Challenge of modelling GLORIA observations of UT/LMS trace gas and cloud distributions at high latitudes: a case study with state-of-the-art models

Florian Haenel[1], Wolfgang Woiwode[1], Jennifer Buchmüller[2], Felix Friedl-Vallon[1], Michael Höpfner[1], Sören Johansson[1], Farahnaz Khosrawi[1], Oliver Kirner[2], Anne Kleinert[1], Hermann Oelhaf[1], Johannes Orphal[1], Roland Ruhnke[1], Björn-Martin Sinnhuber[1], Jörn Ungermann[3], Michael Weimer[1,2,*] and Peter Braesicke[1]

[1]Institute of Meteorology and Climate Research, Karlsruhe Institute of Technology, Karlsruhe, Germany
[2]Steinbuch Centre for Computing, Karlsruhe Institute of Technology, Karlsruhe, Germany
[3]Institute of Energy and Climate Research – Stratosphere (IEK-7), Forschungszentrum Jülich, Jülich, Germany
[*]now at: Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA

Correspondence to: Florian Haenel (florian.haenel@kit.edu)

Abstract. Water vapour and ozone are important for the thermal and radiative balance of the upper troposphere (UT) and lowermost stratosphere (LMS). Both species are modulated by transport processes. Chemical and microphysical processes affect them differently. Thus, representing the different processes and their interactions is a challenging task for dynamical cores, chemical modules and microphysical parameterisations of state-of-the-art atmospheric model components. To test and improve the models, high resolution measurements of the UT/LMS are required. Here, we use measurements taken in a challenging case study by the GLORIA (Gimballed Limb Observer for Radiance Imaging of the Atmosphere) instrument on HALO. The German research aircraft HALO (High Altitude and LOng range research aircraft) performed a research flight on 26 February 2016, which covered deeply subsided air masses of the aged 2015/16 Arctic vortex, high-latitude LMS air masses, a highly textured troposphere-to-stratosphere exchange mixing region, and high-altitude cirrus clouds. Therefore, it provides a multifaceted case study for comparing GLORIA observations with state-of-the-art atmospheric model simulations in a complex UT/LMS region at a late stage of the Arctic winter 2015/16.

Using GLORIA observations in this manifold scenario, we test the ability of the numerical weather prediction (NWP)-model ICON (ICOsahedral Nonhydrostatic) with the extension ART (Aerosols and Reactive Trace gases) and the chemistry-climate model (CCM) EMAC (ECHAM5/MESSy Atmospheric Chemistry) to model the UT/LMS composition of water vapour (H_2O), ozone (O_3), nitric acid (HNO_3) and clouds. Within the scales resolved by the respective model, we find good overall agreement of both models with GLORIA. The applied high-resolution ICON-ART setup involving a R2B7 nest (local grid refinement with a horizontal resolution of about 20 km), covering the HALO flight region, reproduces mesoscale dynamical structures well. An observed troposphere-to-stratosphere exchange connected to an occluded Icelandic low is clearly reproduced by the model. Given the lower resolution (T106) of the nudged simulation of the EMAC model, we find that this model also

reproduces these features well. Overall, trace gas mixing ratios simulated by both models are in a realistic range, and major cloud systems observed by GLORIA are mostly reproduced. However, we find both models to be affected by a well-known systematic moist-bias in the LMS. Further biases are diagnosed in the ICON-ART O_3, EMAC H_2O and EMAC HNO_3 distributions. Finally, we use sensitivity simulations to investigate (i) short-term cirrus cloud impacts on the H_2O distribution

5 (ICON-ART), (ii) the overall impact of polar winter chemistry and microphysical processing on O_3 and HNO_3 (ICON-ART/EMAC), (iii) the impact of the model resolution on simulated parameters (EMAC), and (iv) consequences of scavenging processes by cloud particles (EMAC). We find that changing of the horizontal model resolution results in notable systematic changes for all species in the LMS, while scavenging processes play only a role in case of HNO_3. We need to understand the representativeness of our results. However, this is a unique opportunity to characterise model biases that potentially affect

10 forecasts and projection (adversely), and to discover deficits and define paths for further model improvements.

1 Introduction

Trace gas composition, in particular the vertical distributions of greenhouse gases, and clouds play an important role in the thermal and radiative budget of the upper troposphere/lowermost stratosphere (UT/LMS) (Riese et al., 2012; Hartmann et al., 2013). Stratospheric and, particularly, lowermost stratospheric water vapour has been identified to be an important driver in

15 decadal global surface climate change (Forster and Shine, 2002; Solomon et al., 2010). Also, changes in stratospheric ozone are well known to affect temperature trends and radiative forcing (Forster and Shine, 1997). In the lower stratosphere, ozone depletion is a major contributor to its negative temperature trend. There is also a significant spread in these trends among models while perturbating ozone and other greenhouse gas abundances. Explanations for such differences can be the different responses of individual radiation schemes and different sensitivities in the dynamical forcing in the models to changes in trace

20 gases (Shine et al., 2003). Lowermost stratospheric water vapour distributions show hemispheric differences, thus requiring knowledge on hemispheric and latitudinal distributions and change for accurate climate projections (Kelly et al., 1991; Rosenlof et al., 1997; Pan et al., 1997).

The LMS is the lowest compartment of the stratosphere situated between the local tropopause and the 380 K isentropic level (e.g. Werner et al., 2010). On the winter hemisphere, its composition is mainly affected by air mass contributions from the

25 polar winter vortex, the mid-latitude stratosphere, and the troposphere. While air masses in the polar winter vortex are mostly isolated from the surrounding stratosphere, LMS air masses at the bottom of the polar vortex can be affected significantly by interactions with air masses from lower latitudes (Krause et al., 2018).

Exchange processes including quasi-isentropic and cross-isentropic exchange occur often in the vicinity of jet streams (Holton et al., 1995; Gettelman et al., 2011). They can be accompanied by different kinds of tropopause folds and modulate the trace

30 gas composition of the UT/LMS. Irreversible fluxes between the UT and the LMS can occur in either direction – from stratosphere-to-troposphere and from troposphere-to-stratosphere. Generally, the dominating flux in the extratropics is directed towards the troposphere. Such exchange processes and their effects have been investigated by numerous field observations

(e.g. Ray et al., 1999; Hoor et al., 2002, 2005; Bönisch et al., 2009; Krause et al., 2018) and by many theoretical and modelling studies (e.g. Meloen et al., 2003; Stohl et al., 2003 and references therein).

Cirrus clouds are one of the least understood factors modulating climate change and affecting the composition of the UT/LMS (e.g. Schiller et al., 2008; Barahona and Nenes, 2009). Cirrus clouds absorb upwelling infrared light and reflect sunlight back
5 to space and thereby affect the radiative budget and thus the thermal structure of the tropopause region. Sedimentation of cirrus cloud ice particles redistribute water vertically and changes eventually the water vapour profile. Furthermore, the ice particles are capable of trapping nitric acid and other trace gases (Popp et al., 2004; Voigt et al., 2006; Krämer et al., 2008; Kärcher et al., 2009). Moreover, vertical distributions of H_2O and HNO_3 altered by cirrus cloud processing might affect the availability of reactive nitrogen oxides (NO_x) and hydroxyl radicals, which are again important factors affecting the local concentrations
10 of ozone and methane (Kelly et al., 1991; Krämer et al., 2008; Schiller et al., 2008).

Nowadays, numerical weather prediction and chemistry-climate models (NWPs and CCMs) are capable of resolving the UT/LMS, mesoscale dynamics and cloud processes (in parts) explicitly and by using parameterisations ranging from low to high complexity. Examples of such models are the models ICON (ICOsahedral Nonhydrostatic, see Zängl et al., 2015) with the extension ART (Aerosols and Reactive Trace gases, see Rieger et al., 2015 and Schröter et al., 2018) and EMAC
15 (ECHAM5/MESSy Atmospheric Chemistry, see Jöckel et al., 2006, 2010, 2015 and Roeckner et al., 2006). However, accurate simulations of UT/LMS composition, dynamics and cirrus clouds (and their interactions) remain a challenge and are important building blocks for reliable weather forecasting and climate projections. In particular, LMS water vapour is known to be affected by significant systematic errors in model simulations (e.g. Stenke et al., 2008).

The exceptionally cold Arctic winter 2015/16 was characterised by a stable polar vortex and low temperatures in the UT/LMS
20 region (Matthias et al., 2016). While the winter was the coldest on record from December to early February, complex dynamical processes and a major final stratospheric warming in early March ended the cold phase and resulted in a vortex split in mid-March (Manney and Lawrence, 2016). In the same winter, airborne observations in the framework of the combined POLSTRACC (POLar STRAtosphere in a Changing Climate), GW-LCYCLE (Gravity Wave Life Cycle Experiment) II and SALSA (Seasonality of Air mass transport and origin in the Lowermost Stratosphere using the HALO Aircraft) (PGS) field
25 campaign probed the Arctic UT/LMS region in the period from December 2015 to March 2016 (Oelhaf et al., 2019). During PGS, the GLORIA (Gimballed Limb Observer for Radiance Imaging of the Atmosphere) instrument (Friedl-Vallon et al., 2014; Riese et al., 2014) was deployed on-board the German HALO (High Altitude and LOng Range Research Aircraft). From the GLORIA limb-imaging observations, vertical distributions of temperature, trace gases and clouds are derived and allow detailed model comparisons (e.g. Khosrawi et al., 2017; Braun et al., 2019; Johansson et al., 2019).

30 During the research flight on 26 February 2016 (PGS 14), GLORIA probed subsided LMS air masses of the aged 2015/16 polar vortex in high latitudes, a highly textured troposphere-stratosphere exchange region, and high-altitude cirrus clouds across a long transect spanning from Scandinavia over Greenland to Canada. Here, we use the GLORIA observations during this flight to test the capabilities of EMAC and ICON-ART of modelling mesoscale H_2O, O_3 and HNO_3 distributions and cirrus clouds and to reveal discrepancies and deviations that might be related to (systematic) biases in the trace gas distributions. We

particularly focus on a troposphere-stratosphere exchange region associated with an occluded Icelandic low. Finally, we use sensitivity simulations to investigate (i) short-term cirrus cloud impacts on the H_2O-distribution (ICON-ART), (ii) the impact of polar winter chemistry and microphysical processing on O_3 and HNO_3 (ICON-ART/EMAC), (iii) the impact of model resolution on simulated parameters (EMAC), and (iv) consequences of scavenging processes by cloud particles (EMAC).

5   In Section 2, we introduce our observations, models and diagnostics. An overview of the meteorological situation and the GLORIA observations during PGS 14 is provided in Sect. 3. In Section 4, the 2-dimensional vertical cross sections of modelled cloud and trace gas distributions are compared with the GLORIA observations, discrepancies are diagnosed and investigated, and sensitivity experiments with the models are presented. The results are summarised and discussed in Sect. 5.

2 Data and diagnostics

10   In the following, the characteristics of the GLORIA observations, the used model setups, and the applied diagnostics are introduced. An overview on the cloud and trace gas products used is provided in Tables 1 and 2.

2.1 GLORIA observations

The GLORIA data used here were measured during the HALO flight PGS 14 on 26 February 2016. PGS 14 started in Kiruna, northern Sweden, and covered the Arctic Sea, Greenland, and Eastern Canada (Fig. 1b). GLORIA is a passive infrared limb-

15   imaging spectrometer deployed on-board high-altitude aircrafts (Friedl-Vallon et al., 2014; Riese et al., 2014). GLORIA uses 128 vertical times 48 horizontal pixels of a Mercury Cadmium Telluride (HgCdTe) detector coupled to an interferometer to measure thermal radiation of the atmosphere across the limb (Fig. 1a). The line-of-sight of GLORIA is actively controlled and stabilised by a gimballed frame. GLORIA covers a spectral range from 780 cm^{-1} to 1400 cm^{-1}. Here, we use observations in the high spectral resolution mode (called "chemistry mode"), which involves a spectral sampling of 0.0625 cm^{-1}. In "chemistry

20   mode", one data cube is recorded within ~13 s (~3 km along flight track) and covers 128 vertical angles from ~5 km to flight altitude plus upward viewing angles simultaneously. Within one data cube, spectra of pixel rows are binned to reduce the noise. From the binned and calibrated spectra, vertical profiles of atmospheric parameters are derived. Thereby, one complete set of atmospheric parameter profiles (i.e. temperature, trace gases and cloud parameters) is obtained from one single data cube. For each atmospheric parameter, the obtained profiles are combined to a 2-dimensional time-height cross section along

25   the flight track.

Optical information on vertical cloud coverage is obtained directly from the calibrated spectra by using the cloud index method (Spang et al., 2004). The cloud index uses the colour ratio between the spectral microwindows from 788.20 cm^{-1} to 796.25 cm^{-1} and 832.30 cm^{-1} to 834.40 cm^{-1}. Details on the trace gas retrieval and the data products used here are provided by Johansson et al. (2018a). In the gas-phase H_2O retrieval, one spectral transition in the microwindow from 795.7 cm^{-1} to 796.1

30   cm^{-1} is used. O_3 is retrieved using the spectral microwindows from 780.6 cm^{-1} to 781.7 cm^{-1} and 787.0 cm^{-1} to 787.6 cm^{-1}. Gas-
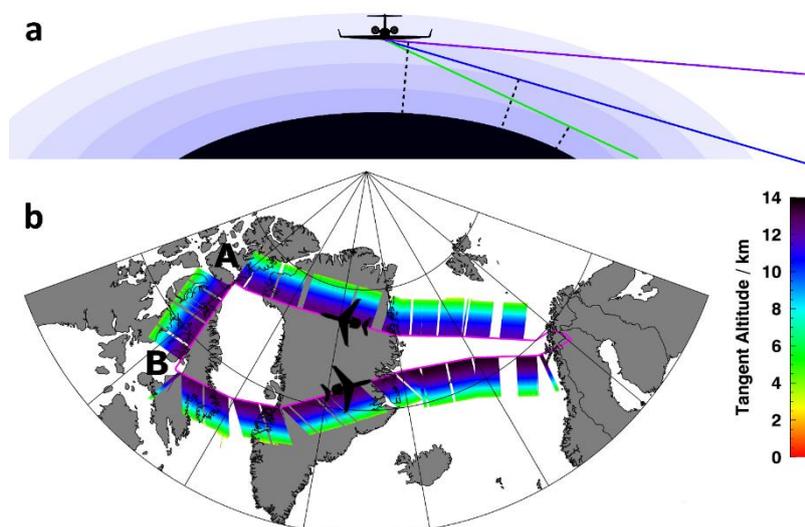
Figure 1. (a) Schematic representation of airborne limb viewing geometry. (b) GLORIA sampling during PGS14 on 26 February 2016. The tangent points of the GLORIA limb views are colour-coded with altitude. Characteristic waypoints are marked (A, B).

phase HNO_3 is retrieved using the spectral microwindows from 862.0 cm^{-1} to 863.5 cm^{-1}, 866.1 cm^{-1} to 867.5 cm^{-1}, and 901.3 cm^{-1} to 901.8 cm^{-1}. As the retrieval of trace gases is not possible in the presence of optically thick clouds, GLORIA limb spectra have been filtered by a dedicated cloud filter based on the cloud index. The estimated accuracy of the GLORIA data amounts to 10 % to 20 % for the respective trace gases (Johansson et al., 2018a). Typical vertical resolutions between 300 and 700 m are achieved for these trace gases.

2.2 ICON-ART chemistry-transport simulations

The state-of-the-art global meteorological forecast system ICON (Zängl et al., 2015) is operational at the German Weather Service (Deutscher Wetterdienst, DWD) since 2015. ICON is developed by the DWD in cooperation with the Max-Planck-Institute for Meteorology, Hamburg. ICON uses a triangular grid, which is well suited for modern computer architectures. Further, it allows efficient scaling of the dynamical core, avoids meridional grid-convergence and singularities at the poles, improves mass conservation and allows efficient local grid refinement (nesting). In the vertical domain, a hybrid height coordinate is used (Leuenberger et al., 2010) that continuously transforms from local topography-following levels to constant height levels at 16 km and above.

The Aerosols and Reactive Trace gases module ART is developed at the Karlsruhe Institute of Technology (KIT). It simulates chemical processes and aerosols, and couples trace gas concentrations and aerosols at each model time step to other relevant processes (Rieger et al., 2015; Schröter et al., 2018). The ICON transport scheme redistributes the tracers, and clouds and
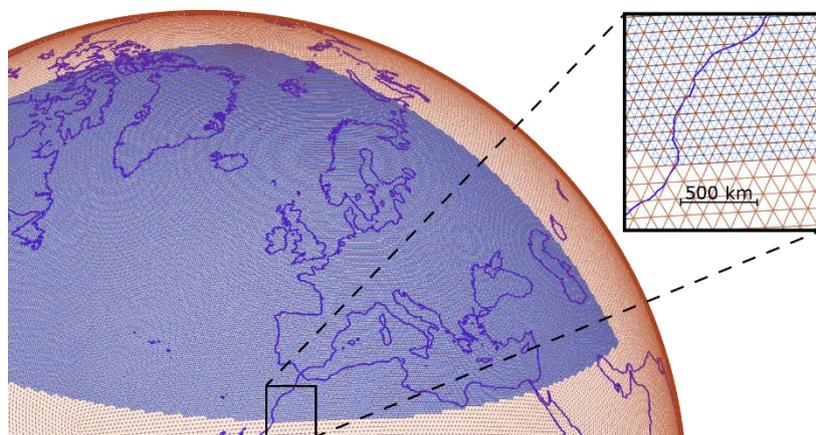
Figure 2. For the POLSTRACC winter, a global ICON-ART simulation with a global R2B6 grid was carried out (red). In the area of the flights, a nest with a R2B7 grid with ~20 km horizontal grid spacing was used to resolve mesoscale processes (blue).

radiation properties are coupled to the meteorological state. ART is capable of simulating chemical and photo-chemical production and loss of reactive trace gases and can be used with defined emission scenarios (Weimer et al., 2017).

For the PGS campaign, a dedicated ICON-ART simulation was performed for the entire polar winter 2015/16 using a R2B6 (~40 km horizontal grid spacing) global grid. In the focus region around Scandinavia and Greenland, a R2B7 nest with a

5   horizontal grid spacing of 20 km was applied (Fig. 2). The potential of the nesting property was recently shown by Weimer et al. (2021). In the vertical, 90 model levels from the ground to 75 km were employed, corresponding to a vertical resolution of ~400 m in the vertical region of interest here. Concerning the meteorology, the simulation was set up in a constrained forecast mode. Every day at 0 UTC, the atmospheric state (pressure, temperature, wind, potential vorticity (PV), as well as specific humidity (q_v), and cloud parameters) was reinitialised using operational ECMWF (European Centre for Medium-Range

10   Weather Forecasts) Integrated Forecast System (IFS) data at horizontal resolution of T1279 (approx. 16 km) and with 137 vertical levels (see Ehard et al., 2018). Therefore, small discontinuities in the meteorological state (including q_v) are possible at the reinitialisation points. To investigate cirrus cloud effects on the LMS water vapour distribution on short forecast time scales, we furthermore use the tracer "H_2O passive". This tracer is mostly identical with q_v (including regular reinitialisation at 0 UTC), but does not account for cloud microphysics (i.e. nucleation and sedimentation of ice particles).

15   Other than the meteorological variables, tracers, such as the ozone tracers, are simulated continuously in a free-running mode after initialisation at the beginning of the winter, using a previous EMAC simulation (Schröter et al., 2018) and are not reinitialised regularly at 0 UTC. The simulation of polar stratospheric ozone loss in the simulated "O_3 tracer" was done using linearised ozone chemistry (LINOZ) and a cold tracer (Schröter et al., 2018; Braesicke and Pyle, 2003), which is activated when temperatures are below a threshold temperature of 195 K. The cold tracer indicates air masses suitable for polar

20   stratospheric clouds, heterogeneous chlorine activation and thus chemical ozone depletion. The cold tracer is characterised by

a life time of 2 days and declines exponentially when temperatures rise above the threshold temperature to account for chlorine deactivation. This way, the full chlorine chemistry on stratospheric clouds is imitated by using the simplified approach of the cold tracer, rather than explicitly calculated. Furthermore, a passive ozone tracer is simulated ("O_3 passive") that is only transported and not affected by chemistry.

5 For qualitative comparisons with clouds observed by GLORIA, the sum of specific cloud ice content (q_i) and snow mixing ratio (q_s) is used to generate a cloud mask (Table 1). Furthermore, we compare the ICON-ART variables specific humidity (q_v), passive specific humidity tracer ("H_2O passive"), ozone tracer ("O_3 tracer"), and passive ozone ("O_3 passive") with the corresponding GLORIA data (Table 2). Since q_v, "H_2O passive", q_i and q_s are reinitialised at 0 UTC, the model data shown in the direct comparisons with GLORIA represent short-term forecasts between ~12 to 21 hours that are interpolated to

10 the corresponding geolocations of the GLORIA observations along the flight track. In contrast, the "O_3 tracer" and "O_3 passive" data are simulated continuously and integrate the effects of transport, mixing, and chemical processes (the latter for "O_3 tracer" only).

2.3 EMAC chemistry-climate simulations

15 The ECHAM/MESSy Atmospheric Chemistry (EMAC) model is a numerical chemistry and climate simulation system that includes submodels describing tropospheric and middle atmosphere processes and their interaction with oceans, land and human influences (Jöckel et al., 2010). It uses the second version of the Modular Earth Submodel System (MESSy2) to link multi-institutional computer codes. The core atmospheric model is the 5th generation European Centre Hamburg general circulation model (ECHAM5, Roeckner et al., 2006). In this study we used EMAC (ECHAM5 version 5.3.02, MESSy version

20 2.52, see Jöckel et al., 2010) in the T42L90MA and T106L90MA-resolution, i.e. with a spherical truncation of T42 (corresponding to a quadratic Gaussian grid of 2.8 by 2.8 degrees in latitude and longitude) and T106 (1.125 by 1.125 degrees) with 90 vertical hybrid pressure levels up to 0.01 hPa (approx. 80 km). A schematic representation of the horizontal model grid is shown in Fig. 3. A Newtonian relaxation technique of the surface pressure and the prognostic variables temperature, vorticity, and divergence above the boundary layer and below 1 hPa towards the ECMWF reanalysis ERA-Interim (Dee et al.,

25 2011) has been used to simulate realistic synoptic conditions.

The applied model setup includes a comprehensive chemistry with gas-phase and heterogeneous reactions on Polar Stratospheric Clouds (PSCs) and comprises about 35 submodels, including the chemistry submodel MECCA (Sander et al., 2011), the photolysis submodel JVAL (Sander et al., 2014), the submodel MSBM, mainly responsible for the simulation of PSCs (Kirner et al., 2011), the submodel CLOUD, based on the ECHAM5 cloud scheme, simulating large scale clouds

30 (Roeckner et al., 2006), the submodel CONVECT, calculating the convection and convective clouds (Tost et al., 2006b), and the submodel SCAV, responsible for scavenging and wet deposition of trace gases and aerosols (Tost et al., 2006a).

We performed three different simulations from 1 July 2015 to 1 April 2016 (initialised with an older EMAC simulation which was started in 1994 and perpetuated to recent years), thus including the Arctic winter 2015/2016 and the PGS campaign. In the first simulation (our "standard" simulation), we use the horizontal resolution of T106 (EMAC-STD). Additionally, we

5 **Table 1.** Data sets and cloud parameters (cirrus/ice clouds).

Dataset	Cloud parameter	Unit
GLORIA	Cloud index	-
EMAC	Large scale cloud snow/ice content (iwc) + convective cloud snow/ice content (cv_iwc)	kg / kg
ICON-ART	Specific cloud ice content (q_i) + snow mixing ratio (q_s)	kg / kg

Table 2. Data sets, trace gas products and sensitivity simulations.

Dataset	Water vapour	Ozone	Nitric acid
GLORIA	H_2O	O_3	HNO_3
EMAC-STD	H_2O H_2O passive[1]	O_3 O_3 passive[1]	HNO_3 HNO_3 passive[1]
EMAC-T42 EMAC-NOSCAV	H_2O	O_3	HNO_3
ICON-ART	Specific humidity[2] (q_v) H_2O passive[3]	O_3 tracer O_3 passive[1]	-
Unit:	ppmv	ppmv	ppbv

10 [1]no chemical sinks and sources, no cloud microphysics
[2]reinitialised daily at 00 UTC using ECMWF IFS, no chemical sinks and sources
[3]reinitialised daily at 00 UTC using ECMWF IFS, no chemical sinks and sources, no cloud microphysics
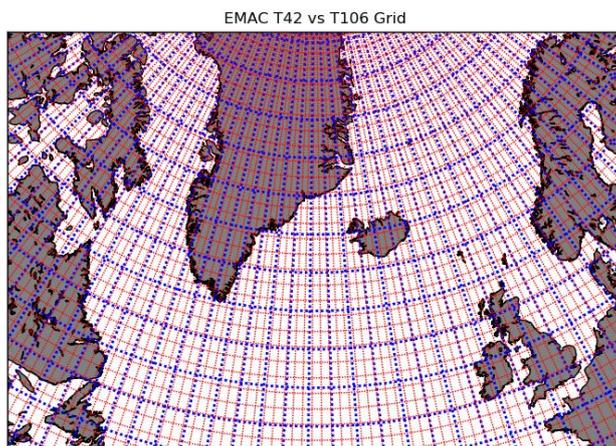
Figure 3. The EMAC standard and sensitivity simulations employed Eulerian grids with 106 (red) and 42 (blue) spectral coefficients. The corresponding T106 (T42) grid corresponds to a horizontal resolution of 125 km (310 km) at the equator. Due to the grid convergence, the zonal grid spacing reduces towards the poles and amounts to ~40 km (~110 km) at 70° N.

performed two sensitivity simulations: First we reduced the horizontal resolution to T42 (EMAC-T42). In the second, we switched off the scavenging processes on ice particles, using again the T106 resolution (EMAC-NOSCAV). For comparisons with clouds observed by GLORIA, the combination of EMAC large scale cloud snow/ice content (iwc) and convective cloud snow/ice content (cv_iwc) is used (see Table 1). With respect to trace gases, the following EMAC variables are used: water vapour (H_2O), ozone (O_3), and gas-phase nitric acid (HNO_3) (Table 2). Furthermore, corresponding passive tracers are simulated, neglecting chemical sinks/sources and cloud microphysics ("H_2O passive", "O_3 passive", and "HNO_3 passive").

2.4 Diagnostics

The vertical profiles of clouds and trace gases are combined to time-height cross sections of these parameters along the HALO flight tracks. For direct comparisons of synoptic and mesoscale patterns with the models, the ICON-ART and EMAC fields of the respective parameters are interpolated to the tangent point geolocations of the GLORIA observations (Fig. 1) to yield the corresponding model cross sections. In the vertical cross sections of the GLORIA data products, PV contours from the corresponding ECMWF reanalysis are superimposed to indicate the dynamical tropopause. For the model cross sections, PV is interpolated from the respective model output.

To quantify biases in the modelled trace gas distributions, the GLORIA and the interpolated model data of the variable under consideration are correlated against each other. In this manner, discrepancies between model simulations and observations can be identified as systematic deviations of data point populations that deviate from the respective 1:1 line. For a vertical assignment, e.g. to identify which data points are associated to the UT or LMS, the data points in the correlations are colour-coded with the corresponding PV values of the models.

3 Flight overview and meteorological analysis

Due to low planetary wave activity the Arctic winter 2015/2016 was extraordinarily cold (concerning the decades before), and a strong polar vortex formed during November and December 2015 (Matthias et al., 2016). Cold conditions prevailed until February 2016. Then, three minor stratospheric warmings led to slightly warmer conditions in the polar vortex, but

5 temperatures remained below the NAT PSC existence temperature (~195 K) on synoptic scales. In early March, the Arctic winter ended by the final stratospheric warming of the season. By mid-March, the vortex was displaced far off the pole and split. The "offspring" vortices decayed rapidly, resulting in a full breakup of the vortex remnants by early April (Manney and Lawrence, 2016).

PGS 14 was performed on 26 February 2016 from Kiruna, northern Sweden. Take off of the HALO aircraft was at 11:19 UTC

10 and landing time was at 20:59 (flight duration of 9:40 h). The HALO flight track (anti-clockwise) and the tangent points of the GLORIA limb observations are shown in Fig. 1b. After take off, HALO headed westwards (GLORIA pointing to northward directions), crossed the Atlantic and Greenland, and continued its flight towards Canada. Then, at waypoint A, it turned to a southward direction (GLORIA pointing to westward directions). Finally, after waypoint B, HALO turned back to eastward directions and headed back towards Scandinavia (GLORIA pointing southwards).

15 Figure 4 shows the meteorological situation on the day before the flight at 12 UTC (left column) and for the flight day at 18 UTC, i.e. during the eastward flight leg back to Kiruna. The colour-coded contour plots in the upper row show ICON-ART q_v at 10 km together with ICON-ART potential temperature (white contours) to visualise the dynamical situation in the UT/LMS region. West of the flight track, dry air masses characterised by high potential temperatures exceeding 340 K to 350 K indicate a deeply subsided air mass of the late polar vortex, which was probed by the GLORIA observations during and around the

20 southward-heading leg. Relatively dry high-latitude LMS air masses are found above Greenland, the Arctic sea, and northern Europe and were probed by GLORIA during the west- and eastward-heading legs (i.e., prior to way-point A and after B, respectively). These high-latitude LMS air masses are interspersed with moist filaments connected to moist upper tropospheric air masses in the south. A broad filament of moist air stretches across the British islands, Iceland and Greenland on 25 February 2016 and partly dissipates on 26 February 2016. During the backward leg to Kiruna, GLORIA pointed subsequently towards

25 upper tropospheric air masses (i.e. high specific humidity >40 ppmv) and into dissipating filaments above central Greenland. The surface weather conditions are shown in the lower row of Fig. 4. On 25 February 2016, a well-defined low-pressure system is located above Scandinavia, and patchy weak high-pressure systems are found around central Greenland and Canada. A strong Azores high is located in the Atlantic Ocean together with a compact Icelandic low located at the southern tip of Greenland, going along with a notably positive North Atlantic Oscillation (NAO) index of +1.61 for February 2016. With

30 a NAO value of +1.62 for the period from October to March, the winter 2015/16 ranks to date just within the top ten of the highest seasonal values on record for this period of the year (both NAO values are retrieved from the record given at https://crudata.uea.ac.uk/cru/data/nao/values.htm; last access: 12 April 2021). An elongated occlusion stretches from South Greenland along Iceland to the Atlantic region near western Ireland. On the flight day, the front associated with the Icelandic
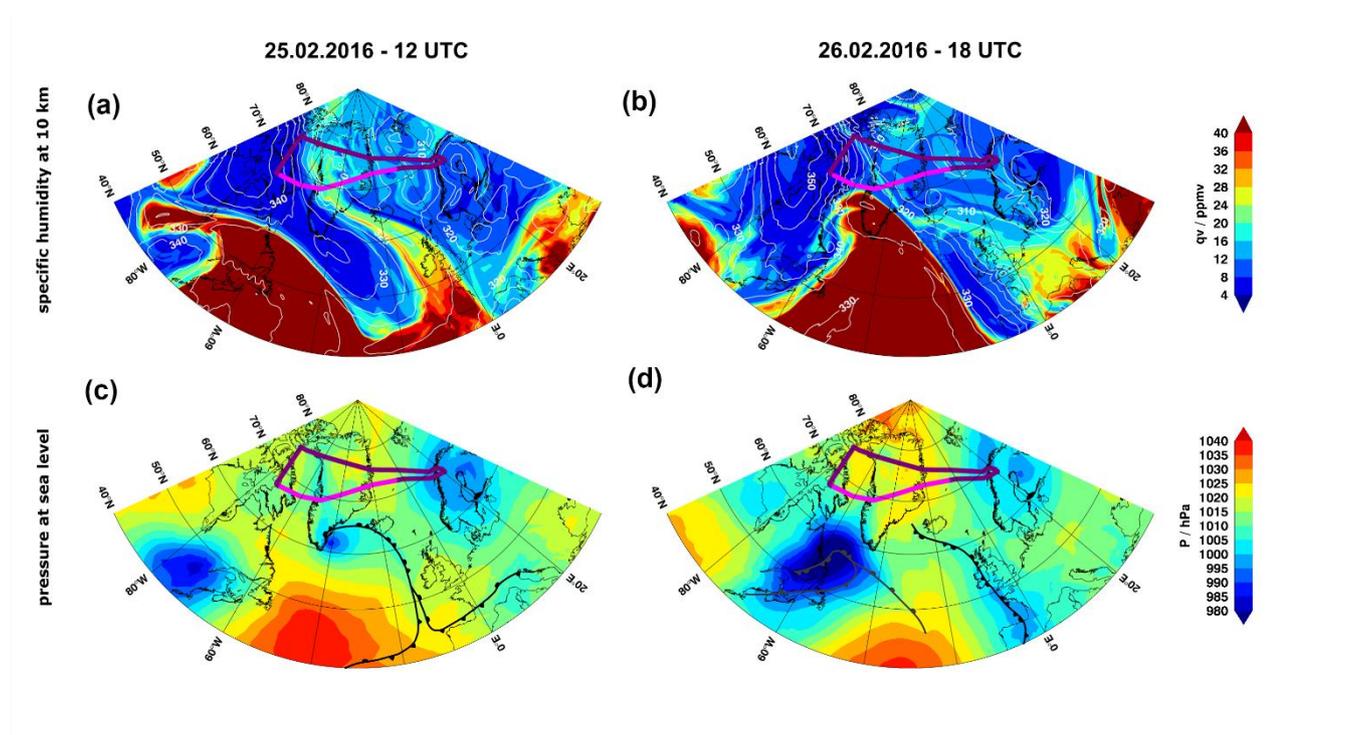
Figure 4. Meteorological conditions in the tropopause region and at sea level on 25 February 2016 (left column) and the flight day, 26 February 2016 (right column) as modelled by ICON-ART. Specific humidity is colour-coded in contour and potential temperature is shown as white contour lines at 10 km altitude (a,b). Pressure at sea level together is shown with selected warm fronts, cold fronts and occlusions (c,d). The HALO flight track on flight day is indicated by a purple line with the section of the flight covering the troposphere-to-stratosphere exchange region highlighted in magenta in all panels.

low is fully occluded, while the situation above Greenland and Canada has only slightly changed. When comparing with the conditions in the UT/LMS region (Fig. 4, upper row), it can be seen clearly that the broad moist filament across Greenland on 25 February 2016 and its remnants on the flight day are connected to the occlusion associated with the Icelandic low. In the occlusion, moist tropospheric air masses are entrained into the surrounding LMS, and filaments of moist air are situated along

5 the viewing direction of GLORIA during the backward leg across Greenland.

Overall, at 10 km the air masses observed by GLORIA on 26 February 2016 subsequently comprise (i) the high-latitude LMS including patchy filaments, (ii) deeply subsided polar vortex air masses above Canada, (iii) upper tropospheric air masses above southern Greenland, (iv) moist air filaments above Greenland and associated with the occluded front of the Icelandic low, and (v) again high-latitude LMS air masses. Therefore, the GLORIA observations provide a unique opportunity to test

10 the capability of ICON-ART and EMAC in simulating the Arctic winter UT/LMS region.

4 Observed and modelled cloud and trace gas distributions

4.1 Clouds

The vertical cross section of the GLORIA cloud index of the entire flight is shown in Fig. 5a. Cloud index (CI) values close to one are indicative of optically thick conditions, i.e. in the presence of clouds, whereas CI values approaching four and higher

5 can be considered as cloud free conditions (Spang et al., 2004). High tropospheric clouds reaching the dynamical tropopause can be clearly identified around 12 UTC to 13 UTC, 14 UTC to 15 UTC, 16:30 to 17:30 UTC, and around 20 UTC, while a lower cloud system coinciding with a lower dynamical tropopause is detected directly at the beginning of the flight (prior to 12 UTC). A narrow band of low CI values is also visible around waypoint A around 8 km altitude. Further individual clouds are identified at lower altitudes between 17:30 and 19:30 UTC. Slightly enhanced cloud index values at flight altitude (12 to

10 13 UTC and after 18 UTC) are the consequence of polar stratospheric clouds above flight altitude (Oelhaf et al., 2019) and are not indicative of cirrus clouds here.

In the following, we compare GLORIA cloud index values with cloud masks generated from the models in a qualitative way. The GLORIA cloud index is an optical quantity, while the model cloud masks are generated from the respective model outputs for condensed water in the solid state (see Table 1). Liquid water is not considered, since the temperatures in the focus region

15 are well below the frost point, and there was no significant contribution of liquid water to the used cloud masks. A quantitative comparison (e.g. conversion of modelled cloud properties into spectral radiances and considering effects related to line-of-sight) is beyond the scope of our study that focuses on the ability of the models to reproduce the smaller scale structures.

We have set the threshold of the cloud mask for the ICON-ART- and EMAC-model at 10^{-9} kg/kg ice/snow water content (cloud parameters, see Table 1). On the one hand this is lower than the estimated sensitivity of 3×10^{-6} g/m^3 for ice water content

20 (IWC) in cirrus clouds of an IR limb sounder (Spang et al., 2015), corresponding to about 1×10^{-8} kg/kg IWC at typical atmospheric conditions at 10 km altitude during the flight. Assuming that the representative concentration for a model grid-box volume is a mean of small-scale patches of enhanced concentrations the choice of a small threshold value for the overall volume seems sensible. On the other hand, it is higher than the lower in situ detection limit of cirrus clouds of 10^{-3} ppmv (Krämer et al., 2020; Schiller et al., 2008) corresponding to 6.2×10^{-10} kg/kg.

25 The ICON-ART cloud mask represents the sum of cloud ice content (q_i) and cloud snow mixing ratio (q_s) interpolated to GLORIA geolocations along the flight track to 2-dimensional time-height cross sections (see Sect. 2.2 and Table 1). It shows the distribution of clouds of the nested ICON-ART simulation in a short consecutive forecast mode along the flight track (see Sect. 2.2).

In the ICON-ART simulation (Fig. 5b), three of the four major cloud systems seen in the GLORIA observations can be

30 identified, with differences in the vertical and horizontal extent. However, the observed cloud system around 14 to 15 UTC below 10 km altitude is missing in this ICON-ART representation. Modelled cloud systems below approx. 10 km around 12 UTC to 13 UTC, 16:30 UTC to 17:30 UTC, and around 20 UTC agree well with GLORIA in the horizontal domain. Discrepancies in the large cloud system around 20 UTC below 6 km can be explained by the fact that no robust information
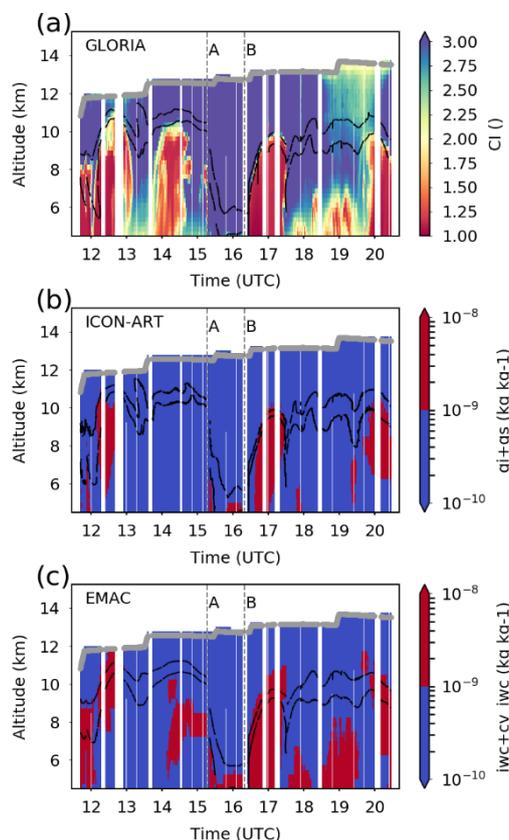
Figure 5. Qualitative comparison of clouds along flight track observed by GLORIA and cloud masks generated from ICON-ART and EMAC. (a) Vertical cross section of cloud index (CI) derived from GLORIA spectra. (b) Cloud mask constructed from ICON-ART specific cloud ice content plus snow mixing ratio. (c) Cloud mask constructed from EMAC large scale snow/ice content plus convective cloud snow/ice content. Black lines: 2 PVU and 4 PVU isolines (lower and higher lines, respectively) from ECMWF reanalysis (a), ICON-ART (b) and EMAC (c) as indicators for the dynamical tropopause. Grey lines: HALO flight altitude.

on vertical cloud structure can be derived from GLORIA, if optically dense cloud layers are located above. In such cases, lower limb views can be optically saturated, and low cloud index values may result although cloud-free conditions are present below. We explain the fact, that the vertically extended cloud system detected by GLORIA around 14 UTC to 15 UTC is not reproduced by the nested ICON-ART simulation by a temporal mismatch in the simulated cloud systems (see Appendix A).

5   Furthermore, the discrepancies might be explained partly by line-of-sight-related effects, since GLORIA accumulates light along extended limb views, while the model is interpolated at a certain geolocation. For the observed cloud systems at lower altitudes between 17:30 UTC and 19:30 UTC only weak indications are found in the nested ICON-ART simulation. Further high cloud systems prior to 12 UTC appear more sharply in the ICON-ART simulation, while a simulated cloud at 16 UTC below 6 km is not confirmed by GLORIA.

10   The respective cloud mask of the EMAC-standard simulation (STD) with the T106L90MA-resolution was generated by using the sum of the large scale cloud snow/ice content (iwc) and the convective cloud snow/ice content (cv_iwc) (see Table 1).

Recalling that the EMAC-simulation uses a continuously nudged meteorology (see Sect. 2.3), however the cloud variables are not nudged. As can be seen in Fig. 5c, the EMAC standard simulation (STD) reproduces the cloud patterns observed by GLORIA well. All of the observed cloud systems can be found in the cross section along the flight path generated from the EMAC simulation. Especially, the observed cloud system between 14 to 15 UTC, which is not reproduced by ICON-ART, is

5   reproduced by EMAC, but with a different morphology and slightly displaced horizontally and vertically. Also, the lower clouds observed between 17:30 and 19:30 UTC are reproduced well by the EMAC-simulation. As in the case of ICON-ART, a simulated low cloud system at 16 UTC is not confirmed by GLORIA.

In the EMAC-simulation the modelled horizontal and vertical extents are mostly larger when compared to ICON-ART (e.g. prior to 12 UTC and 16:30 UTC to 19:30 UTC). The lower model resolution and lower time resolution of the output (1h for

10   EMAC versus 0.25 h for ICON-ART) could be one possible explanation, making a positive cloud detection more likely (concerning the spatial coverage). Furthermore, the lower grid spacing is better comparable to the horizontal extensions of the GLORIA-limb views, which results in a more consistent comparison in certain cases when cloud systems are located along the line-of-sight. The high cloud system prior to 12 UTC matches the GLORIA cloud index better than in the case of ICON-ART, while the cloud system in EMAC at 12 to 13 UTC appears higher than in the GLORIA- and ICON-ART-data, even

15   exceeding the 2 PVU- and 4 PVU-isoline and reaching the GLORIA-flight altitude. The clouds at 16:30 to 17:30 UTC and at 20 UTC also reach higher in the atmosphere in the EMAC-cross section compared to the GLORIA- and ICON-ART data, and again notably higher than the respective local dynamical tropopause. In these cases the ICON-ART cloud mask agrees better with the GLORIA observations.

Another proxy for the characterisation of detectable cloud systems in the model, is looking at the precipitation events. This is

20   done in the following in the case of ICON-ART by using a passive water vapour tracer forecast in the constrained forecast mode as a reference. In addition, this analysis sheds light on to which degree precipitation affects the modelled water vapour in the UT/LMS (cf. Sect. 4.2). The passive water vapour tracer does not account for cloud microphysics and therefore no nucleation, sedimentation and evaporation of hydrometeors. Residuals between the ICON-ART specific humidity forecast (see Sect. 4.2) and the passive reference tracer show where microphysical processes altered UT/LMS humidity within the time

25   frame of the forecast (i.e., between ~12 h to ~20 h, depending on the flight section). It shows the accumulated effect of clouds and therefore indirectly the presence of cloud systems at the respective GLORIA geolocations during the time of the forecast at the day of PGS Flight 14.

Figure 6 shows the residual, i.e. the difference between ICON-ART specific humidity and the passive tracer without cloud microphysics. Negative residuals indicate regions which are depleted in water vapour due to cloud processes. Positive residuals

30   show regions enriched in gas-phase water vapour due to evaporation of ice/snow particles. Negative and positive residuals clearly prove the generation and transformation of hydrometeors in the UT/LMS during the entire flight. Before the waypoint A, a strong pattern with residuals exceeding (~12 UTC to ~13 UTC) and approaching (around 15 UTC) ±20 ppmv is found, with weaker signatures in-between. After waypoint B, a sequence of distinct anomalies well exceeding ±20 ppmv is found until the end of the flight.
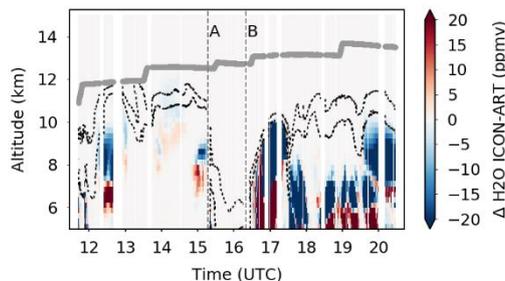
Figure 6. Modelled short-term changes in specific humidity due to cloud processes. Residuals between ICON-ART specific humidity and corresponding H_2O tracer without cloud microphysics. Black dashed lines: ICON-ART 2 PVU and 4 PVU isolines (lower and higher lines, respectively) as indicators for the dynamical tropopause. Grey lines: HALO flight altitude.

The comparison of Figure 6 with Figure 5a shows that this idealised ICON-ART diagnostic is a good proxy for the simulation of clouds in the model and does not require a threshold approach (as discussed above). However, it is an integrated quantity showing the history of "cloud events" on the respective day, whereas the cloud masks show "snapshots" of simulated hydrometeors at the geolocations and time of the measurement. At a closer look all of the observed cloud systems coincide

5   qualitatively with a corresponding precipitation pattern at the respective geolocations in the ICON-ART-data. This means that there is evidence for the existence of all observed cloud systems in the ICON-ART-simulation. In particular, at 14 UTC to 15 UTC, where a cloud system detected by GLORIA is not reproduced by the cloud mask of ICON-ART (as described above and cf. Fig. 5a and 5b), weak negative residuals reaching up to about -5 ppmv stretch even into the LMS and hint at drying of the uppermost troposphere and LMS by high altitude cirrus cloud ice particle sedimentation. Positive residuals of the same

10   magnitude are found below between 9 km to 10 km, and another precipitation pattern in the direct vicinity is found at 14 UTC and reaches further down to below 8 km. Therefore, these cumulative patterns found in Fig. 6 support that a cloud system has been present in the simulation at some time before the measurement during the day of PGS Flight 14.

There is also evidence in the ICON-ART-data for the lower cloud system observed between 17:30 and 19:30 UTC (cf. Fig. 6 with Fig. 5a). Even though this cloud system is underestimated in the simulation (see Fig. 5b), Fig. 6 suggests that it has been

15   present at these locations at some time prior to the measurement in the simulation at the day of the flight.

The narrow cloud band at waypoint A, detected by GLORIA around 8 km, and also evident in the EMAC-cross section (see Fig. 5c), is not visible in the ICON-ART-cross section (cf. Fig 5a/c with 5b). However, again a strong signal of vertical redistribution of water vapour is visible in Fig. 6 at this geolocation, which again, hints at the presence of this cloud system in the ICON-ART simulation at some time prior to the measurement. Thus, uncertainties in the timing of the ICON-ART forecast

20   might partly explain the discrepancies between GLORIA and ICON-ART here beside the other reasons discussed above.

In the Appendix A we will further investigate this issue by sampling the models at the respective GLORIA geolocations with a negative time offset, to shed light on the history and development of the cloud systems in the models at the day of the flight and to prove that seemingly "missing clouds" in the ICON-ART-data based on the cloud mask can be identified in the simulations just a few hours prior to the measurements.

Overall, the simulated precipitation patterns in Fig. 6 are consistent with the observed and modelled cloud systems in Fig. 5 and clearly show that modelled water vapour distributions in the UT/LMS are significantly modulated by more than ± 20 ppmv in the UT and -5 ppmv in the LMS. Therefore, precipitation clearly is a significant factor in modelled UT humidity on short-term time scales and affects also significantly LMS humidity.

5 In summary, most of the major cloud systems can be identified qualitatively in both models. Remaining discrepancies between GLORIA and the models can be explained by horizontal and temporal mismatches of the cloud systems in the simulations and line-of-sight related effects of the GLORIA observation. In particular the fact that the ICON-ART lacks the simulation of the observed large cloud system at 14 to 15 UTC will be addressed in the Appendix A.

4.2 Trace gas distributions

10 In the following, we compare observations of water vapour, ozone and nitric acid with the respective simulated trace gases by ICON-ART and EMAC. For the former only water vapour, i.e. q_v and ozone have been simulated.

Figure 7a-c show the water vapour, ozone and nitric acid distributions observed by GLORIA along the flight track. Before waypoint A, moist tropospheric air masses extend to the dynamical tropopause, which is located mostly around ~10 km in Fig. 7a. Some moist "patches" are also found in the LMS here. In contrast, dry stratospheric air masses reaching down to ~6 km

15 indicate a deeply subsided polar vortex remnant after waypoint A to slightly behind waypoint B (cf. with Fig. 4). Afterwards, again a high tropopause around ~10 km is found. The cloud system from 16:30 UTC to 17:30 UTC (cf. with Fig. 5a) is related to the moist tropospheric air masses above south-western Greenland (cf. with Fig. 4). In the subsequent part above central Greenland between 17:30 to 19:00 UTC, a highly textured LMS is found. Narrow moist filaments of tropospheric air reach by ~2 km up into the LMS, and the dynamical tropopause altitude oscillates along the flight track. Afterwards, a more homogenous

20 tropopause and water vapour distribution is found until the end of the flight.

The ozone distribution (Fig. 7b) shows a complimentary pattern compared to water vapour. At tropospheric altitudes, low ozone mixing ratios are found, while ozone mixing ratios above the tropopause increase with altitude. Also, in the ozone distribution, the deeply subsided polar vortex remnant from waypoint A to slightly after waypoint B can be clearly identified by high ozone mixing ratios reaching down towards ~6 km. From 17:30 UTC to ~19:00 UTC, filaments of low ozone

25 correspond to the structures of enhanced water vapour (Fig. 7a) and reach up to the flight altitude therefore even deeper into the LMS. For nitric acid (Fig. 7c), a similar pattern is found as for ozone, but with a higher contrast and more pronounced filaments from 17:00 UTC to ~19:00 UTC. Furthermore, the nitric acid distribution shows a local maximum at and below flight altitude from 14 UTC until the end of the flight, which are a consequence of nitrification of the LMS in the same winter (see Braun et al., 2019).

30 As shown in Figure 7d, the overall distribution and mesoscale structures in the ICON-ART specific humidity forecast agree well with water vapour detected by GLORIA. Recall that q_v was reinitialised using operational ECMWF IFS data at 00 UTC. The location of the strongest gradient in water vapour (roughly the transition from red to yellow shadings) is matched well
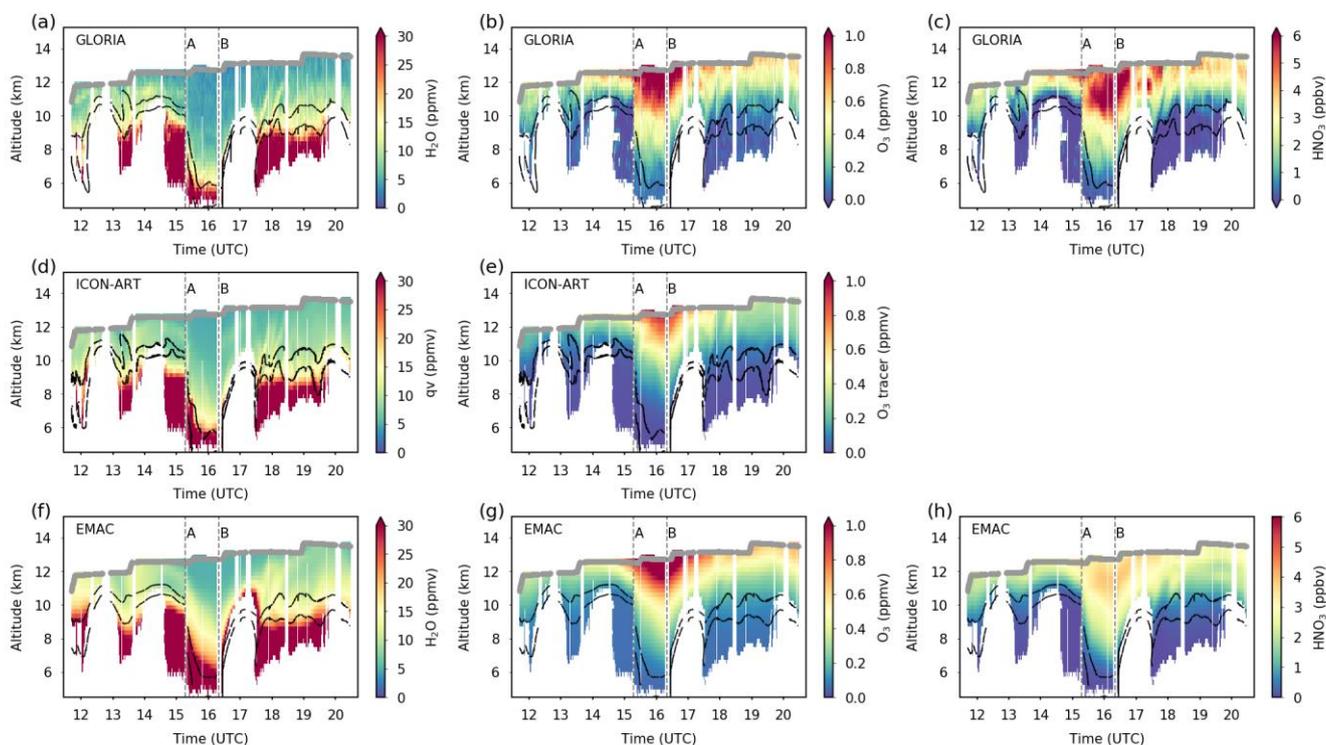
Figure 7. Observed and modelled trace gas distributions. GLORIA observations of water vapour, ozone and nitric acid (a-c). ICON-ART (nested simulation) short-term forecast of specific humidity (d) and free-running simulation of ozone using simplified ozone depletion parameterisation (e). EMAC free-running simulations of water vapour, ozone and nitric acid (f-h). Black lines: 2 PVU and 4 PVU isolines (lower and higher lines, respectively) from ECMWF reanalysis (a-c), ICON-ART (d,e) and EMAC (f-h) as indicators for the dynamical tropopause. Grey lines: HALO flight altitude.

during the entire flight. This applies also for subsided air masses from waypoint A to slightly behind B. Therefore, the water vapour distribution suggests that the dynamical structure of the late vortex air masses is modelled in a realistic way by ICON-ART. During the backward leg to Kiruna, excellent agreement is found for the narrow moist filaments and structures stretching into the LMS between 17:30 UTC and 19:00 UTC.

5 Keeping in mind that water vapour is simulated by EMAC continuously (i.e. no reinitialisation at 00 UTC and not nudged), the EMAC-STD simulation also reproduces the observed water vapour distribution well (Fig. 7f). Naturally, less details are found in the EMAC simulation due to the lower horizontal resolution. The subsided air mass from A to slightly behind B is reproduced by EMAC. However, moister air masses with water vapour > 20 ppmv reach here higher up by 1-2 km. Furthermore, stratospheric air masses above the dynamical tropopause appear slightly moister in the EMAC simulation when

10 compared to GLORIA and ICON-ART, and moist air masses reach above the dynamical tropopause in the vicinity of the cloud system around 17 UTC. Surprisingly, the moist filaments and structures seen in the GLORIA and ICON-ART data between 17:30 UTC and 19:00 UTC can be identified schematically in the EMAC simulation.

The continuous ICON-ART ozone simulation (i.e. no reinitialisation at 00 UTC) also matches the mesoscale patterns seen in the GLORIA observations (Fig. 7e), with however systematically lower volume mixing ratios. Again, the deeply subsided air

masses from waypoint A to slightly after waypoint B can be clearly identified by higher ozone mixing ratios reaching down to lower altitudes. Similar filaments and structures as seen in the GLORIA observation between 17:30 UTC and 19:00 UTC are identified, with however less details and fine-structures. The EMAC ozone distribution (Fig. 7g) matches the GLORIA observations well within the limitations of the model resolution, as already discussed by Johansson et al., (2019). Here, absolute

5 mixing ratios agree quite well with the GLORIA observations. All major structures are reproduced, and weak indications are found again for the filaments and structures from 17:30 UTC to 19:00 UTC. The overall ozone mixing ratios in the EMAC simulation are higher when compared to ICON-ART and closer to the absolute values observed by GLORIA.

The nitric acid distribution simulated by EMAC (Fig. 7h) matches the overall structure seen in the GLORIA data only qualitatively. Systematically lower mixing ratios are found in the EMAC data, and local maxima seen in the GLORIA

10 observations between 14 and 19 UTC are hardly reproduced. This is probably due to the fact that EMAC underestimates nitrification of the LMS in this particular winter. A similar underestimation of nitric acid simulated by EMAC was found for the Arctic winters 2009/2010 and 2010/2011 as discussed in Khosrawi et al., (2018), and also in the comparison to GLORIA measurements of research flight 21 on 18 March 2016, described in Khosrawi et al., (2017). However, the observed narrow filaments with low nitric acid reaching into the LMS between 17:30 UTC and 19:00 UTC are again reproduced well by the

15 model.

In summary, the dynamical situation is represented well by both models (with either consecutive ICON-ART forecasts or continuously nudged EMAC simulations) within the limitations of their horizontal resolution. Both models clearly reproduce the observed strongly subsided air masses in the western part of the flight and the narrow filaments between 17:30 UTC and 19:00 UTC. Here, complimentary patterns are found in the water vapour distribution when compared to ozone and nitric acid.

20 Water vapour in the LMS is overestimated by EMAC, and ozone is underestimated by ICON-ART. Furthermore, EMAC clearly underestimates nitric acid and hardly reproduces nitrification patterns seen in the GLORIA data.

4.3 Troposphere-to-stratosphere exchange region

Close-ups of the GLORIA, ICON-ART and EMAC-STD trace gas distributions are presented in Fig. 8. In Figure 8a, two stronger moist filaments reaching into the LMS up to ~12 km are seen between 17:30 and 18:30 UTC, with a weaker filament in-between at ~18 UTC. The typical horizontal extent of the filaments along the flight direction is only 50-100 km just above

25 the tropopause. During the further course of the flight warped regions of the dynamical tropopause are identified until ~19 UTC.

ICON-ART specific humidity reproduces the vertical and horizontal extent as well as maximum mixing ratios very well (Fig. 8d). Even the weak filament in-between the more developed filaments can be clearly identified. However, overall water vapour

30 mixing ratios are slightly higher when compared to GLORIA. In the EMAC simulation, the two major filaments can be weakly identified, and warping of the dynamical tropopause is weaker (Fig. 8f). Thereby, it has to be remembered that the horizontal resolution of the EMAC simulation is T106, which is by a factor of ~5 lower than that the ICON-ART R2B7 nest. Overall absolute water vapour mixing ratios are clearly overestimated by EMAC.
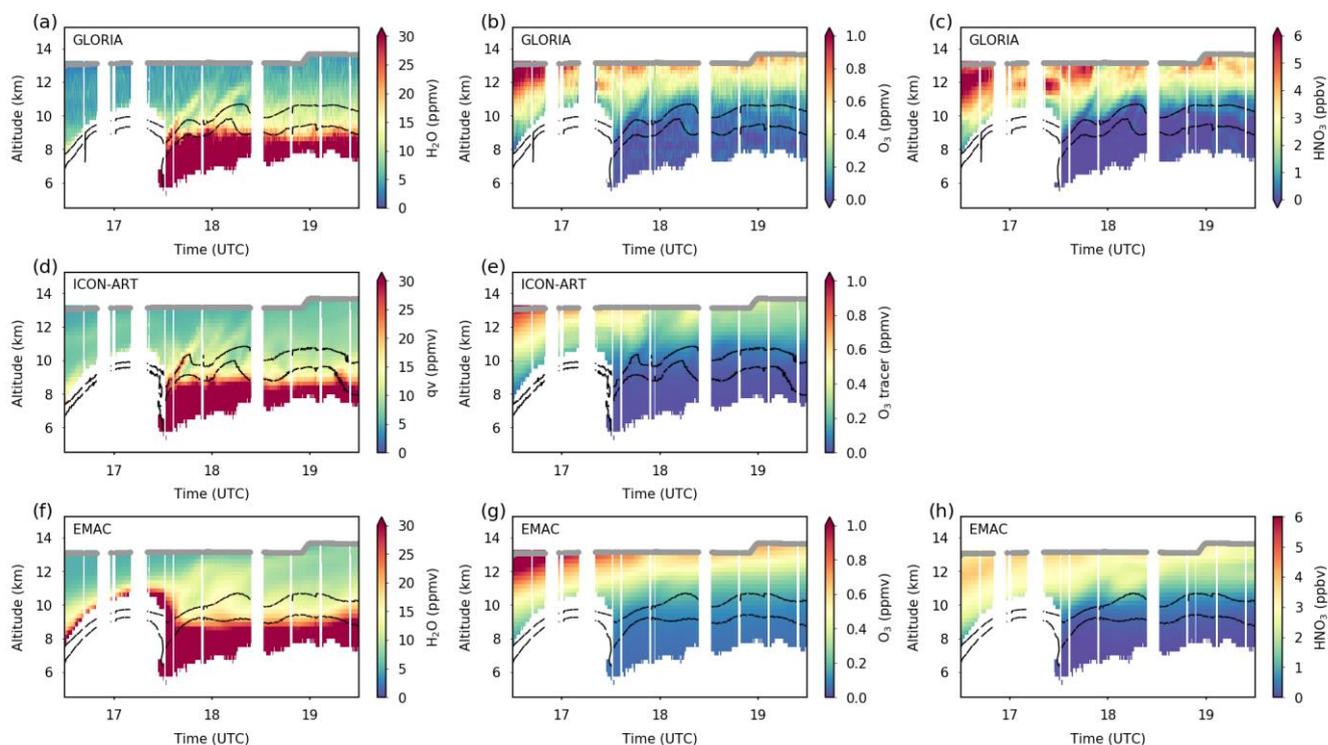
Figure 8. Close-ups of troposphere-to-stratosphere exchange region between 17:30 and 19:30 UTC. For legend, see Fig. 7.

The GLORIA ozone distribution shows detailed fine structures close to the flight altitude. Structures low in ozone correspond to the respective high water vapour structures and extend further to flight altitude (Fig. 8b). The combination of ozone and water vapour data clearly shows that folded airmass structures reach deeply into the LMS and that active troposphere-to-stratosphere exchange takes place here. ICON-ART ozone reproduces the same sequence of filaments, with however lower mixing ratios and less fine structure. EMAC reproduces the filaments around 17:30 UTC to 18:30 UTC only faintly, while observed absolute mixing ratios are matched well. Finally, the GLORIA close-up in Figure 8c shows highly structured nitric acid distribution. EMAC again resolves the filaments in principle, while mixing ratios are clearly underestimated and local maxima are hardly reproduced.

In summary, Figure 8 shows that ICON-ART using the R2B7 (~20 km resolution) nest is able to resolve mesoscale fine structures with a horizontal extent of less than 100 km. In case of specific humidity, a similar degree of detail is achieved when compared to GLORIA, while less details are found in the simulation in the case of ozone. Given the lower resolution of the nudged T106 simulation of the EMAC model, we find that this model also reproduces dynamical structures at the lower edge of its resolution. Clear evidence for troposphere-to-stratosphere exchange in connection with the occlusion seen in Fig. 4 is found. Deviations in the trace gas distributions by both models are found and are quantified in the following section.

4.4 Quantification of model discrepancies and sensitivity studies

By scattering and correlating modelled mixing ratios with the observed values, model discrepancies (and likely biases) can be quantified as deviations from the ideal 1:1 line (Fig. 9). Furthermore, a dynamical context in the vertical domain is provided by colour-coding the data points with corresponding PV values.

5   To quantify the simulated cumulative impact of ozone depletion and nitrification of the LMS in the ICON-ART and EMAC simulations during the entire winter until the flight date, corresponding passive tracers are simulated (Fig. 10). Residuals between the "active" tracers (i.e. chemical and microphysical processes activated) and the corresponding passive tracers (only dynamical processes act on it) indicate the cumulative net changes due to the processes considered in the "active" case.

For ICON-ART specific humidity, excellent agreement is found for high tropospheric water vapour levels (Fig. 9a). At PV

10   levels higher than ~4 PVU, a systematic moist bias is evident in the ICON-ART model data. The systematic offset at the high PV levels is attributed to the same systematic moist bias, which is known for the ECMWF and other weather forecast systems (e.g. Stenke et al., 2008). It is not unexpected that this bias is translated into the ICON-ART simulation, since the simulation is done in a constrained forecast mode reinitialised from ECMWF IFS data. The correlation of EMAC H_2O with GLORIA
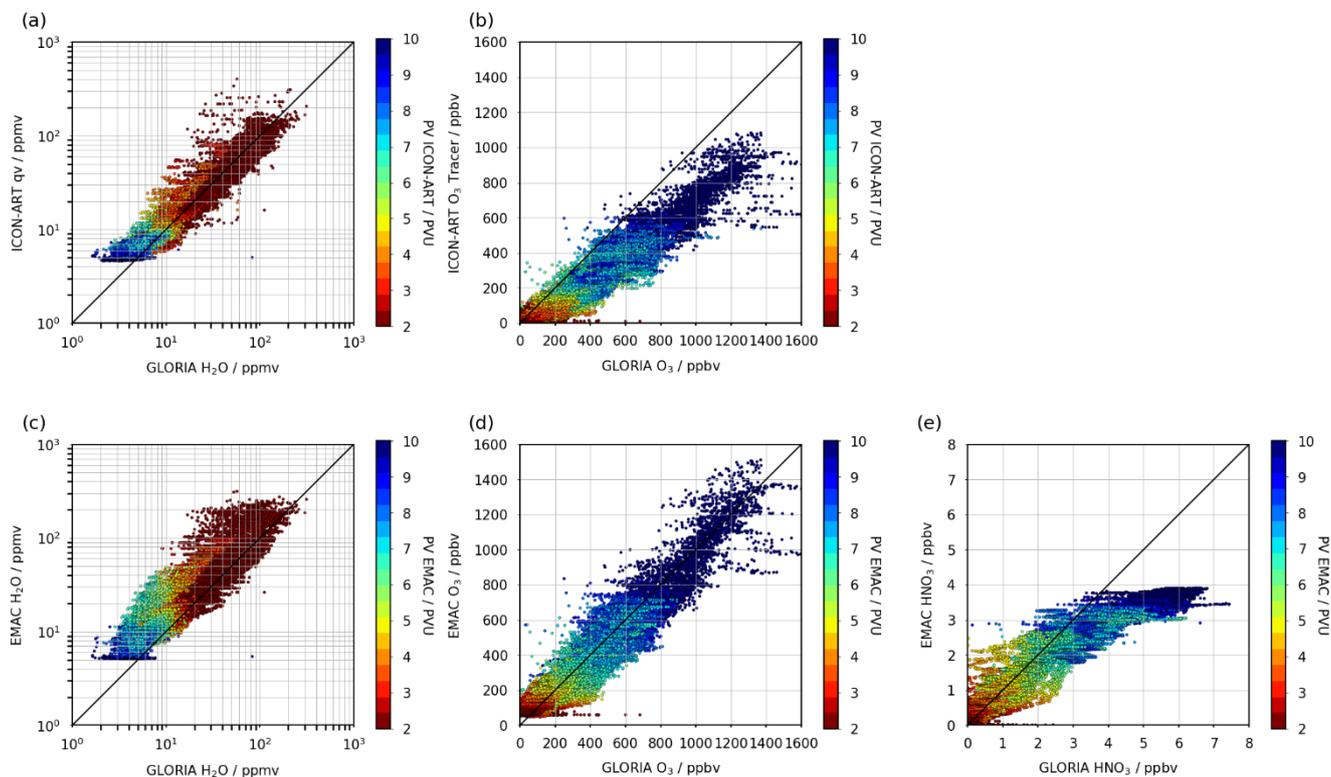
Figure 9. Correlation of GLORIA H_2O, O_3 and HNO_3 to corresponding EMAC and ICON-ART output variables. Colour-coding: PV from corresponding model.

water vapour (Fig. 9c) shows a systematic moist bias in the model from the troposphere (low PV values, red) up to the highest stratospheric air masses accessed (high PV values, blue). At 6 PVU and higher, the pattern of the moist bias is similar to the bias in the ICON-ART forecast (Fig. 9a). Only at highest PV levels, both model data sets move again somewhat towards the ideal 1:1 line.

5   For the nested ICON-ART ozone (Fig. 9b), a systematic low bias is found and increases with PV. This is attributed to the simplified ozone depletion parameterisation. For the T106 EMAC simulation the agreement in ozone with GLORIA measurements is very good (Fig. 9d). Here, the data points are well scattered around the 1:1 line at all PV levels. Small groups of data points with larger deviations at high PV values are attributed to fine-structures in the LMS, which are seen in the GLORIA data, but which are not resolved by the model (cf. e.g. Fig. 8b versus 8g).
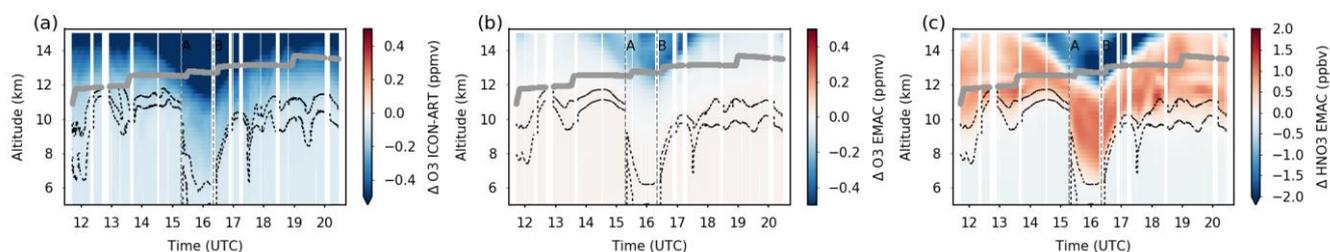
Figure 10. Modelled ozone depletion and changes in nitric acid due to chemical and microphysical processes. Residuals between (a) ICON-ART O_3 tracer and passive O_3 tracer, (b) EMAC O_3 and O_3 passive, and (c) EMAC HNO_3 and HNO_3 passive. Black dashed lines: ICON-ART/EMAC 2 PVU and 4 PVU isolines (lower and higher lines, respectively) as indicators for the dynamical tropopause. Grey lines: HALO flight altitude.

10   In the ICON-ART simulation, the "active" ozone tracer simulation shows systematically lower mixing ratios than the "passive" ozone tracer (Fig. 10a) at all altitudes due to modelled ozone depletion. Above the dynamical tropopause, the difference increases from -0.1 ppmv to more than -0.4 ppmv and shows that ozone deficit increases in the late polar vortex.

In the EMAC simulation (Fig. 10b), the residual is close to zero in the troposphere, in the tropopause region and also at lower levels of the LMS. Only in the deeply subsided vortex remnant around waypoint A and B, ozone is significantly lower in the

15   "active" simulation, which is indicated by residuals exceeding -0.2 ppmv. The fact that ICON-ART residuals are more negative in the LMS than in the case of EMAC and recalling that overall EMAC ozone agrees well with GLORIA (see Fig. 9d) suggests that the simplified ozone scheme by ICON-ART overestimates ozone depletion in the LMS.

While EMAC nitric acid agrees well with GLORIA in the troposphere, a systematic low bias is found above the troposphere and strongly increases with altitude (Fig. 9e). The bias amounts ~50 % at the highest PV levels of ~10 PVU under consideration

20   and suggests that the observed nitrification of the LMS is hardly reproduced. A similar bias has been identified by Khosrawi et al., (2017) while comparing EMAC with GLORIA results (PGS Flight 21).

The EMAC nitric acid residual shown in Fig. 10c clearly shows that this species is enhanced in the simulation by ~ 1 to 1.5 ppbv in the LMS and depleted above due chemical processes and/or (de-) nitrification by evaporating HNO_3-containing particles in polar stratospheric clouds (see also Khosrawi et al., 2017, 2018).
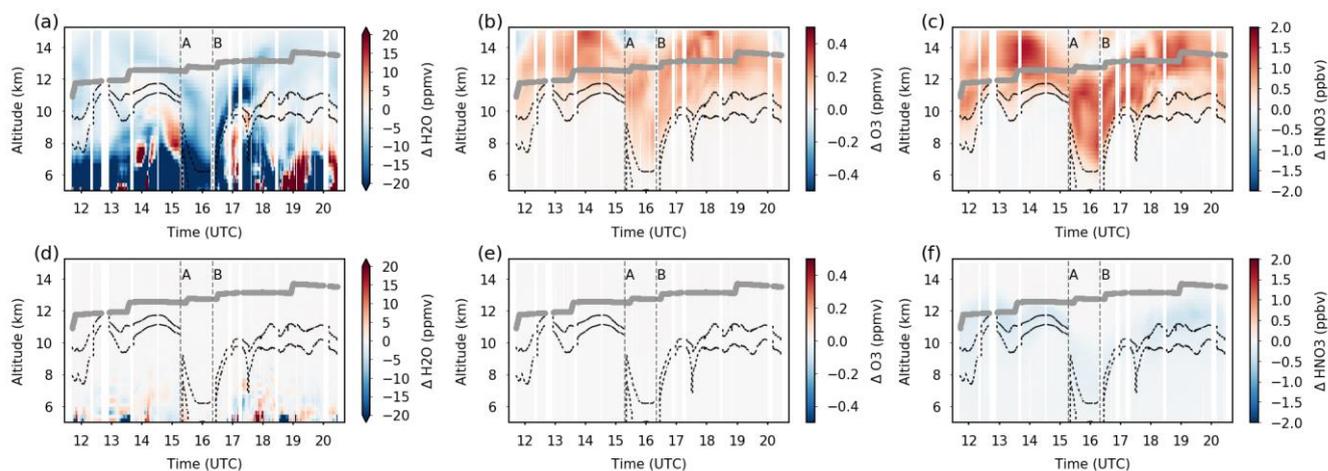
Figure 11. Modelled differences in H$_2$O, O$_3$ and HNO$_3$ due to lower resolution (a-c, T106 vs T42 resolution) and neglecting scavenging processes in clouds (d-f, EMAC-STD minus EMAC-NOSCAV). Black dashed lines: EMAC 2 PVU and 4 PVU isolines (lower and higher lines, respectively) as indicators for the dynamical tropopause. Grey lines: HALO flight altitude.

Finally, the EMAC sensitivity simulations presented in Fig. 11 show that changing the model resolution from T106 to T42 enhances the LMS moist bias in the water vapour distribution (Fig 11a, compare Stenke et al., 2008) and results in significant low biases in the LMS ozone (Fig 11b) and nitric acid distributions (Fig 11c) in the T42 simulation. A similar behaviour is documented in Khosrawi et al. (2017), stating that the T106 simulation agrees slightly better with Aura/MLS observations for both species.

Simulated scavenging processes result in noticeable changes in the LMS only in the case of nitric acid (Fig 11d-f). HNO$_3$ mixing ratios in a band of ~±1 km around the 4 PVU isoline are slightly lower by up to about 0.5 ppbv in the standard simulation. Reminding that EMAC simulates here absolute mixing ratios of ~ 2 ppbv, this suggests that nitric acid is significantly higher in the LMS if scavenging processes by clouds are neglected. Even though HNO$_3$ in EMAC is underestimated in most parts of the LMS (see Sect. 4.2 and above), it is, however, overestimated in most parts of a region between the respective 2 PV- and 4 PV-isoline and ~1 km beneath (see Fig. B 1f and B 1i in Appendix B). This, in turn, means, that HNO$_3$ mixing ratios in the EMAC-simulation are closer to GLORIA measurements in this region, if scavenging processes are considered, and it hints that trapping of HNO$_3$ by high altitude cirrus clouds could play a significant role.

5 Discussion and conclusions

Using GLORIA observations during the HALO long-range flight on 26 February 2016, we test the ability of the ACM ICON-ART and the CCM EMAC to model mesoscale dynamical features, the chemical composition and cirrus clouds and their impacts in the UT/LMS. The flight constitutes a multifaceted test case, covering deeply subsided air masses of the aged 2015/16

5 Arctic vortex, high-latitude LMS air masses, a highly textured troposphere-to-stratosphere exchange region, and high-altitude cirrus clouds.

In both models, even though very different in their character, the dynamical situation, in particular, with the strongly subsided air masses in the western part of the flight, is simulated well. Here, the observed stratospheric air masses, characterised by low water vapour, high ozone and enhanced nitric acid mixing ratios, are reproduced.

10 The high-resolution ICON-ART setup (in a short consecutive forecast mode) involving a R2B7 nest (approx. 20 km) reproduces mesoscale dynamical structures also quite well. The observed troposphere-to-stratosphere exchange connected to an occluded front associated with a strong Icelandic low is clearly reproduced by the model. Given the lower resolution of the nudged T106 simulation of the EMAC model, we find that this model also reproduces these features at the limit of the used model resolution in a very reasonable way.

15 All major cloud systems detected by GLORIA can be identified qualitatively in both models by generated cloud masks from the respective ice water content variables interpolated to the GLORIA geolocations. Remaining discrepancies between GLORIA and the models as well as between the models are reproduced to uncertainties in the modelled geolocations or timing of cloud scenarios as well as limitation of the comparison. We have demonstrated that residuals between the active water vapour tracer and the respective tracer neglecting cloud microphysics in the ICON-ART simulation can be used for an

20 alternative proxy for the presence of clouds, in terms of an integrated picture of the short forecast. In particular, this proxy hinted at a cloud system observed by GLORIA at 14 to 15 UTC, which is not present in the ICON-ART simulation at this particular time. However, a corresponding cloud system is found in the model data a few hours prior to the measurement at this particular geolocation. Both models tend to simulate cloud systems reaching higher above the tropopause than observed by GLORIA and suggest that LMS humidity is significantly affected by cloud microphysics in the simulations. This is

25 supported by the ICON-ART short-term sensitivity forecast neglecting cloud microphysics, which shows that LMS humidity can be depleted locally by cloud processes by 1-2 ppmv within less than 20 hours.

Overall magnitudes of UT/LMS humidity are reproduced well by the consecutive ICON-ART short-term forecasts (reinitialised at 00 UTC with ECMWF IFS) and the continuous simulations of EMAC water vapour. However, a systematic moist bias is found in the LMS in both models. The same moist bias is known for the ECMWF and other weather/atmospheric

30 forecast systems and is a contributing factor to a cold bias there in medium-range forecasts with these systems (Stenke et al., 2008). The fact that both models tend to simulate cirrus clouds reaching higher above the tropopause than observed by GLORIA might be related to the moist bias. Here, enhanced saturation versus the ice phase in the model simulations might be

a reason for the cloud systems reaching to higher altitudes. Consistent with other studies (Roeckner et al., 2006; Polichtchouk et al., 2019), we find a higher moist bias in an EMAC simulation with a lower resolution (T42 instead of T106).

While the overall ozone mixing ratios of EMAC are in good agreement with GLORIA, the simplified ICON-ART O_3 depletion scheme LINOZ and the use of a cold tracer (Braesicke and Pyle, 2003) to imitate heterogeneous chemistry on PSCs

5 systematically overestimate ozone depletion in the LMS by ~0.2 ppmv. This bias might be reduced by tuning of the LINOZ-scheme and/or the threshold temperature and life time of the cold tracer. Furthermore, EMAC nitric acid does not clearly show the observed nitrification of the LMS. This bias has already been documented in Khosrawi et al. (2017, 2018) with comparing EMAC to satellite data. The same problem has been found in a previous study for the same winter using the CLaMS model (Braun et al., 2019) and suggests that microphysical properties of HNO_3-containing particles in polar stratospheric clouds

10 resulting in denitrification of the stratosphere and nitrification of lower layers are not parameterised in a sufficiently realistic way.

We find that LMS composition modelled by EMAC is notably affected by model resolution. In addition to the enhanced moist bias, a change in horizontal resolution from T106 to T42 leads to a low bias in ozone, and an even more pronounced low bias in nitric acid. This effect, concerning ozone and nitric acid, has been also found in Khosrawi et al. (2017), when compared to

15 satellite data. These discrepancies might be overcome by resolution-dependent model tuning. Finally, our EMAC simulations show that neglecting scavenging processes by clouds show practically no impact on water vapour and ozone in the LMS, while nitric acid is noticeably depleted by ~0.5 ppbv if scavenging processes are activated in the simulation.

Overall, we find that ICON-ART and EMAC T106 are well suited for comparison to high resolution remote sensing aircraft data. Fine structures like the troposphere-stratosphere exchange region are reproduced well by ICON-ART and even modelled

20 schematically by EMAC despite the much coarser resolution.

However, we find that accurate simulations of UT/LMS composition remains challenging and both models need to be further improved. We speculate that the reported biases and sensitivities might help to provide better forecasts and long-term projections by these and other models. Furthermore, continuous high resolution measurements of atmospheric trace gases and clouds are required, to continuously test and further improve the models, so that they can be used for reliable projections of

25 temperature trends in the UT/LMS and surface climate.

Appendix A

In this section we want to get back to the comparison of observed clouds by GLORIA and modelled clouds by ICON-ART and EMAC. To prove that seemingly "missing cloud systems" in the ICON-ART model, in particular the cloud system at 14 to 15 UTC, had been present at some time prior to the measurement at the respective geolocations in the model, and to examine

5 the evolution of clouds during the day of PGS Flight 14, we have sampled the model output of ICON-ART cloud variables (q_i and q_s) and the EMAC cloud variables (iwc and cv_iwc) at the GLORIA geolocations, but with negative time offsets varying from -1 to -10 hours.

Figure A1 to A3 show the evolution of clouds in the ICON-ART- (panels d-f) and EMAC-model (panels g-i) between -10 hours to -1 hours prior to the GLORIA measurements interpolated to the GLORIA geolocations which are defined by altitude

10 and time of measurement (in UTC) along the flight. For better comparison Figure 5, which corresponds to no time offset in the models, is again attached in Fig. A3.

The cloud system detected by GLORIA at 14 to 15 UTC corresponds to geolocations along the westward flight leg between central Greenland and approx. the west coast of northern Greenland (see Fig. 1), with GLORIA pointing to the north.
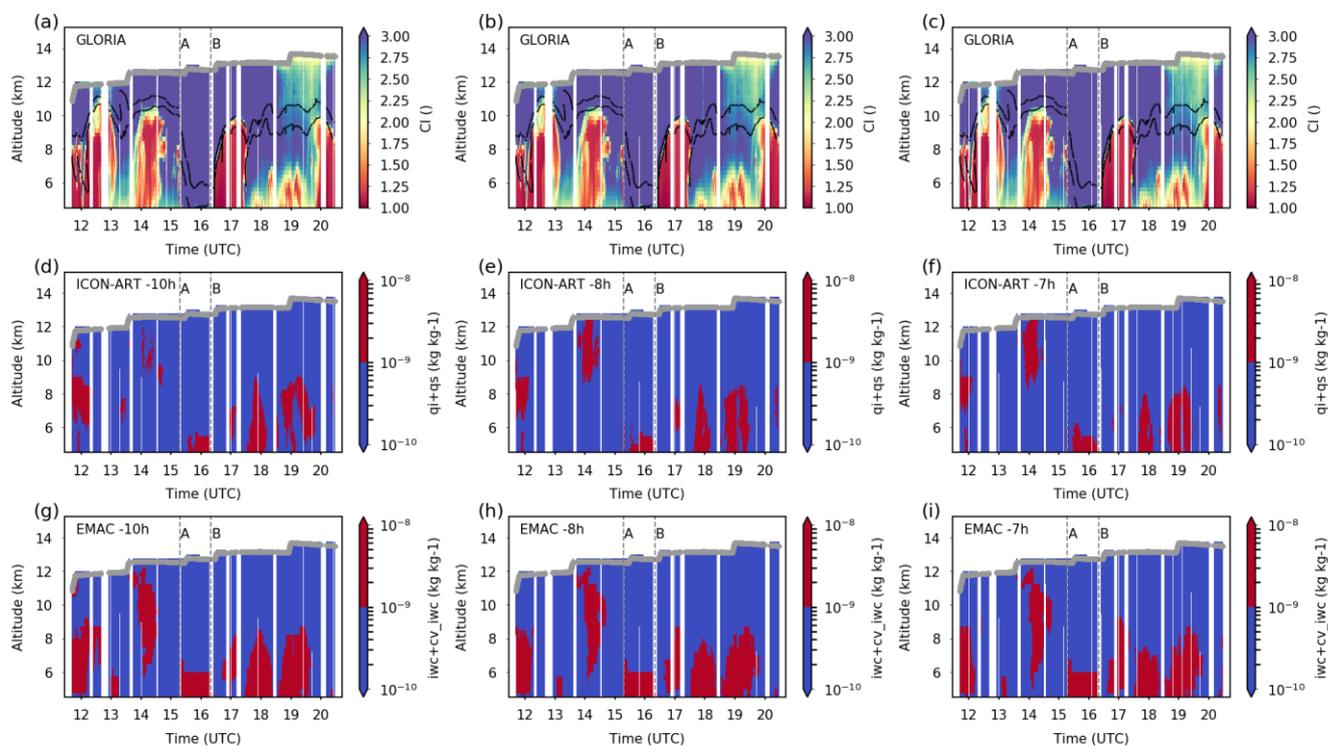
Figure A 1. Same as Figure 5, but model data (ICON-ART and EMAC) have been sampled with a constant time offset of -10, -8 and -7 hours during the interpolation to the GLORIA geolocations along the flight.
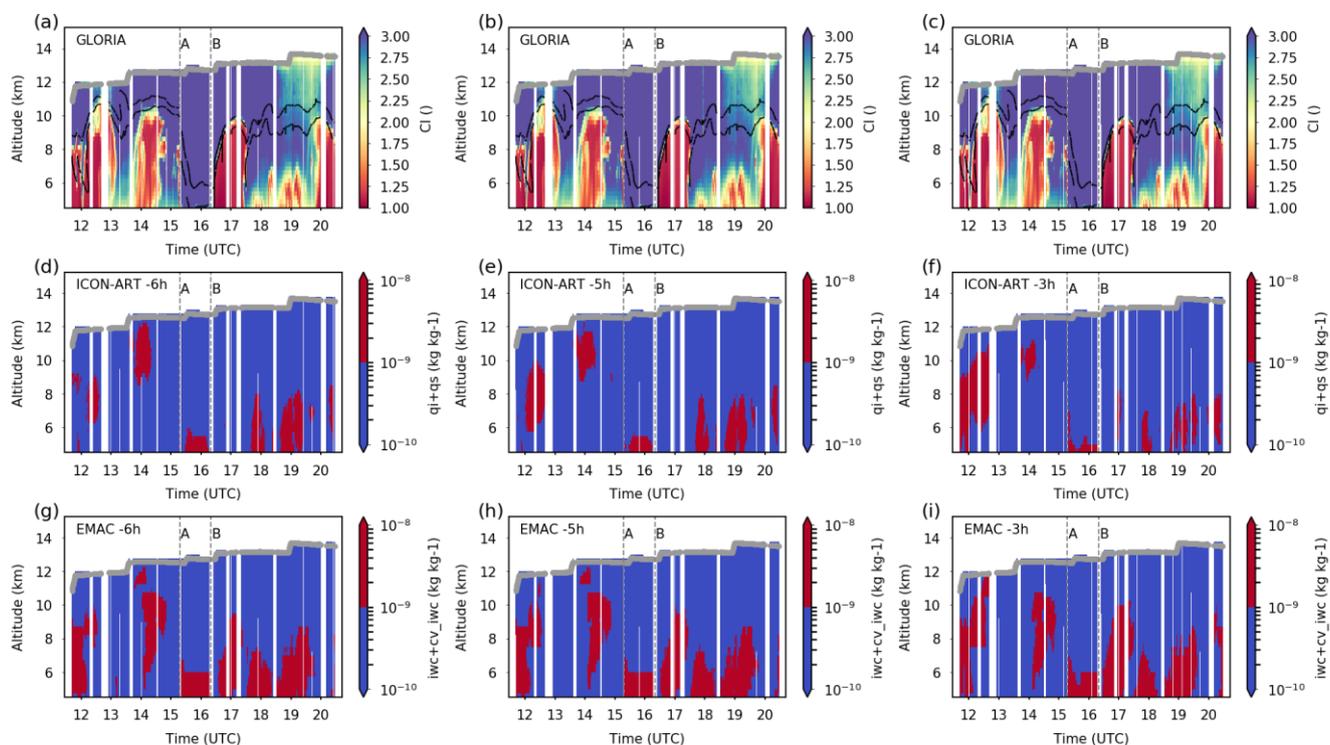
Figure A 2. Same as Figure 5, but model data (ICON-ART and EMAC) have been sampled with a constant time offset of -6, -5 and -3 hours during the interpolation to the GLORIA geolocations along the flight.

Inspection of the panels d-f in Figures A1 to A3 shows that a corresponding cloud system is forming about 10 hours before the measurement in the ICON-ART-model and it is growing until it reaches its maximum vertical and horizontal extent at about a time offset of -7 hours (Fig. A1f). It is also centered higher in the atmosphere than the measured cloud system by GLORIA.

Afterwards (from -6 hours to -2 hours) the cloud system is dissolving while subsiding into supposably warmer layers, until it completely vanishes at -1h hour (Fig. A3e). This proves that a corresponding cloud system is also present in the model data, however it appears a few hours earlier at the particular geolocation.

The corresponding cloud system in the EMAC-simulation (Fig. A1-A3, panels g-i) appears with slightly different shape, but with remarkably larger vertical extent, reaching down deep into the troposphere to about 6 km altitude. It remains approx. constant from -10 to about -6 hours, where it breaks apart into two pieces (Fig. A2g). Afterwards, the upper part is also dissolving and vanishes at about -3 hours (Fig A2i), while the lower part is also subsiding and decreasing in shape to its tropospheric remnants at the time of the GLORIA measurements as depicted in Fig. 5.

In Sect. 4.1 we also found hints that the lower cloud system between 17:30 and 19:30 UTC, which was underestimated in the ICON-ART-cross section, is more pronounced at some time prior to the measurement. Inspection of Figure A1 to A3 yields
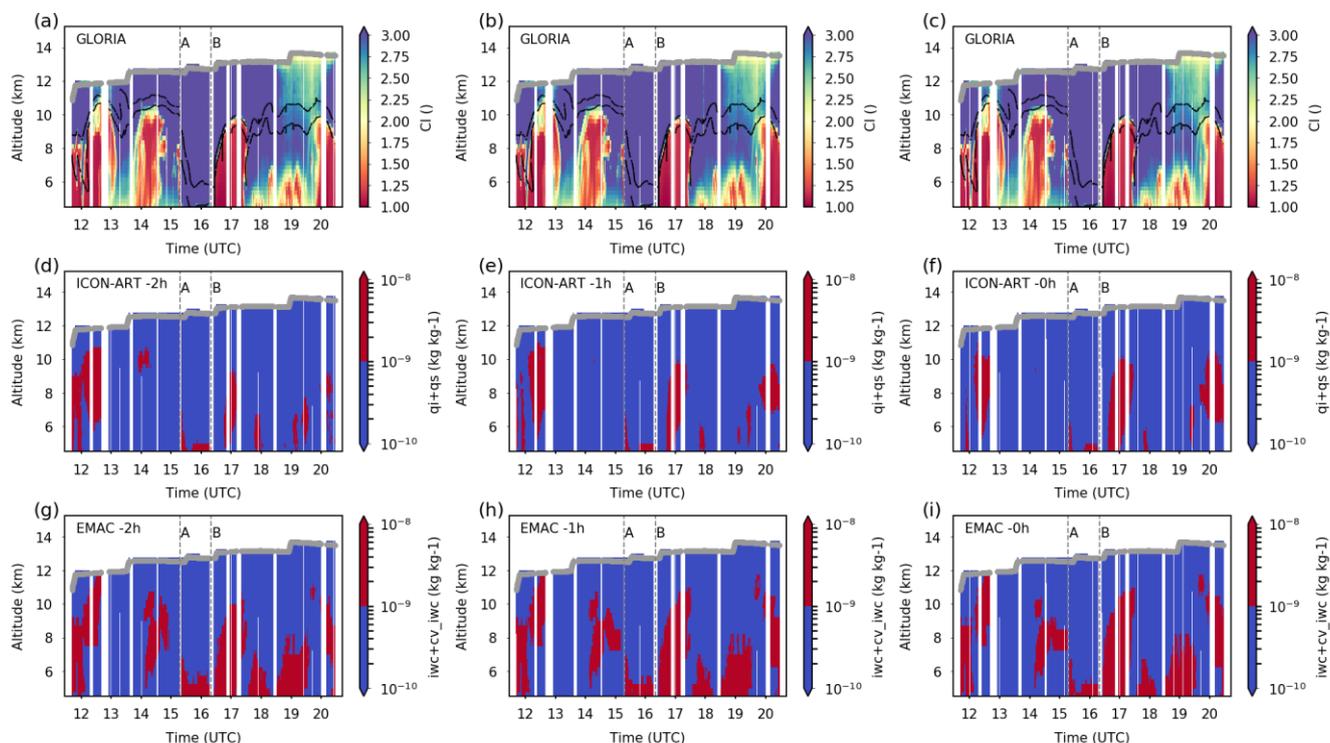
Figure A 3. Same as Figure 5, but model data (ICON-ART and EMAC) have been sampled with a constant time offset of -2, -1 and 0 hours during the interpolation to the GLORIA geolocations along the flight.

that the corresponding cloud system has been more developed at these geolocations during the day of the flight reaching its best resemblance to the GLORIA cloud index around -3 hours prior to the measurement.

However, we do not find any indications in Fig. A1 to A3 in the interpolated ICON-ART-data (panels d-f) for a cloud located at waypoint A around 8 km altitude, which would be responsible for the large precipitation signal in Fig. 6, and which is also

5 visible in the EMAC-data (cf. with Fig. 5c).

In summary, this analysis yields that better resemblance of the ICON-ART cloud data to the GLORIA observations and EMAC-simulations is found in some cases, if model data of an earlier time step is considered.

In particular, the large cloud system observed by GLORIA at 14 to 15 UTC is reproduced in both the ICON-ART and EMAC-model, however its vertical extent is much more pronounced in the EMAC-model.

10 Both models show that this cloud system is subsiding with time, which is in accordance to the meteorological situation above Central Greenland (a high pressure system cf. Sect. 3).

Appendix B
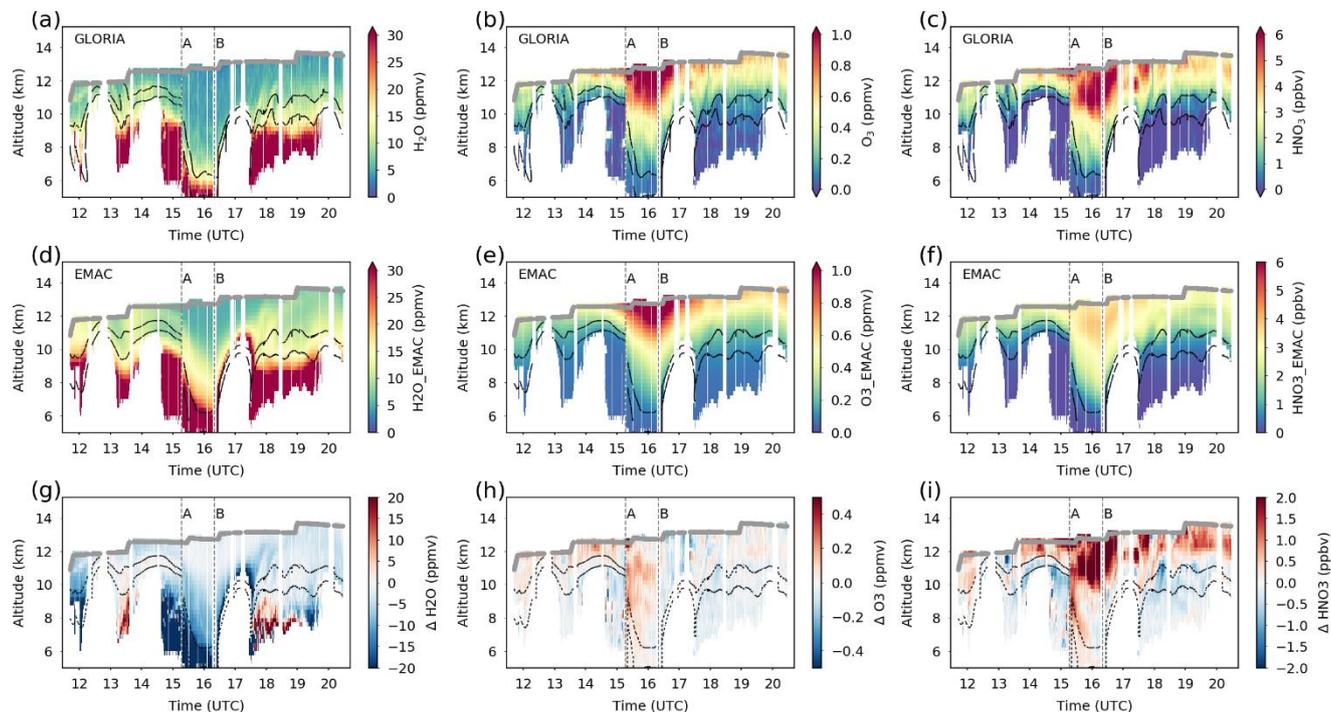
Figure B 1. Observed and modelled trace gas distributions. GLORIA observations of water vapour, ozone and nitric acid (a-c). EMAC continuous simulations of water vapour, ozone and nitric acid (d-f), and respective residuals between GLORIA and EMAC (g-i) Black lines: 2 PVU and 4 PVU isolines (lower and higher lines, respectively) from ECMWF reanalysis (a-c), EMAC (d-i) as indicators for the dynamical tropopause. Grey lines: HALO flight altitude.

Figure B1 shows the GLORIA observations of water vapour, ozone and nitric acid and the corresponding EMAC simulation as in Figure 7. In addition, the residuals between the GLORIA observations and the respective modelled trace gases by EMAC

5 are depicted in panels g-i.

Data availability. The data used here are available at the repository radar4KIT (https://doi.org/10.35097/454, Haenel et al., 2021). The GLORIA observations can also be accessed at the HALO database (https://doi.org/10.17616/R39Q0T, HALO

10 consortium, 2016, last access: 16 April 2020) and at the KITopen repository (https://doi.org/10.5445/IR/1000086506, Johansson et al., 2018b, last access: 16 April 2020). The complete data of the ICON-ART and EMAC simulations are available on the Large Scale Data Facility (LSDF) of SCC. Access can be granted upon request.

Author contributions. FH and WW designed the study. FH and WW analysed the data with support from MH, PB, RR, SJ, AK, MW, OK, FK and JB. JB and MW performed the ICON-ART simulations. OK performed the EMAC simulations. HO, B-MS and WW coordinated the HALO activities during PGS. FF-V and the GLORIA team performed the GLORIA measurements and operations. JU, AK and SJ performed the GLORIA level-1 data analysis. SJ performed the GLORIA

5   chemistry mode trace gas retrievals (level-2 data) used here. MH, AK, JU and SJ contributed to the discussion concerning the GLORIA cloud detection. FH and WW prepared the manuscript with contributions from the other authors. FH, WW, MW and FK designed the figures. All authors helped with discussions and with finalising the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. Atmospheric research with HALO is supported by the Priority Programme SPP 1294 of the Deutsche

10   Forschungsgemeinschaft (DFG). The work of F. Haenel has been funded by the DFG project no. 316735585 (WO 2160/1-1) and by the project "Advanced Earth System Modelling Capacity" (ESM) with project no. ZT-0003 funded by the Helmholtz Association. We thank the GLORIA-Team and DLR-FX for performing the measurements and HALO flights during PGS. The EMAC and ICON-ART simulations were performed on the supercomputer ForHLR and with the help of the Large Scale Data Facility at the Karlsruhe Institute of Technology, both funded by the Ministry of Science, Research and the Arts Baden-

15   Württemberg and by the Federal Ministry of Education and Research. The interpolations of EMAC and ICON-ART simulation data to the GLORIA tangent geolocations were performed on the bwUniCluster (2.0). The authors acknowledge support by the state of Baden-Württemberg through bwHPC. We thank ECMWF for providing the meteorological data used here.

References

20   Barahona, D. and Nenes, A.: Parameterizing the competition between homogeneous and heterogeneous freezing in cirrus cloud formation – monodisperse ice nuclei, Atmos. Chem. Phys., 9, 369-381, doi:10.5194/acp-9-369-2009, 2009.

Bönisch, H., Engel, A., Curtius, J., Birner, Th., and Hoor, P.: Quantifying transport into the lowermost stratosphere using simultaneous in-situ measurements of SF6 and CO2, Atmos. Chem. Phys., 9, 5905-5919, https://doi.org/10.5194/acp-9-5905-2009, 2009.

25   Braesicke, P. and Pyle, J. A.: Changing ozone and changing circulation in northern mid-latitudes: Possible feedbacks? Geophys. Res. Lett., 30, 1059, doi:10.1029/2002GL015973, 2003.

Braun, M., Grooß, J.-U., Woiwode, W., Johansson, S., Höpfner, M., Friedl-Vallon, F., Oelhaf, H., Preusse, P., Ungermann, J., Sinnhuber, B.-M., Ziereis, H., and Braesicke, P.: Nitrification of the lowermost stratosphere during the exceptionally cold Arctic winter 2015–2016, Atmos. Chem. Phys., 19, 13681–13699, https://doi.org/10.5194/acp-19-13681-2019, 2019.

30   Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bid-lot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M.,

Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteorol. Soc., 137, 553–597, 2011.

5   Ehard, B., Malardel, S., Dörnbrack, A., Kaifler, B., Kaifler, N., Wedi, N.: Comparing ECMWF high-resolution analyses with lidar temperature measurements in the middle atmosphere, Q. J. Roy. Meteorol. Soc., 144, 633–640, https://doi.org/10.1002/qj.3206, 2018.

Forster, P. M. de F., and Shine, K. P.: Radiative forcing and temperature trends from stratospheric ozone changes, J. Geophys. Res., 102, 10841-10855, 1997.

10 Forster, P. M. de F., and Shine, K. P.: Assessing the climate impact and its uncertainty for trends in stratospheric water vapor, Geophys. Res. Lett., 29, 6, doi:1029/2001GL013909, 2002.

Friedl-Vallon, F., Gulde, T., Hase, F., Kleinert, A., Kulessa, T., Maucher, G., Neubert, T., Olschewski, F., Piesch, C., Preusse, P., Rongen, H., Sartorius, C., Schneider, H., Schönfeld, A., Tan, V., Bayer, N., Blank, J., Dapp, R., Ebersoldt, A., Fischer, H., Graf, F., Guggenmoser, T., Höpfner, M., Kaufmann, M., Kretschmer, E., Latzko, T., Nordmeyer, H., Oelhaf, H.,

15 Orphal, J., Riese, M., Schardt, G., Schillings, J., Sha, M. K., Suminska-Ebersoldt, O., and Ungermann, J.: Instrument concept of the imaging Fourier transform spectrometer GLORIA, Atmos. Meas. Tech., 7, 3565–3577, https://doi.org/10.5194/amt-7-3565-2014, 2014.

Gettelman, A., Hoor, P., Pan, L. L., Randel, W. J., Hegglin, M. I., and Birner, T.: The extratropical upper troposphere and lower stratosphere, Rev. Geophys., 49, RG3003, https://doi.org/10.1029/2011RG000355, 2011.

20 Haenel, F., Johansson, S., Woiwode, W. and Höpfner, M.: POLSTRACC F14 ICECREAM, Institute of Meteorology and Climate Research, Atmospheric Trace Gases and Remote Sensing (IMK-ASF), Karlsruhe Institute of Technology (KIT), https://doi.org/10.35097/454, 2021.

HALO consortium: HALO database, Registry of Research Data Repositories, https://doi.org/10.17616/R39Q0T, 2016.

Hartmann, D. L., Klein Tank, A. M. G., Rusticucci, M., Alexander, I. V., Brönnimann, S., Charabi, Y., Dentener, F. J., Dlugo-

25 kencky, E. J., Easterling, D. R., Kaplan, A., Soden, B. J., Thorne, P. W., Wild, M., and Zhai, P. M.: Observations: atmosphere and surface, in: Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, UK and New York, NY, USA, 2013.

30 Holton, J. R., Haynes, P. H., McIntyre, M. E., Douglass, A. R., Rood, R. B., and Pfister, L.: Stratosphere-Troposphere Exchange, Rev. Geophys., 33, 403–439, https://doi.org/10.1029/95rg02097, 1995.

Hoor, P., Fischer, H., Lange, L., Lelieveld, J., and Brunner, D.: Seasonal Variations of a mixing layer in the lowermost stratosphere as identified by the CO-O_3 correlation from in-situ measurements, J. Geophys. Res., 107, 4044, doi:10.1029/2001JD000575, 2002.

Hoor, P., Fischer, H., and Lelieveld, J.: Tropical and extratropical tropospheric air in the lowermost stratosphere over Europe: A CO-based budget, Geophys. Res., Lett., 32, L07802, doi:10.1029/2004GL022018, 2005.

Johansson, S., Woiwode, W., Höpfner, M., Friedl-Vallon, F., Kleinert, A., Kretschmer, E., Latzko, T., Orphal, J., Preusse, P., Ungermann, J., Santee, M. L., Jurkat-Witschas, T., Marsing, A., Voigt, C., Giez, A., Krämer, M., Rolf, C., Zahn, A., Engel, A., Sinnhuber, B.-M., and Oelhaf, H.: Airborne limb-imaging measurements of temperature, HNO3, O3, ClONO2, H2O and CFC-12 during the Arctic winter 2015/2016: characterization, in situ validation and comparison to Aura/MLS, Atmos. Meas. Tech., 11, 4737–4756, https://doi.org/10.5194/amt-11-4737-2018, 2018a.

Johansson, S., Woiwode, W., and Höpfner, M.: GLORIA Chemistry Mode temperature and trace gas retrievals from the POLSTRACC/GW-LCYCLEII/GWEX/SALSA campaigns 2015/2016, https://doi.org/10.5445/IR/1000086506, 2018b.

Johansson, S., Santee, M. L., Grooß, J.-U., Höpfner, M., Braun, M., Friedl-Vallon, F., Khosrawi, F., Kirner, O., Kretschmer, E., Oelhaf, H., Orphal, J., Sinnhuber, B.-M., Tritscher, I., Ungermann, J., Walker, K. A., and Woiwode, W.: Unusual chlorine partitioning in the 2015/16 Arctic winter lowermost stratosphere: observations and simulations, Atmos. Chem. Phys., 19, 8311–8338, https://doi.org/10.5194/acp-19-8311-2019, 2019.

Jöckel, P., Tost, H., Pozzer, A., Brühl, C., Buchholz, J., Ganzeveld, L., Hoor, P., Kerkweg, A., Lawrence, M. G., Sander, R., Steil, B., Stiller, G., Tanarhte, M., Taraborrelli, D., van Aardenne, J., and Lelieveld, J.: The atmospheric chemistry general circulation model ECHAM5/MESSy1: consistent simulation of ozone from the surface to the mesosphere, Atmos. Chem. Phys., 6, 5067-5104, doi:10.5194/acp-6-5067-2006, 2006.

Jöckel, P., Kerkweg, A., Pozzer, A., Sander, R., Tost, H., Riede, H., Baumgaertner, A., Gromov, S., and Kern, B.: Development cycle 2 of the Modular Earth Submodel System (MESSy2), Geosci. Model Dev., 3, 717-752, doi:10.5194/gmd-3-717-2010, 2010.

Jöckel, P., Tost, H., Pozzer, A., Kunze, M., Kirner, O., Brenninkmeijer, C. A. M., Brinkop, S., Cai, D. S., Dyroff, C., Eckstein, J., Frank, F., Garny, H., Gottschaldt, K.-D., Graf, P., Grewe, V., Kerkweg, A., Kern, B., Matthes, S., Mertens, M., Meul, S., Neumaier, M., Nützel, M., Oberländer-Hayn, S., Ruhnke, R., Runde, T., Sander, R., Scharffe, D., and Zahn, A.: Earth System Chemistry Integrated Modelling (ESCiMo) with the Modular Earth Submodel System (MESSy, version 2.51), Geosci. Model Dev. Discuss., 8, 8635-8750, doi:10.5194/gmdd-8-8635-2015, 2015.

Kärcher, B., Abbatt, J. P. D., Cox, R. A., Popp, P. J., and Voigt, C.: Trapping of trace gases by growing ice surfaces including surface-saturated adsorption, J. Geophys. Res., 114, doi:10.1029/2009JD011857, 2009.

Kelly, K. K., Tuck, A. F., and Davies, T.: Wintertime asymmetry of upper tropospheric water between the Northern and Southern Hemispheres, Nature, 353, 244-247, 1991.

Khosrawi, F., Kirner, O., Sinnhuber, B.-M., Johansson, S., Höpfner, M., Santee, M. L., Froidevaux, L., Ungermann, J., Ruhnke, R., Woiwode, W., Oelhaf, H., and Braesicke, P.: Denitrification, dehydration and ozone loss during the 2015/2016 Arctic winter, Atmos. Chem. Phys., 17, 12893–12910, https://doi.org/10.5194/acp-17-12893-2017, 2017.

Khosrawi, F., Kirner, O., Stiller, G., Höpfner, M., Santee, M. L., Kellmann, S., and Braesicke, P.: Comparison of ECHAM5/MESSy Atmospheric Chemistry (EMAC) simulations of the Arctic winter 2009/2010 and 2010/2011 with

Envisat/MIPAS and Aura/MLS observations, Atmos. Chem. Phys., 18, 8873–8892, https://doi.org/10.5194/acp-18-8873-2018, 2018.

Kirner, O., Ruhnke, R., Buchholz-Dietsch, J., Jöckel, P., Brühl, C., and Steil, B.: Simulation of polar stratospheric clouds in the chemistry-climate-model EMAC via the submodel PSC, Geosci. Model Dev., 4, 169–182, https://doi.org/10.5194/gmd-4-169-2011, 2011.

Krämer, M., Schiller, C., Voigt, C., Schlager, H., and Popp, P. J.: A climatological view of HNO3 partitioning in cirrus clouds, Q. J. Roy. Meteorol. Soc., 134, 905–912, doi: 10.1002/qj.253, 2008.

Krämer, M., Rolf, C., Spelten, N., Afchine, A., Fahey, D., Jensen, E., Khaykin, S., Kuhn, T., Lawson, P., Lykov, A., Pan, L. L., Riese, M., Rollins, A., Stroh, F., Thornberry, T., Wolf, V., Woods, S., Spichtinger, P., Quaas, J., and Sourdeval, O.: A microphysics guide to cirrus – Part 2: Climatologies of clouds and humidity from observations, Atmos. Chem. Phys., 20, 12569–12608, https://doi.org/10.5194/acp-20-12569-2020, 2020.

Krause, J., Hoor, P., Engel, A., Plöger, F., Grooß, J.-U., Bönisch, H., Keber, T., Sinnhuber, B.-M., Woiwode, W., and Oelhaf, H.: Mixing and ageing in the polar lower stratosphere in winter 2015–2016, Atmos. Chem. Phys., 18, 6057–6073, https://doi.org/10.5194/acp-18-6057-2018, 2018.

Leuenberger, D., Koller, M., Fuhrer, O., and Schär, C.: A Generalization of the SLEVE Vertical Coordinate. Mon. Wea. Rev., 138, 3683–3689, https://doi.org/10.1175/2010MWR3307.1, 2010.

Manney, G. L. and Lawrence, Z. D.: The major stratospheric final warming in 2016: dispersal of vortex air and termination of Arctic chemical ozone loss, Atmos. Chem. Phys., 16, 15371–15396, https://doi.org/10.5194/acp-16-15371-2016, 2016.

Matthias, V., Dörnbrack, A., and Stober, G.: The extraordinarily strong and cold polar vortex on the early northern winter 2015/16, Geophys. Res. Lett., 43, 12287–12294, https://doi.org/10.1002/2016GL071676, 2016.

Meloen, J., Siegmund, P., van Velthoven, P., Kelder, H., Sprenger, M., Wernli, H., Kentarchos, A. S., Roelofs, G. J., Feichter, J., Land, C., Forster, C., James, P., Stohl, A., Collins, W., and Cristofanelli, P.: Stratosphere-troposphere exchange: A model and method intercomparison, J. Geophys. Res., 108, 8526, doi:10.1029/2002JD002274, 2003.

Oelhaf, H., Sinnhuber, B.-M., Woiwode, W., Bönisch, H., Bozem, H., Engel, A., Fix, A., Friedl-Vallon, F., Grooß, J., Hoor, P., Johansson, S., Jurkat-Witschas, T., Kaufmann, S., Krämer, M., Krause, J., Kretschmer, E., Lörks, D., Marsing, A., Orphal, J., Pfeilsticker, K., Pitts, M., Poole, L., Preusse, P., Rapp, M., Riese, M., Rolf, C., Ungermann, J., Voigt, C., Volk, C. M., Wirth, M., Zahn, A., and Ziereis, H.: Polstracc: Airborne Experiment for Studying the Polar Stratosphere in a Changing Climate with the High Altitude and Long Range Research Aircraft (HALO). Bull. Amer. Meteor. Soc., 100, 2634–2664, https://doi.org/10.1175/BAMS-D-18-0181.1, 2019.

Pan, L., Solomon, S., Randel, W., Lamarque, J.-F., Hess, P., Gille, J., Chiou, E.-W., and McCormick, M. P.: Hemispheric asymmetries and seasonal variations of the lowermost stratospheric water vapor and ozone derived from SAGE II data, J. Geophys. Res., 102, 28177–28184, 1997.

Polichtchouk, I., Stockdale, T., Bechtold, P., Diamantakis, M., Malardel, S., Sandu, I., Vána, F., and Wedi, N.: Control on stratospheric temperature in IFS: resolution and vertical advection, ECMWF Technical Memorandum no 847, doi: 10.21957/cz3t12t7e, 2019.

Popp, P. J., Gao, R. S., Marcy, T. P., Fahey, D. W., Hudson, P. K., Thompson, T. L., Kärcher, B., Ridley, B. A., Weinheimer, A. J., Knapp, D. J., Montzka, D. D., Baumgardner, D. G., Garrett, T. J., Weinstock, E. M., Smith, J. B., Sayres, D. S., Pittman, J. V., Dhaniyala, S., Bui, T. P., and Mahoney, M. J.: Nitric acid uptake on subtropical cirrus cloud particles, J. Geophys. Res., 109, D06302, doi:10.1029/2003JD004255, 2004.

Ray, E. A., Moore, F. L., Elkins, J. W., Dutton, G. S., Fahey, D. W., Vömel, H., Oltmans, S. J., and Rosenlof, K. H.: Transport into the Northern Hemisphere lowermost stratosphere revealed by in situ tracer measurements, J. Geophys. Res., 104(D21), 26565-26580, doi:10.1029/1999JD900323, 1999.

Rieger, D., Bangert, M., Bischoff-Gauss, I., Förstner, J., Lundgren, K., Reinert, D., Schröter, J., Vogel, H., Zängl, G., Ruhnke, R., and Vogel, B.: ICON-ART 1.0 – a new online-coupled model system from the global to regional scale, Geosci. Model Dev., 8, 1659–1676, https://doi.org/10.5194/gmd-8-1659-2015, 2015.

Riese, M., Ploeger, F., Rap, A., Vogel, B., Konopka, P., Dameris, M., and Forster, P.: Impact of uncertainties in atmospheric mixing on simulated UTLS composition and related radiative effects, J. Geophys. Res., 117, D16305, doi:10.1029/2012JD017751, 2012.

Riese, M., Oelhaf, H., Preusse, P., Blank, J., Ern, M., Friedl-Vallon, F., Fischer, H., Guggenmoser, T., Höpfner, M., Hoor, P., Kaufmann, M., Orphal, J., Plöger, F., Spang, R., Suminska-Ebersoldt, O., Ungermann, J., Vogel, B., and Woiwode, W.: Gimballed Limb Observer for Radiance Imaging of the Atmosphere (GLORIA) scientific objectives, Atmos. Meas. Tech., 7, 1915–1928, https://doi.org/10.5194/amt-7-1915-2014, 2014.

Roeckner, E., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kornblueh, L., Manzini, E., Schlese, U., and Schulzweida, U.: Sensitivity of simulated climate to horizontal and vertical resolution in the ECHAM5 atmosphere model, J. Climate, 19, 3771–3791, https://doi.org/10.1175/jcli3824.1, 2006.

Rosenlof, K. H., Tuck, A. F., Kelly, K. K. Russel, J. M., and Mc-Cormick, M. P.: Hemispheric asymmetries in water vapor and inferences about transport in the lower stratosphere, J. Geophys. Res., 102(D11), 13213–13234, doi:10.1029/97JD00873, 1997.

Sander, R., Baumgaertner, A., Gromov, S., Harder, H., Jöckel, P., Kerkweg, A., Kubistin, D., Regelin, E., Riede, H., Sandu, A., Taraborrelli, D., Tost, H., and Xie, Z.-Q.: The atmospheric chemistry box model CAABA/MECCA-3.0, Geosci. Model Dev., 4, 373–380, https://doi.org/10.5194/gmd-4-373-2011, 2011.

Sander, R., Jöckel, P., Kirner, O., Kunert, A. T., Landgraf, J., and Pozzer, A.: The photolysis module JVAL-14, compatible with the MESSy standard, and the JVal PreProcessor (JVPP), Geosci. Model Dev., 7, 2653–2662, https://doi.org/10.5194/gmd-7-2653-2014, 2014.

Schiller, C., Krämer, M., Afchine, A., Spelten, N., and Sitnikov, N.: The ice Water Content in Arctic, Midlatitude and Tropical Cirrus, J. Geophys. Res., 113, D24208, doi:10.1029/2008JD010342, 2008.

Schröter, J., Rieger, D., Stassen, C., Vogel, H., Weimer, M., Werchner, S., Förstner, J., Prill, F., Reinert, D., Zängl, G., Giorgetta, M., Ruhnke, R., Vogel, B., and Braesicke, P.: ICON-ART 2.1: a flexible tracer framework and its application for composition studies in numerical weather forecasting and climate simulations, Geosci. Model Dev., 11, 4043–4068, https://doi.org/10.5194/gmd-11-4043-2018, 2018.

5   Shine, K. P., Bourqui, M. S., Forster, P. M. F., Hare, S. H. E., Langematz, U., Braesicke, P., Grewe, V., Ponater, M., Schnadt, C., Smith, C. A., Haigh, J. D., Austin, J., Butchart, N., Shindell, D. T., Randel, W. J., Nagashima, T., Portmann, R. W., Solomon, S., Seidel, D. J., Lanzante, J., Klein, S., Ramaswamy, V., and Schwarzkopf, M. D.: A comparison of model-simulated trends in stratospheric temperatures, Q. J. R. Meteorol. Soc., 129, 1565–1588, https://doi.org/10.1256/qj.02.186, 2003.

10 Solomon, S., Rosenlof, K. H., Portmann, R. W., Daniel, J. S., Davis, S. M., Sanford, T. J., and Plattner, G.-K.: Contributions of Stratospheric Water Vapor to Decadal Changes in the Rate of Global Warming, Science, 327, 1219–1223, 2010.

Spang, R., Remedios, J. J., and Brarkley, M. P.: Colour indices for the detection and differentiation of cloud types in infrared limb emission spectra, Adv. Space Res., 33, 1041–1047, https://doi.org/10.1016/S0273-1177(03)00585-4, 2004.

Spang, R., Günther, G., Riese, M., Hoffmann, L., Müller, R., and Griessbach, S.: Satellite observations of cirrus clouds in the
15 Northern Hemisphere lowermost stratosphere, Atmos. Chem. Phys., 15, 927–950, https://doi.org/10.5194/acp-15-927-2015, 2015.

Stenke, A., Grewe, V., and Ponater, M.: Lagrangian transport of water vapour and cloud water in the ECHAM4 GCM and its impact on the cold bias, Clim. Dynam., 31, 491–506, doi:10.1007/s00382-007-0347-5, 2008.

Stohl, A., Bonasoni, P., Cristofanelli, P., Collins, W. J., Feichter, J., Frank, A., Forster, C., Gerasopoulos, E., Gäggeler, H.,
20 James, P., Kentarchos, T., Kromp-Kalb, H., Krüger, B., Land, C., Meloen, J., Papayannis, A., Priller, A., Seibert, P., Sprenger, M., Roelofs, G. J., Scheel, H. E., Schnabel, C., Siegmund, P., Tobler, L., Trickl, T., Wernli, H., Wirth, V., Zanis, P., and Zerefos, C.: Stratosphere-troposphere exchange: A review, and what we have learned from STACCATO, J. Geophys. Res., 108, 8516, doi:10.1029/2002JD002490, 2003.

Tost, H., Jöckel, P., Kerkweg, A., Sander, R., and Lelieveld, J.: Technical note: A new comprehensive SCAVenging submodel
25 for global atmospheric chemistry modelling, Atmos. Chem. Phys., 6, 565–574, doi:10.5194/acp-6-565-2006, 2006a.

Tost, H., Jöckel, P., and Lelieveld, J.: Influence of different convection parameterisations in a GCM, Atmos. Chem. Phys., 6, 5475–5493, doi:10.5194/acp-6-5475-2006, 2006b.

Voigt, C., Schlager, H., Ziereis, H., Kärcher, B., Luo, B. P., Schiller, C., Krämer, M., Popp, P. J., Irie, H., and Kondo, Y.: Nitric acid uptake in cirrus clouds, Geophys. Res. Lett., 33, L05803, doi:10.1029/2005GL025159, 2006.

30 Weimer, M., Schröter, J., Eckstein, J., Deetz, K., Neumaier, M., Fischbeck, G., Hu, L., Millet, D. B., Rieger, D., Vogel, H., Vogel, B., Reddmann, T., Kirner, O., Ruhnke, R., and Braesicke, P.: An emission module for ICON-ART 2.0: implementation and simulations of acetone, Geosci. Model Dev., 10, 2471-2494, doi: 10.5194/gmd-10-2471-2017, 2017.

Weimer, M., Buchmüller, J., Hoffmann, L., Kirner, O., Luo, B., Ruhnke, R., Steiner, M., Tritscher, I., and Braesicke, P.: Mountain-wave-induced polar stratospheric clouds and their representation in the global chemistry model ICON-ART, Atmos. Chem. Phys., 21, 9515–9543, https://doi.org/10.5194/acp-21-9515-2021, 2021.

Werner, A., Volk, C. M., Ivanova, E. V., Wetter, T., Schiller, C., Schlager, H., and Konopka, P.: Quantifying transport into the Arctic lowermost stratosphere, Atmos. Chem. Phys., 10, 11623-11639, https://doi.org/10.5194/acp-10-11623-2010, 2010.

Zängl, G., Reinert, D., Rípodas, P., and Baldauf, M.: The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core, Q. J. Roy. Meteorol. Soc., 141, 563–579, https://doi.org/10.1002/qj.2378, 2015.