K).

Comparing the mean E_p-profiles of the two IFS versions to the CORAL measurements (Fig. 3a), a similar increase of E_p with increasing altitude between 30 and 40 km altitude can be seen in all three data sets. The IFS data agree quantitatively very well with the 1 km × 1 h E_p-values derived from CORAL measurements (black dashed line in Fig. 3a). The effect of the time and space averaging is documented by three additional lines. Stronger vertical averaging (3 km compared to 1 km) decreases the lidar-derived E_p whereas the a shorter time averaging shifts the curves to higher E_p-values. Above about 40 km altitude, both IFS data sets show rapidly decreasing E_p values. As expected, cycle 41r1 shows an even stronger decrease of E_p than cycle 41r2. The enhancement of gravity wave acticity is consistent with the reduced strength of the numerical diffusion applied within the sponge layer.

Both IFS data sets accurately reproduce the general temporal development of the mean E_p between 30 and 40 km altitude (Fig. 3b). In particular, they capture very well the period with enhanced E_p during December 2015 as well as the strong decrease of E_p during January 2016 when Sodankylä was located in the centre of the polar vortex (Manney and Lawrence 2016). The overall temporal correlation between colocated data points is about 0.80 for both cycles (52 data points in each data set). Additionally, individual events such as the sudden increase of E_p around the minor SSW on 15 February 2016 are well captured by both cycles.

5. Sensitivity experiments with the IFS

A series of sensitivity experiments were conducted to better evaluate the influence of the model resolution versus the importance of the initial conditions or the model cycle. In this section, the lidar data are compared with a series of 4 lead day forecasts starting from HRES analyses which were produced either by IFS cycle 41r1 or by cycle 41r2 running in parallel in the e-suite. The evolution of the kinetic energy spectra in the mesosphere (not shown) reveals that the small-scale characteristics of a given model configuration are re-created in less than a day. The feedback on the large scale may be much slower, but after four days, a general tendency is usually visible.

For the sensitivity experiments, the week from 10 to 17 December 2015 was selected, since the data coverage by the

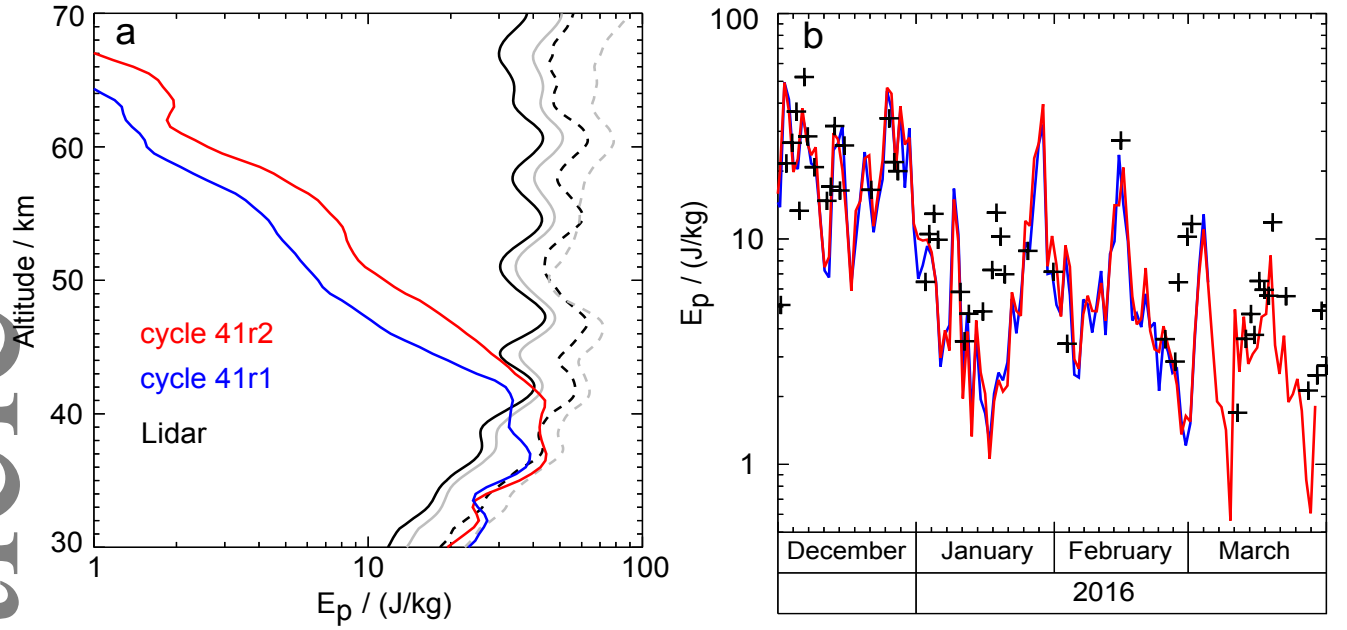

Figure 3. Left panel: Mean vertical profiles of the gravity wave potential energy density E_p above Sodankylä, Finland, during December 2015 derived from CORAL measurements (black and grey lines), the IFS cycle 41r1 (blue line) and the IFS cycle 41r2 (red line). The lidar profiles are differently averaged: black (grey) lines are 60 (10) min avarages; solid (dashed) lines refer to 3 km (1 km) averages. Only IFS profiles with simultaneous lidar measurements were used. Right panel: Daily mean E_p between 30 and 40 km altitude over Sodankylä derived from 1 km × 1 h CORAL measurements (black crosses), the IFS cycle 41r1 (T_L1279, blue line) and the IFS cycle 41r2 ($T_{Co}1279$, red line). Here, all IFS profiles above Sodankylä from December 2015 to March 2016 were used.

Table 1. Nomenclature of the sensitivity runs. All sensitivity runs were conducted with the IFS cycle 41r2.

	horizontal grid	
analysis	$T_{Co}1279$	T_L1279
41r2	S1	S2
41r1	S3	S4

lidar is largest during this period. In total 64 h of lidar temperature observations are available within this time frame. All sensitivity runs were conducted with the IFS cycle 41r2 installed to a new supercomputer at the ECMWF. Hourly outputs from +97 to +108 h lead time of simulations initialized from the respective 00 UTC or 12 UTC analyses were combined to form a continuous hourly data set. As in the previous comparison, the IFS data were interpolated to the geographical lidar position and only IFS profiles with coinciding lidar measurements are used in the comparison.

Four sensitivity runs were carried out. We varied the horizontal grid (T_L1279 versus $T_{Co}1279$) and started the sensitivity runs from the two available analyses (IFS cycle 41r1 versus cycle 41r2 analysis) for each grid setting. The names of the four runs are given in Table 1. As the sensitivity runs were conducted with the same IFS cycle 41r2, the results of the sensitivity run S1 should be nearly identical to results of the IFS cycle 41r2 presented in Fig. 2.

The mean temperature profiles of the four sensitivity runs S1 to S4 in comparison to the CORAL measurements and the 4 lead day forecasts of IFS cycle 41r1 (blue dashed line) and the IFS cycle 41r2 running in the e-suite (red dashed line) are shown in Figure 4. Below 45 km altitude, all simulations agree well with the CORAL measurements. Above 50 km altitude, all IFS profiles, regardless of their resolution and initial condition, have a similar mean lapse rate and exhibit significantly lower temperatures compared to the lidar measurements during the selected time period. The maximum offset between the temperature profiles of the four

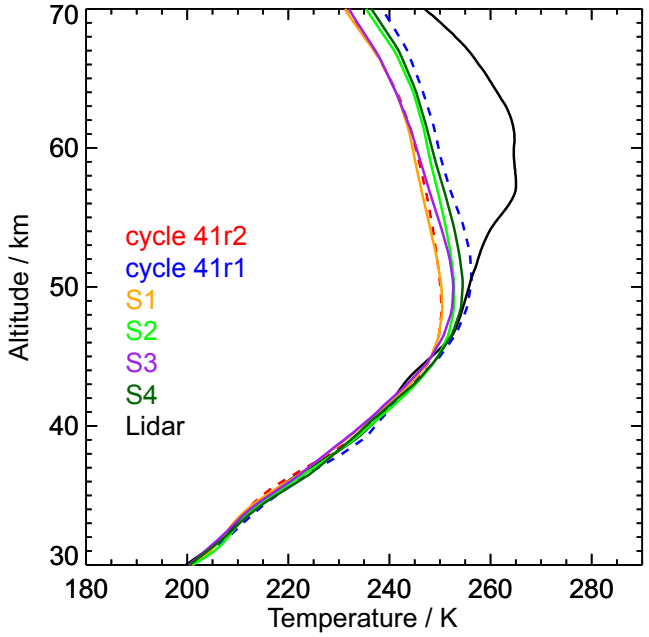

Figure 4. Mean temperature profiles above Sodankylä from 10 to 17 December 2015. Black line: CORAL measurements; dashed red and blue lines: 4 day forecasts from IFS cycles 41r2 and 41r1, respectively. Colored solid lines: 4 day forecasts of the sensitivity runs S1 to S4, see Table 1. Only IFS profiles with simultaneous lidar measurements are used.

sensitivity runs is about 8 K. This value is about a factor of 2 smaller than the maximum temperature difference between the IFS cycles 41r1 and 41r2 as shown in Figure 2b. This suggests that the data assimilation configuration introduced with the IFS cycle 41r2 amplifies the cooling of the polar mesosphere.

As expected, the IFS cycle 41r2 e-suite run (red dashed line) and the sensitivity experiment S1 (orange line) yield almost identical mean temperature profiles. In contrast, the sensitivity experiment S4 (dark green line) exhibits a mean temperature which is 2 to 3 K lower than the mean temperature of the IFS cycle 41r1 forecast (blue dashed line). Although both simulations

run on the same horizontal grid and were initialized with the same analyses, the additional cooling must be again attributed to the science changes in the data assimilation configuration of the IFS cycle 41r2. Modified ozone climate files implemented in this cycle suggest another possible cause of the colder temperatures found in simulations S1 to S4. Other tuning of the physics parametrization, in particular of the non-orographic gravity wave drag may also contribute to this additional cooling near the model top.

All the three runs using the higher $T_{Co}1279$ resolution (red, orange and purple lines in Fig. 4) are about 5 K cooler than the lower resolved T_L1279 profiles (blue, light and dark green lines in Fig. 4) above 50 km. After four days, runs starting from a warmer initial condition (analysis of cycle 41r1) adjust to the cooler profile which is characteristic for the high resolution $T_{Co}1279$ runs. This clearly indicates that the extra cooling in the mesosphere above northern Finland brought out by the comparison between IFS cycles 41r1 and 41r2 comes mainly from the change of resolution.

It is difficult to isolate the processes responsible for this prevailing cooling of the polar mesosphere in the $T_{Co}1279$ runs. Recent attempts to assess systematically the state of the middle atmosphere circulation and its dependency on model physics and numerics on seasonal timescales are provided by Polichtchouk *et al.* (2017). There, a persistent global warm bias of 20 K in the mesosphere and upper stratosphere against MLS and ERA-Interim were found. It must be noted that our observed negative temperature deviations are probably not representative for the mean polar middle atmosphere as they were taken at a particular high-latitude location and during a time period when Sodankylä was inside of a strong polar vortex with anomalous cold stratospheric temperatures, see Dörnbrack *et al.* (2017, Fig. 3).

6. Summary

The temperature profiles up to 45 km altitude of the operational IFS cycle 41r1 and the upgraded IFS cycle 41r2 running in parallel as e-suite show a remarkable agreement with lidar observations above Sodankylä, Finland, during December 2015. At higher altitudes, the polar mesosphere simulated by the ECMWF IFS is colder compared to the lidar measurements, which enhanced temperature deviation occurring within the IFS cycle 41r2.

Hydrostatic gravity waves are well resolved by the IFS up to 45 km altitude (approximately 1 hPa). Above 45 km altitude, the gravity wave activity is dampened by the diffusion in the sponge layer which is designed to eliminate the divergence motion at all wavenumbers. Furthermore, it was shown that the IFS is capable of reproducing the overall temporal evolution of the observed gravity wave activity between 30 and 40 km altitude above Northern Finland.

Sensitivity runs conducted with the IFS cycle 41r2 suggest that both the scientific changes between cycle 41r1 and 41r2 and the resolution upgrade from T_L1279 to $T_{Co}1279$ contribute to the colder polar mesosphere detected in this study. The data assimilation system introduced with the new cycle seems to amplify this negative temperature deviations even further.

Middle atmospheric lidar measurements are very valuable observations for the validation of the ECMWF IFS model at higher altitudes where none of the data available on the Global Telecommunication System permit a reliable anchorage of the atmospheric state in the IFS data assimilation system. Although lidar observations are only conducted at one point and they are limited in time, the high accuracy of the lidar data is an important asset for model validation.

The comparison between IFS data and the lidar observations shows the difficulty of modeling the atmospheric state near the model top, in a layer where the vertical resolution is very coarse and where unphysical numerical filters are necessary to eliminate reflection from the top boundary. Work is in progress at ECMWF to improve the mean representation of the HRES and re-analysis data sets in the mesosphere (Polichtchouk *et al.* 2017). However, most of the resolved variability in the upper mesosphere will probably still have to be filtered out as long as the physics described by the model (in particular the radiation parametrization) does not contain the processes which would be necessary to move the top model level to a higher altitude.

Acknowledgements

This work was supported by the project "Processes and Climatology of Gravity Waves" (PACOG, grant RA 1400/6-1) in the framework of the research unit "Multiscale Dynamics of Gravity Waves" (MS-GWaves) funded by the German Research Foundation.

References

Bauer P, Thorpe A, Brunet G. 2015. The quiet revolution of numerical weather prediction. *Nature* **525**: 47–55, doi:10.1038/nature14956, URL http://dx.doi.org/10.1038/nature14956.

Beljaars ACM, Brown AR, Wood N. 2004. A new parametrization of turbulent orographic form drag. *Quart. J. Roy. Meteor. Soc.* **130**(599): 1327–1347, doi:10.1256/qj.03.73.

Blum U, Fricke KH, Baumgarten G, Schöch A. 2004. Simultaneous lidar observations of temperatures and waves in the polar middle atmosphere on the east and west side of the Scandinavian mountains: A case study on 19/20 January 2003. *Atmos. Chem. Phys.* **4**: 809–816, doi:10.5194/acpd-4-969-2004.

Dörnbrack A, Birner T, Fix A, Flentje H, Meister A, Schmid H, Browell EV, Mahoney MJ. 2002. Evidence for intertia gravity waves forming polar stratospheric clouds over Scandinavia. *J. Geoph. Res.* **107**: 8287–8295, doi:10.1029/2001JD000452.

Dörnbrack A, Gisinger S, Pitts MC, Poole LR, Maturilli M. 2017. Multilevel cloud structures over Svalbard. *Mon. Wea. Rev.* doi: 10.1175/mwr-d-16-0214.1.

Ehard B, Kaifler B, Dörnbrack A, Preusse P, Eckermann S, Bramberger M, Gisinger S, Kaifler N, Liley B, Wagner J, Rapp M. 2017. Horizontal propagation of large-amplitude mountain waves into the polar night jet. *J. Geophys. Res.* **122**(3): 1423–1436, doi:10.1002/2016JD025621.

Ehard B, Kaifler B, Kaifler N, Rapp M. 2015. Evaluation of methods for gravity wave extraction from middle-atmospheric lidar temperature measurements. *Atmos. Meas. Tech.* **8**(11): 4645–4655, doi:10.5194/amt-8-4645-2015.

Hauchecorne A, Chanin M. 1980. Density and temperature profiles obtained by lidar between 35 and 70 km. *Geophys. Res. Lett.* **7**: 565–568, doi:10.1029/GL007i008p00565.

Hólm E, Forbes R, Lang S, Magnusson L, Malardel S. 2016. New model cycle brings higher resolution. *ECMWF Newsletter* **147**: 14–19, URL http://www.ecmwf.int/sites/default/files/elibrary/2

Jablonowski C, Williamson DL. 2011. The pros and cons of diffusion, filters and fixers in atmospheric general circulation models. In: *Numerical Techniques for Global Atmospheric Models*, Lauritzen P, Jablonowski C, Taylor M, Nair R (eds), Springer Berlin Heidelberg, pp. 381–493.

Jewtoukoff V, Hertzog A, Plougonven R, de la Camara A, Lott F. 2015. Comparison of gravity waves in the southern hemisphere derived from balloon observations and the ECMWF analyses. *J. Atmos. Sci.* **72**(9): 3449–3468, doi:10.1175/JAS-D-14-0324.1.

Kaifler B, Kaifler N, Ehard B, Dörnbrack A, Rapp M, Fritts D. 2015. Influences of source conditions on mountain wave penetration into the stratosphere and mesosphere. *Geophys. Res. Lett.* **42**: 9488–9494, doi:10.1002/2015GL066465.

Kaifler N, Kaifler B, Ehard B, Gisinger S, Dörnbrack A, Rapp M, Kivi R, Kozlovsky A, Lester M, Liley B. 2017. Observational indications of downward-propagating gravity waves in middle atmosphere lidar data. *J. Atmos. Solar-Terr. Phys.* doi:10.1016/j.jastp.2017.03.003.

Le Pichon A, Assink J, Heinrich P, Blanc E, Charlton-Perez A, Lee C, Keckhut P, Hauchecorne A, Rüfenacht R, Kämpfer N, Drob D, Smets P, Evers L, Ceranna L, Pilger C, Ross O, Claud C. 2015. Comparison of co-located independent ground-based middle atmospheric wind and temperature measurements with numerical weather prediction models. *J. Geophys. Res. Atmos.* **120**(16): 8318–8331, doi:10.1002/2015JD023273.

Leutbecher M, Volkert H. 2000. The propagation of mountain waves into the stratosphere: quantitative evaluation of three-dimensional simulations. *J. Atmos. Sci.* **57**: 3090–3108.

Lott F, Miller MJ. 1997. A new subgrid-scale orographic drag parametrization: Its formulation and testing. *Quart. J. Roy. Meteor. Soc.* **123**(537): 101–127, doi:10.1002/qj.49712353704.

Malardel S, Wedi NP. 2016. How does subgrid-scale parametrization influence nonlinear spectral energy fluxes in global NWP models? *J. Geophys. Res.* **121**(10): 5395–5410, doi:10.1002/2015JD023970.

Manney GL, Lawrence ZD. 2016. The major stratospheric final warming in 2016: dispersal of vortex air and termination of Arctic chemical ozone loss. *Atmos. Chem. Phys.* **16**(23): 15 371–15 396, doi:10.5194/acp-16-15371-2016.

Matthias V, Dörnbrack A, Stober G. 2016. The extraordinarily strong and cold polar vortex in the early northern winter 2015/2016. *Geophys. Res. Lett.* **43**(23): 12,287–12,294, doi:10.1002/2016gl071676.

Orr A, Bechtold P, Scinocca J, Ern M, Janiskova M. 2010. Improved Middle Atmosphere Climate and Forecasts in the ECMWF Model through a Nonorographic Gravity Wave Drag Parameterization. *J. Climate* **23**(22): 5905–5926, doi:10.1175/2010jcli3490.1.

Plougonven R, Hertzog A, Alexander MJ. 2015. Case studies of nonorographic gravity waves over the Southern Ocean emphasize the role of moisture. *J. Geophys. Res. Atmos.* **120**(21): 9488–9494, doi:10.1002/2014JD022332.

Plougonven R, Teitelbaum H. 2003. Comparison of a large-scale inertia-gravity wave as seen in the ECMWF analyses and from radiosondes. *Geophys. Res. Lett* **30**(18): 1954, doi:10.1029/2003GL017716.

Polichtchouk I, Hogan R, Shepherd T, Bechtold P, Stockdale T, Malardel S, Lock SJ, Magnusson L. 2017. What influences the middle atmosphere circulation in the IFS? *ECMWF Technical Memorandum* **809**.

Preusse P, Eckermann S, Ern M, Oberheide J, Picard R, Roble R, Riese G, Russell III J, Mlynczak M. 2009. Global ray tracing simulations of the SABER gravity wave climatology. *J. Geophys. Res.* **114**(D8), doi:10.1029/2008JD011214.

Preusse P, Ern M, Bechtold P, Eckermann S, Kalisch S, Trinh Q, Riese M. 2014. Characteristics of gravity waves resolved by ECMWF. *Atmos. Chem. Phys.* **14**: 10 483–10 508, doi:10.5194/acp-14-10483-2014.

Schroeder S, Preusse P, Ern M, Riese M. 2009. Gravity waves resolved in ECMWF and measured by SABER. *Geophys. Res. Lett.* **36**: L10 805, doi:10.1029/2008GL037054.

Sutherland B. 2010. *Internal Gravity Waves*. Cambridge University Press, 1st edn.

Wagner J, Dörnbrack A, Rapp M, Gisinger S, Ehard B, Bramberger M, Witschas B, Chouza F, Rahm S, Mallaun C, Baumgarten G, Hoor P. 2017. Observed versus simulated mountain waves over Scandinavia – improvement by enhanced model resolution? *Atmos. Chem. Phys.* doi:10.5194/acp-2016-765.

Wedi NP. 2014. Increasing horizontal resolution in numerical weather prediction and climate simulations: illusion or panacea? *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **372**(2018), doi:10.1098/rsta.2013.0289.

Wu D, Eckermann S. 2008. Global Gravity Wave Variances from Aura MLS: Characteristics and Interpretation. *J. Atmos. Sci.* **65**(12): 3695–3718, doi:10.1175/2008JAS2489.1.