

Replies to review comments

We thank the reviewers, the co-editor, as well as Cameron Homeyer and his students for the thoughtful comments and the time and effort spent on the manuscript. Please find our point-by-point replies below (in blue color and italics). A revised manuscript with tracked changes was attached.

In this initial reply we would like to address the main issue regarding the assimilation of the Concordiasi data into the (re)analyses.

Following the first review comment by Andreas Stohl, we contacted scientists and support staff at ECMWF, NASA, and NCEP/NCAR to clarify the question if the Concordiasi data have been assimilated by the respective centres.

Unfortunately and unexpectedly, it turned out that this was case for the ECMWF data sets and MERRA, but not for NCEP/NCAR. In particular, we learned that 15-min time averaged data from the Concordiasi balloons have been transmitted over the Global Telecommunication System (GTS) (Rabier, 2013). The data transmitted over GTS were then assimilated by the respective centers.

We agree that the Concordiasi data can not be considered as an independent data source for direct validation of the large-scale state of the ECMWF products and MERRA. They can only be used for validation of NCEP/NCAR. Furthermore, we consider the data useful for the assessment of small-scale structure (e. g., gravity waves), because the Concordiasi data have been subject to downsampling and data thinning before they were assimilated.

Meteorological analyses are a result of combining various observations (satellite and in-situ), a forecast model, and a data assimilation procedure. A comparison with the assimilated balloon data does not provide validation in a strict sense, but it still provides information regarding the performance of the overall system. Our study showed notable differences between the ECMWF operational analysis, ERA-Interim, and MERRA, despite the fact that the balloon data have been assimilated.

After seeking consent with the co-editor, we therefore conducted a major revision of the manuscript. In particular, we made the following changes:

- 1. We added a new paragraph in Sect. 2.2 describing the role of data assimilation in this study.*
- 2. We replaced the term “validation” by “comparison” or “evaluation” throughout the manuscript.*
- 3. We significantly shortened Sect. 3.1 of the paper that dealt with the “direct validation” of the meteorological analysis.*
- 4. We removed some of the comparisons to other balloon campaigns (Vorcore and Pre-Concordiasi), where data have not been considered for data assimilation.*

Reviewer #1

Summary: This study provides validation of four lower stratospheric temperature and wind analyses through comparisons with independent (not incorporated into the analyses) long duration constant pressure balloon observations from Sep 2010 to Dec 2011. In addition to direct temperature and wind comparisons, the study compares trajectory errors, a quantity that is important to understand for the interpretation of stratospheric transport and chemistry. The observations, analyses, and methods are clearly presented and the results clearly explained. Past work is appropriately referenced, figures are clear, and the paper is very well written. The balloons provide an excellent source of independent, in situ, stratosphere observations, that are well exploited by this study. Overall this work should be of interest to many readers of ACP.

Main Point:

MERRA-2 was released in the Fall of 2015 and should really be included in place of, or better yet, in addition to the MERRA results. MERRA-2 replaces MERRA and will be used in more future studies than the previous, and no longer produced, MERRA system. Including MERRA-2 comparisons should make the paper much more useful and more widely cited. It is difficult to recommend publications as is, with only the out-of-date MERRA system included.

MERRA-2 data were quite new when we conducted the study, but we included them in this revision.

Minor Points:

Table 2: The MERRA and MERRA-2 products are also provided on the 72 model vertical levels. The higher vertical resolution available may change the balloon comparison results. Was there a reason for examining the pressure level output?

Following Hoffmann et al. (2016), we initially focused on MERRA output on pressure levels, because the MPTRAC model uses pressure as vertical coordinate. However, for the assessment of MERRA-2, we implemented new code to process meteorological data on hybrid sigma/pressure levels.

Figure 4, Caption: Add text for the latitude of the outer circle and the longitude orientation.

We added corresponding text in the caption.

Line 23, "40 m/s": There are still some 50 m/s contours in the 2010-12-01 figure. Is the 40 m/s number an average?

This was an error in the text. We replaced "early December" by "mid of December".

Line 21, "grid-scale variances". How are the grid-scale variances calculated? Are they space or time variances?

For each particle position the grid-scale variance was calculated considering values of the 8 nearest grid points and 2 nearest time steps of the meteorological data. We rephrased the text to clarify.

Figures 5 and 6: The meridional wind bias plot might show differences more clearly with a different vertical scale. The values are small and the circulation will change over the balloon record and with latitude, however the small average meridional wind error is still of interest.

We adjusted the vertical scale to make small values visible.

Figures in general: The multi panel figures should be labelled a, b, c, ... and referred to as such in the figure captions and text.

These labels will be inserted during the copy-editing process as needed.

Reviewer #2 (A. Stohl)

This is an interesting study that compares meteorological data from several (re)analysis data sets to balloon measurements in the stratosphere. While the results are perhaps not entirely unexpected and the methods applied are not particularly novel, the study is solid and it is always good to see validation of re-analysis data (and trajectory calculations based on them) against independent data. The paper would, however, benefit from some shortening of the Results section (see also comments below). Most of the results are relatively easy to interpret and do not need to be discussed in such great detail as there is now – the figures are to a large extent self-explanatory, given that the analysis methods and statistical quantities presented are relatively simple. I have some more detailed comments below but if they are addressed adequately, I am in favor of publication of this paper.

Major points to consider:

Please confirm explicitly (both in your response and in the paper) that the Concordiasi balloon data was not assimilated in any of the data sets that you are using. I assume this is the case but if the data were assimilated, of course your study would not be very meaningful as the data could not be considered independent.

Please see initial reply regarding the role of data assimilation.

Section 3.1: What is the motivation for comparing the (re)analysis data sets against smoothed balloon data rather than against the unfiltered data? This reduces the relative differences between the high-resolution operational analyses and the coarser resolution re-analysis data because the operational data captures some of the high resolution variability, while the coarser-resolution re-analyses capture very little of it. Thus, by filtering the performance of the higher-resolution data sets is artificially degraded (relative to the other data sets), and that doesn't seem to be very logical. I think it would be much more

meaningful to do the standard/main/reference comparisons against the unfiltered data and use the filtering as a sensitivity study to show that the poorer performance of the coarser-resolution data sets is due to their inability to capture some of the fine-resolution details.

The filter was applied to achieve a consistent separation between large-scale dynamics (e. g., zonal temperature gradients and planetary waves) and small-scale features (mainly gravity waves). The cut-off period of 15 h was selected to cover the longest possible periods of inertial gravity waves at high latitudes. Please note that the application of a low-pass filter for detrending is a standard technique for gravity wave analyses. We think that such a scale separation is useful, because we may expect that large-scale features are significantly affected by data assimilation of the 15 min-downsampled balloon data, whereas for the small-scale features this is less relevant. We rephrased the text to clarify.

The results section is in many ways too detailed. For example, in section 3.1 (but also in other sections) there are too many numbers that the reader can't all remember. These numbers are all available in Figures and if summaries are needed, this information could be put in tables. However, putting so many numbers into the text, makes it very exhausting to read. I would suggest to substantially shorten this by concentrating on the main findings/messages and the conclusions that can be drawn from these results, rather than listing all individual results. References to the figures should be enough.

Section 3.1 was significantly shortened as many of the results were related to comparisons of large-scale dynamics of the reanalysis and the balloon data. Those results have been shortened or removed, because the data are not independent in most cases. Instead, we focus on the assessment of representation of small-scale structures in the analyses due to gravity waves.

Section 3.2: Here, it is stated that the best results are found when constraining the vertical position of the trajectories to the actual pressure heights of the balloons. But this should not be presented as a finding, as it is a trivial result. Anything else than that would indicate some error in the calculation! Notice also that this method has been used before and this may be mentioned. I think the first publication of this was by Baumann and Stohl (1997) but there have been other uses, e.g. Riddle et al. (2006). Related to this, it is also a trivial result that the AVTDs of the isopycnic approach increase with time (page 10, line 25). This does not indicate any real errors, but just shows the fact that the trajectory height is not constrained to the balloon altitude, and since the model does not (cannot) account for altitude variations of the balloon, of course there are errors.

Earlier studies using stratospheric superpressure balloon observations (e. g. Hertzog et al., 2004; Boccara et al., 2008) used only the isopycnic approach for trajectory evaluation. Our study shows for the first time that the isopycnic approach provides very similar results compared to trajectories constrained by pressure observations for this type of balloon. The studies of Baumann and Stohl (1997) and Riddle et al. (2006) refer to other types of altitude-controlled balloons, which are operating in the lower troposphere. However, we

rephrased the text in Sect. 3.2 and added the references to clarify.

Section 3.4: It would be interesting to calculate the AHTDs also for the ensemble-mean trajectory of all the trajectories with superimposed subgrid-scale fluctuations. Are the AHTDs for this trajectory larger than for the reference trajectory without sub-grid wind fluctuations?

The AHTDs for both cases are very similar, but not necessarily always larger if we consider superimposed subgrid-scale fluctuations. Because the AHTDs are so similar, we decided to show a plot of the differences in Fig. 10.

Minor points:

Introduction, first few lines: When speaking of the polar vortex, make clear you mean the southern hemisphere.

We clarified this in the revised manuscript.

Page 3, line 33: you say that GPS positions are recorded at each observation time. I suppose this means every 30 seconds, but where you write this, it is not really clear, as you mention the other observations only later. So explicitly say how often GPS data are recorded.

The GPS positions were recorded every 60s and interpolated to 30s time intervals to combine them with the Tsen data. We added this information in the text.

Figures 5 and 6: The ranges used for the y axes are not ideal. This is extreme for the upper right panels (meridional wind bias) where the data range occupies clearly less than 10% of the available space. This makes it very difficult to actually read the values and makes the figure almost meaningless.

We adjusted the y axes ranges to make small values visible.

Is figure 10 really needed?

Considering the information already given in the text, Figure 10 was somewhat redundant. We removed it from the revised manuscript.

Typos, etc.:

Page 4, line 23: evelopE

Fixed.

Page 12, line 26: ... because theY are ...

Fixed.

References:

Baumann, K., and A. Stohl (1997): Validation of a long-range trajectory model using gas balloon tracks from the Gordon Bennett Cup 95. J. Appl. Meteor. 36, 711-720.

Riddle, E. E., P. B. Voss, A. Stohl, D. Holcomb, D. Maczka, K. Washburn, and R. W. Talbot (2006): Trajectory model validation using newly developed altitude-controlled balloons during the International Consortium for Atmospheric Research on Transport and Transformations 2004 campaign. *J. Geophys. Res.* 111, D23S57, doi:10.1029/2006JD007456.

Reviewer #3

The manuscript is generally well written with nice figures and a clear presentation of the methods applied. But unfortunately I consider the applied method to be flawed. By having access to the operational ECMWF analysis feedback data and to the ECMWF ERA-Interim analysis feedback archive, I can confirm that the Concordiasi temperature and wind observations were assimilated by both data assimilation systems. ECMWF assimilated the data distributed on the GTS (15 minutes frequency). 68% of the data was assimilated and 32% removed by thinning.

The ERA-Interim statistics for the whole Concordiasi campaign showed balloon data minus analysis departures for temperature: Standard deviation 0.7K, bias -0.3K. Compared against 12-hour background fields the values were: Standard deviation 0.8K, bias -0.5K. The similar statistics for zonal wind: analysis departures: Standard Deviation 1.2 m/s, bias -0.1 m/s. Background departures: Standard deviation 1.9 m/s, bias -0.1 m/s. Meridional wind: analysis departures: standard deviation 1.2 m/s, bias 0 m/s. Background departures: Standard deviation: 1.9 m/s, bias -0.1 m/s. Around 41000 temperature Concordiasi measurements were assimilated during the three months. 41500 zonal and 41500 Concordiasi meridional measurements were assimilated.

These detailed statistics are included here to confirm that the Concordiasi data was fitted well by the ERA-Interim analysis and therefore cannot be considered independent data. Similarly can be said for the operational ECMWF analysis (not shown). This means that this is not a valid comparison of the four (re)analysis systems, if it is true the data was not assimilated in NCEP reanalysis and the MERIS reanalysis. This means that the core part of the manuscript, the inter-comparison, would not make much sense and would not be fair. Based on this I would recommend that the editor rejects the paper.

Additional costly assimilation experiments without assimilation of the Concordiasi in the ECMWF systems would be required for a fair comparison. It would require a very significant rewrite of the manuscript to remove all the parts that relates to inter-comparison, or clearly split the description and evaluation of the ECMWF systems results and NCEP/MERIS. No matter what it would not provide a proper inter-comparison.

At this stage it does not make sense for me to provide detailed comments. The two main issues I have are related to use of interpolated model data and the 15 hour time filtering.

Please see initial reply regarding the role of data assimilation.

Short comment by Cameron Homeyer et al.

Disclaimer: This is a summary of a group peer review exercise in my senior undergraduate research class at the University of Oklahoma. 39 students participated in this review.

The authors present a validation and transport analysis of multiple large-scale models (both operational and reanalysis) using a (hopefully independent) set of long-duration stratospheric balloon observations over the Antarctic. Errors in stratospheric temperatures and winds are examined along the path of many balloon flights and found to be dependent on latitude and model grid resolution. Through the use of trajectory calculations with balloon locations as initial particle conditions, the authors identify errors in transport calculations using the model wind fields through comparisons with the observed balloon paths. Horizontal displacement errors in the trajectory calculations are found to scale considerably with grid resolution. Furthermore, multiple trajectory calculations for differing sources of vertical motion are calculated and show that horizontal and vertical displacement errors of trajectories also depend significantly on this choice.

Part of this study is a demonstration that the authors' recently developed trajectory model MPTRAC produces reasonable results, while the main focus is on comparisons between observed balloon flight paths and trajectory calculations driven by wind fields from models with varying complexity and grid resolution. One of the strengths of this manuscript is the quality of the figures included. We find that while the paper is generally well written, there are some areas of the technical description and analysis that are unclear or too vague. In some cases, this casts doubt on the results. Detailed comments are provided below.

General Comments

1. On the calibration of the balloon temperature sensor and its accuracy: It is mentioned in Section 2.1 that due to daytime heating by the sun, data from the thermistors on the balloon undergo an empirical correction. However, no detail on exactly how the data are corrected and how this impacts the uncertainty of the measurement is given. Precision of the temperature data set is given, but it seems that understanding its uncertainty and the effect of the empirical correction on the model validation carried out here are required.

For solar zenith angles $\alpha \leq 94.5^\circ$ (daytime measurements) the corrected temperature T_c is deduced from the raw measurement T through

$$T_c = T - A \exp\left(\frac{\alpha - 94.5^\circ}{B}\right),$$

with empirical coefficients A and B for each thermistor. Nighttime measurements ($\alpha > 94.5^\circ$) are not corrected, i. e., $T_c = T$. For further details we would refer to the detailed

description provided by Hertzog et al. (2004), which is cited in our manuscript. Note that we conducted a cross check by calculating day- and nighttime statistics separately, but did not find any significant differences (not shown in the paper). This suggests that the empirical temperature correction does not introduce any large uncertainties.

2. Euclidean distances are used to determine horizontal displacement errors in the trajectories, but is this an appropriate choice? Since the curvature of zonal and meridional winds is most pronounced at the pole, shouldn't distances be calculated using a geodesic approach? Using a Euclidean approach may introduce unwanted errors that bias the results.

The approach used here approximates spherical distances with $\geq 99\%$ accuracy for distances up to 3000 km (Rößler et al., 2017). Therefore, no significant biases were introduced in our results.

3. There is a substantial amount of unnecessary detail in the abstract, much of which (including lists of numerical values) seems better left to the main sections and tables of the paper.

We tried to shorten the abstract in the revised manuscript, but we feel that numerical values should still be presented, because they summarize most of the rather detailed statistics presented in the paper.

4. While the polar vortex was used to motivate this work, it would be nice to see some connection between the findings of this study and the polar vortex in Section 4. For example, how might the results from this work be leveraged in future studies examining dynamics and transport in the vortex?

Following a suggestion of the co-editor, we revised Sect. 4 in order to better relate this work to studies of Lawrence et al. (2015) and Manney et al. (1996, 2003, 2005), which deal with the evaluation of reanalyses in the polar stratosphere.

5. Is the Concordiasi dataset independent of those used for assimilation in the set of models analyzed here? This is an important point that was not discussed in the manuscript.

Please see initial reply regarding the role of data assimilation.

Specific Comments

Page 4, line 23: "envelop" should be "envelope"

Fixed.

Page 5, line 10: For clarity, it would be good to point out that the QBO positive phase is westerly here.

We added this information.

Page 8, lines 20-21: "deviations do not" should be "deviations either do not"

Fixed.

Page 11, line 18: Change "get" to "become"

Fixed.

Page 12, line 26: "because the are generally" should be "because they are generally"

Fixed.

Page 13, line 21: Change "but only by 15" to "but only of 15"

Fixed.

Figure 1: The objective of the colored balloon path is appreciated, but a scale should be given so the reader knows how changes in color correspond to changes in time.

We added a color scale.

Figure 3: While this figure is not a leading element of the analysis, it would be good to provide more detail on the dataset this is based on than including a link in the caption.

We added the information that this analysis is based on MERRA-2.

Figure 4: While the figure caption states the dataset used for this sequence of maps is ERA-Interim, it would be good to specify this in the corresponding text.

As this figure is not dealing with an intercomparison of the different meteorological data sets, we mention the name of the specific data set only in its caption.

Co-editor comment (F. Khosrawi)

Dear authors and referees,

I would like to thank the referees for their thorough review and the authors for their clarifications. The fact that Concordiasi balloon data has been assimilated into the meteorological analyses cannot be ignored. However, in my opinion this does not justify a rejection. I am quite confident that the study can be brought into a publishable form with major revisions.

The authors presented already some ideas how they could improve their manuscript. Most important is that the fact that the Concordiasi balloon data is assimilated in the meteorological analyses is considered when performing the assessment and drawing conclusions. Nevertheless, the Concordiasi data are not the only data that is assimilated and one should not forget that meteorological analyses are based on model simulations. Thus, even with the Concordiasi data assimilated into the analyses the impact cannot be that severe that one cannot do a meaningful assessment of the performance of the meteorological data sets.

Please see initial reply regarding the role of data assimilation.

Contrary to the suggestion by the authors to remove the NCAR/NCEP I would suggest to keep this data set to have one "independent" data set in the comparison. To make the assessment then more concise the section could be split into two comparisons: one between NCEP and Concordiasi and another one comparing the Concordiasi data with ECMWF OP, ERA-Interim, MERRA and MERRA-2. A second option would be to just compare the meteorological analyses without comparing these to the Concordiasi balloon data. Further, I would appreciate if MERRA would not just be replaced with MERRA-2, but rather that both data sets would be used in the assessment. Another third option would be to include another independent data set into the comparison.

We followed the advise and kept NCEP/NCAR data in the paper and added MERRA-2 in addition to MERRA. We decided to not split the analysis in two parts (independent versus dependent data), but tried to make very clear in the discussion which parts are affected by data assimilation.

Another point that could be improved is the references to previous studies. There are a lot of studies comparing the performance of meteorological analyses by Gloria Manney and her colleagues (see list below). Additionally, I would suggest to change the term validation in evaluation or assessment throughout the manuscript.

We tried to improve the discussion in Sect. 4 by relating our work to these references. We also rephrased the term "validation" as suggested.

Based on the suggestions for improvements given by the referees, by myself and by the authors themselves I would like to encourage the authors to conduct major revisions and resubmit their manuscript.

Best regards Farahnaz Khosrawi

References:

Lawrence et al. (2015), Comparisons of polar processing diagnostics from 34 years of the ERA-Interim and MERRA reanalyses, Atmospheric Chemistry and Physics, Vol. 15, Issue 7, 38723-3892.

Manney et al. (2005): Diagnostic comparison of meteorological analyses during the 2002 Antarctic winter, Monthly Weather Review, Vol. 133. Issue 5, 1261-1278.

Manney et al. (2003): Lower stratospheric temperature differences between meteorological analyses in two cold Arctic winters and their impact on polar processing studies, Journal of Geophysical Research, Vol. 108, Issue 5.

Manney et al. (1996): Comparison of U. K. Meteorological Office and U. S. National Center stratospheric analyses during northern and southern winter, Journal of Geophysical Research, Vol. 101, Issue D6, 10311-10334.

References

- Boccara, G., A. Hertzog, C. Basdevant, and F. Vial (2008), Accuracy of NCEP/NCAR reanalyses and ECMWF analyses in the lower stratosphere over Antarctica in 2005, *J. Geophys. Res.*, *113*(D20), doi:10.1029/2008JD010116.
- Hertzog, A., C. Basdevant, F. Vial, and C. Mechoso (2004), The accuracy of stratospheric analyses in the northern hemisphere inferred from long-duration balloon flights, *Quart. J. Roy. Meteorol. Soc.*, *130*(597), 607–626.
- Hoffmann, L., A. W. Grimsdell, and M. J. Alexander (2016), Stratospheric gravity waves at southern hemisphere orographic hotspots: 2003–2014 AIRS/Aqua observations, *Atmos. Chem. Phys. Discuss.*, *2016*, doi:10.5194/acp-2016-341.
- Röbber, T., O. Stein, Y. Heng, and L. Hoffmann (2017), Regional and seasonal truncation errors of trajectory calculations using ECMWF high-resolution operational analyses and forecasts, *Geosci. Model Dev. Discuss.*, *2017*, 1–27, doi:10.5194/gmd-2016-314.

~~Validation~~ Intercomparison of meteorological analyses and trajectories in the Antarctic lower stratosphere ~~using~~ with Concordiasi superpressure balloon observations

Lars Hoffmann¹, Albert Hertzog², Thomas Rößler¹, Olaf Stein¹, and Xue Wu^{1,3}

¹Jülich Supercomputing Centre, Forschungszentrum Jülich, Jülich, Germany

²Laboratoire de Météorologie Dynamique, École Polytechnique, IPSL, Palaiseau, France

³Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

Correspondence to: L. Hoffmann (l.hoffmann@fz-juelich.de)

Abstract. In this study we ~~validated~~ compared temperatures and horizontal winds of meteorological analyses in the Antarctic lower stratosphere, a region of the atmosphere that is of major interest regarding chemistry and dynamics of the polar vortex. The ~~validation was performed with long-duration observations from 19 superpressure balloon flights during the Concordiasi field campaign in September 2010 to January 2011. Our intercomparison study~~ covers the European Centre for Medium-Range
5 Weather Forecasts (ECMWF) operational analysis, the ERA-Interim reanalysis, the Modern-Era Retrospective analysis for Research and Applications (~~MERRA version 1 and 2 (MERRA and MERRA-2)~~), and the National Centers for Environmental Prediction and National Center for Atmospheric Research (NCEP/NCAR) reanalysis. The comparison was performed with respect to long-duration observations from 19 superpressure balloon flights during the Concordiasi field campaign in September 2010 to January 2011. Most of the balloon measurements were conducted at altitudes of 17 – 18.5 km and latitudes of 60 – 85°S.
10 We found that large-scale state temperatures of the analyses have a mean precision of 0.40, 5 – 1.4 K and a warm bias of 0.4 – 2.1 K ~~at about 17 – 18.5 km altitude and 60 – 85°S~~ with respect to the balloon data. Zonal and meridional winds have a mean precision of 0.9 – 2.3 m s⁻¹ and a bias below ±0.5 m s⁻¹ ~~in the same region~~. Standard deviations related to small-scale fluctuations ~~such as due to~~ gravity waves are reproduced at levels of 15 – 60% for temperature and 30 – 60% for the horizontal winds. Considering the fact that the balloon observations have been assimilated into all analyses, except for NCEP/NCAR, notable
15 differences found here indicate that other observations, different forecast models, and different data assimilation procedures have significant impact on the analyses as well. We also used the balloon observations to ~~validate trajectory calculations~~ evaluate trajectory calculations with our new Lagrangian transport model Massive-Parallel Trajectory Calculations (MPTRAC), where vertical motions of simulated trajectories were nudged to pressure measurements of the balloons ~~to take into account changes in the overall mass configuration of the balloon-gondola system~~. We found relative horizontal transport deviations of
20 4.54 – 12% and error growth rates of 60 – 170 km day⁻¹ for 15-day trajectories. Dispersion simulations revealed some difficulties with the representation of subgrid-scale wind fluctuations in ~~our Lagrangian transport model~~ MPTRAC, as the spread of air parcels simulated with different analyses was not consistent. ~~Although~~ However, although case studies suggest that the accuracy of trajectory calculations is influenced by meteorological complexity, diffusion generally does not contribute significantly to transport deviations in our analysis. Overall, ~~validation~~ evaluation results are satisfactory and compare well to earlier

studies using superpressure balloon observations. ~~In most cases, best performance was achieved by the ECMWF operational analysis, having the best spatiotemporal resolution, followed by ERA-Interim, MERRA, and finally NCEP/NCAR, having the lowest spatiotemporal resolution. Future work applying Eulerian or Lagrangian models to study the chemistry and dynamics of the polar vortex may use our validation results as additional guideline for error analyses.~~

5 1 Introduction

The seasonal formation and decay of the [southern hemisphere](#) polar vortex is likely the most prominent feature of the extratropical stratospheric circulation (e. g., Schoeberl and Hartmann, 1991; Newman and Schoeberl, 2003; Waugh and Polvani, 2010). The structure and dynamics of the polar vortex play a key role in the winter and spring stratospheric circulation and coupling between the stratosphere and troposphere. A number of studies demonstrated that the polar vortex can influence tropospheric weather and climate (Baldwin and Dunkerton, 1999; Polvani and Kushner, 2002; Thompson et al., 2002; Baldwin et al., 2003). Furthermore, the polar vortex acts as a cold trap for stratospheric air, which plays a critical role in polar ozone depletion and the annual formation of the Antarctic ozone hole (Solomon, 1999, and references therein). These topics motivated various observational and modeling studies in recent years to better understand the structure and dynamics of the polar vortex as well as implications on polar ozone loss in the stratosphere. [Among those, a number of studies focused on the evaluation of the representation of the southern hemisphere polar vortex in meteorological reanalyses \(Manney et al., 1996, 2005; Lawrence et al., 2015\).](#)

Lagrangian particle dispersion models are indispensable tools to study atmospheric transport processes (e. g., Lin et al., 2012). Trajectory calculations in Lagrangian transport simulations are commonly driven by wind fields from global meteorological reanalyses. The accuracy of trajectory calculations depends on various factors, including interpolation and sampling errors related to the finite spatial resolution of the meteorological data as well as errors of the wind field itself, which are introduced during the data assimilation process (e. g., Stohl, 1998; Bowman et al., 2013). In this study we ~~aimed at direct validation~~ [conducted an intercomparison](#) of temperature and wind data as well as trajectory calculations for the Antarctic lower stratosphere using different meteorological data sets. We ~~assessed the performance of three meteorological~~ [considered four](#) reanalyses, including the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim reanalysis (Dee et al., 2011), the Modern-Era Retrospective analysis for Research and Applications (~~MERRA~~) [reanalysis \(Rienecker et al., 2011\) version 1 and 2 \(MERRA and MERRA-2\) reanalysis \(Rienecker et al., 2011; Bosilovich et al., 2015\)](#), and the National Centers for Environmental Prediction and the National Center for Atmospheric Research (NCEP/NCAR) reanalysis (Kalnay et al., 1996). Furthermore, we ~~compare~~ [compared](#) with the ECMWF operational analysis (OA), which is produced with significantly higher spatial resolution.

~~For validation we utilized~~ [The analyses data are compared with](#) superpressure balloon observations during the Concordiasi field campaign (Rabier et al., 2010) in September 2010 to January 2011. During the campaign 19 superpressure balloons were launched from McMurdo station (78°S, 166°E), Antarctica. Each balloon flew in the mid- and high-latitude lower stratosphere for a typical period of three months. The sensors aboard the balloons provide position, pressure, and temperature at high accuracy and high temporal sampling. Various studies demonstrated that superpressure balloon observations constitute an

excellent source of data for the ~~validation-evaluation~~ of meteorological analyses (Knudsen et al., 1996, 2002; Hertzog et al., 2004, 2006; Knudsen et al., 2006; Boccara et al., 2008; Podglajen et al., 2014; Friedrich et al., 2017). ~~This paper presents an update on earlier work, in particular to Boccara et al. (2008), who performed a validation analysis based on superpressure balloon observations during the Vorecore campaign in Antarctica in September 2005 to February 2006. The results are also~~
5 ~~compared with findings of the PreConcordiasi campaign (Podglajen et al., 2014), which took place at tropical latitudes in February 2010.~~

~~Our new study may contribute directly to current research activities that focus on intercomparisons of different reanalyses, e. g., the Stratosphere-troposphere Processes And their Role in Climate (SPARC) Reanalysis Intercomparison Project (S-RIP) (Fujiwara et al., 2016).~~ Here we applied the Lagrangian particle dispersion model Massive-Parallel Trajectory Calculations
10 (MPTRAC) (Hoffmann et al., 2016a) to conduct the trajectory calculations for the balloon observations. MPTRAC is a rather new model and our study ~~also-mainly~~ serves the purpose of ~~validating-the-evaluating-this~~ model. However, the ~~methods and~~ results are transferable ~~also~~ to other Lagrangian models for the stratosphere ~~as well~~, e. g., the Chemical Lagrangian Model of the Stratosphere (CLaMS) (McKenna et al., 2002a, b) or the Alfred Wegener Institute Lagrangian Chemistry/Transport System (ATLAS) (Wohltmann and Rex, 2009). The results of the ~~direct-validation-of-the-temperature-and-wind-data-of~~
15 ~~the-meteorological-analyses-are-of-interest-for-studies-using-chemistry-transport-models-to-assess-polar-ozone-loss-in-the-stratosphere (e. g., Chipperfield, 1999; Grooß et al., 2002, 2005; Wohltmann et al., 2013).~~ ~~The results of the trajectory-validation trajectory evaluation~~ are of particular interest for studies applying the ‘Match’ technique (von der Gathen et al., 1995; Rex et al., 1997) to assess polar ozone loss. In order to distinguish between chemically and transport-induced changes of ozone abundance, the Match approach uses trajectory calculations to relate ozone observations within the same air mass at different locations
20 to each other. ~~The results of the intercomparison of the temperature and wind data of the meteorological analyses may be of interest for studies using chemistry-transport models to assess polar ozone loss in the stratosphere (e. g., Chipperfield, 1999; Grooß et al., 2002, 2005; Wohltmann et al., 2013).~~ ~~Our new study also contributes to current research activities that focus on intercomparisons of different reanalyses, including the Stratosphere-troposphere Processes And their Role in Climate (SPARC) Reanalysis Intercomparison Project (S-RIP) (Fujiwara et al., 2017).~~

25 In Sect. 2 we introduce the superpressure balloon observations during the Concordiasi campaign. We also describe the ~~four~~ ~~five~~ meteorological data sets and discuss the meteorological conditions during the campaign. Furthermore, we introduce the Lagrangian particle dispersion model MPTRAC and the approach used for trajectory ~~validation-evaluation~~. The results of our study are provided in Sect. 3. In the first part we compare temperatures and horizontal winds of the different meteorological data sets directly at the position of the balloon measurements. In the second part we focus on the ~~validation-evaluation~~ of
30 trajectory calculations, ~~where we assess different types of vertical motions, the impact of the different meteorological data sets, and the impact of subgrid-scale wind fluctuations~~. Finally, Sect. 4 provides a summary and conclusions.

2 Data and methods

2.1 Superpressure balloon observations

Superpressure balloons are aerostatic balloons, which are filled with a fixed amount of lifting gas, and [where for which](#) the maximum volume of the balloon is kept constant by means of a closed, inextensible, spherical envelope. After launch, the balloons ascend and expand until they reach a float level where the atmospheric density matches the balloon density. On this isopycnic surface a balloon is free to float horizontally with the motion of the wind. Hence, superpressure balloons behave as quasi-Lagrangian tracers in the atmosphere. In this study we analyzed superpressure balloon observations in the lower stratosphere during the Concordiasi field campaign in Antarctica in September 2010 to January 2011. The Concordiasi field campaign aimed at making innovative atmospheric observations to study the circulation and chemical species in the polar lower stratosphere and to reduce uncertainties in diverse fields in Antarctic science (Rabier et al., 2010). During the field campaign 19 superpressure balloons with 12 m diameter were launched from McMurdo station (78°S, 166°E), Antarctica by the French space agency, Centre National d'Etudes Spatiales (CNES). Balloons of this size typically drift at pressure levels of ~ 60 hPa and altitudes of ~ 18 km. The balloons were launched between 8 September and 26 October 2010, and each balloon flew in the mid- and high-latitude lower stratosphere for a typical period of two to three months. The flight dates are summarized in Table 1 and the balloon trajectories are shown in Fig. 1.

The positions of the balloons were tracked [over time every 60 s](#) by means of global positioning satellite (GPS) receivers. At each observation time the components of the horizontal wind are computed by finite differences between the GPS positions. The uncertainty is about 1 m for the GPS horizontal position and 0.1 m s^{-1} for the derived winds (Podglajen et al., 2014). Each balloon launched during Concordiasi was equipped with a meteorological payload called Thermodynamical sensor (TSEN). TSEN makes in-situ measurements of atmospheric pressure and temperature every 30 s during the whole flight. The pressure is measured with an accuracy of 1 Pa and a precision of 0.1 Pa. The air temperature is measured via two thermistors. During daytime, the thermistors are heated by the sun, leading to daytime temperature measurements [being](#) warmer than the real air temperature. An empirical correction has been used to correct for this effect ([Hertzog et al., 2004](#)), [which is described in detail by Hertzog et al. \(2004\)](#). The precision of [the corrected](#) temperature observations is ~ 0.25 K during daytime and ~ 0.1 K during nighttime. Note that technical issues aboard the scientific gondola caused a few data gaps in the TSEN data set, but most of them were shorter than 15 min.

In order to quantify the coverage of the balloon observations during the free-flying phases, we independently calculated the 5% and 95% quantiles of various parameter distributions. All statistics presented in this paper are most representative for the parameter ranges reported below. Any findings for parameters outside these ranges need to be considered carefully, because only few measurements are available to support them. We found that most of the measurements (i. e., more than 90%) took place between 25 September and 22 December 2010, at an altitude range of 17.0–18.5 km, and within a latitude range of 59–84°S. The pressure measurements are mostly within a range of 58.2–69.1 hPa and the temperature measurements within 189–227 K. The density of air, calculated from pressure and temperature, varies between 0.099 – 0.120 kg m^{-3} . The zonal

winds are predominately westerly and mostly within a range of $1-44 \text{ m s}^{-1}$. The meridional wind distributions are nearly symmetric, with meridional winds being in the range of $\pm 17 \text{ m s}^{-1}$. Horizontal wind speeds are mostly within $5-47 \text{ m s}^{-1}$.

As an example, Fig. 2 shows time series of density, temperature, zonal wind, and meridional wind as measured during flight number 4 of the Concordiasi campaign. The density time series shows decreasing density during the first 20 days, but remains rather stable thereafter. This initial decrease in density is due to the release of dropsondes, which are another part of the balloon payloads on flight number 1 – 13. The release of dropsondes changes the overall mass configuration of the balloon-gondola system, which is compensated by changes in density. A closer inspection of the time series reveals also diurnal variations in the balloon density. During the day the balloon ~~envelop~~envelope is heated by the sun, which increases the temperature and pressure of the gas inside the balloon. The balloon slightly expands in return, which decreases its equilibrium density. In addition to this regular daily pattern, the time series show notable variability on even shorter time scales, including semi-diurnal oscillations of the horizontal winds, which are attributed to near-inertial gravity waves and semi-diurnal tides. As we do not expect the reanalyses to reproduce those fluctuations with great accuracy, we applied a ~~low-pass band-pass~~ filter with 15 h cut-off period to ~~suppress all oscillations caused by separate between small-scale features (e. g., pure and inertia-gravity waves, so that only perturbations due to)~~ and the large-scale ~~dynamics remain visible in the time series (cf. Fig. 2)~~state (e. g., zonal temperature gradients and planetary waves). The cut-off period of the ~~low-pass band-pass~~ filter was selected to cover the longest inertial ~~period periods~~ in the balloon data set, $T = 2\pi f^{-1}$, with Coriolis parameter f , ranging from about 12.0 h at 85°S to 13.9 h at 60°S . Figure 2 illustrates the effect of low-pass filtering to extract the large-scale state.

2.2 Meteorological data

In this study we considered ~~four~~five meteorological data sets, the ECMWF operational analysis, ERA-Interim (Dee et al., 2011), MERRA (Rienecker et al., 2011), MERRA-2 (Bosilovich et al., 2015), and the NCEP/NCAR reanalysis (Kalnay et al., 1996). ~~Fujiwara et al. (2016)~~Fujiwara et al. (2017) provides a review of key aspects of ~~these data sets~~the reanalyses. Table 2 summarizes information on spatial and temporal resolution and coverage. ~~The four data sets considered here of the data sets as considered in this study. Note that the five data sets~~ vary substantially in resolution, i. e., by a factor of 2 in temporal resolution, by a factor of 5 in vertical resolution, and by a factor of 20×20 in horizontal resolution. ~~Note that we~~We retrieved the data sets at the temporal and spatial resolution at which they are typically provided to the users by the respective centers. ~~The same data sets have been considered by Hoffmann et al. (2016a), who provide a more detailed description of data preprocessing~~Following Hoffmann et al. (2016a), both ECMWF data sets were retrieved on hybrid sigma-pressure levels and converted to pressure levels by means of the Climate Data Operators (Schulzweida, 2014) whereas MERRA and NCEP/NCAR data were retrieved directly on pressure levels. For MERRA-2 we implemented new code in our Lagrangian transport model in order to be able to process meteorological data directly on hybrid sigma-pressure levels, which finally allowed us to consider MERRA-2 data with higher spatial resolution in this study.

An important aspect that needs to be taken into account in a comparison of the Concordiasi balloon observations and the meteorological data sets is that the balloon observations have been subject to data assimilation. In particular, 15-min time averaged data from the Concordiasi balloons have been transmitted over the Global Telecommunication System (GTS)

(Rabier et al., 2013). The data transmitted over GTS were then assimilated by the respective centers. The Concordiasi balloon observations have been assimilated into the ECMWF data sets, MERRA, and MERRA-2, but they were not considered for the NCEP/NCAR reanalysis. The observations therefore provide an independent data source for validation only for the NCEP/NCAR data set. However, as meteorological analyses are a result of combining various satellite and in-situ observations, a forecast model, and a data assimilation procedure, a comparison of the meteorological data with the Concordiasi observations still provides information on the performance of the overall system. As the observational data have been subject to downsampling and data thinning before they were assimilated, an assessment of the representation of small-scale structures due to gravity waves also remains meaningful.

The Concordiasi balloon measurements covered the final stratospheric warming and decay of the southern hemisphere polar vortex during 2010/2011 austral spring to summer. Although a mid-winter minor sudden stratospheric warming during July and early August 2010 resulted in an off-pole displacement and weakening of the stratospheric polar vortex (De Laat and van Weele, 2011; Klekociuk et al., 2011), the polar vortex returned to be relatively stable from mid-August to October, except for a second short warming that began in early September. This pattern was primarily attributed to the quasi-biennial oscillation being in a strong westerly or positive phase that helped to maintain a persistent polar vortex. According to NASA Ozone Watch and the World Meteorological Organisation Antarctic Ozone Bulletins (see <http://www.wmo.int/pages/prog/arep/gaw/ozone/index.html>; last access: 30 September 2016), the longitudinally averaged poleward eddy heat flux between 45°S and 75°S, which is an indicator of disturbance in polar stratosphere, was much smaller than the long-term mean (Fig. 3), indicating that the vortex was relatively unperturbed from mid-September to December.

Figure 4 illustrates that the polar vortex was typically quite symmetric and stable in September and October. Afterwards, the polar vortex elongated and weakened gradually through November, was displaced off the pole in mid-December and broke down by mid-January 2011. The vortex breakup was marked when the winds around the vortex edge decreased below 15 m s^{-1} on the 475 K potential temperature surface. From an analysis of temperatures on the levels where most of the balloon measurements were attained (about 50–60 hPa, $\sim 475 \text{ K}$), the final warming started from mid-October with development of strong zonal asymmetries in temperature. The cold pool over the South Pole declined and displaced, and until end of November, minimum temperatures over Antarctica increased from around 180 to 220 K. A warm pool with temperatures of 230–240 K dominated Antarctica from end of December. Consistent with the warming process, the polar jet showed a pronounced reduction in wind speed from 70 m s^{-1} at the beginning of September to 40 m s^{-1} in early by mid of December and then further weakened to less than 20 m s^{-1} from beginning of January.

2.3 Trajectory calculations

We conducted the trajectory calculations for the Concordiasi balloon observations with the Lagrangian particle dispersion model MPTRAC (Hoffmann et al., 2016a). MPTRAC has been developed to support analyses of atmospheric transport processes in the free troposphere and stratosphere. In previous studies it was used to perform transport simulations for volcanic eruptions and to reconstruct time- and height-resolved emission rates for these events (Heng et al., 2016; Hoffmann et al., 2016a). Transport is simulated by calculating trajectories for large numbers of air parcels based on given wind fields

from global meteorological reanalyses. [The numerical accuracy and efficiency of trajectory calculations with MPTRAC was assessed by Rößler et al. \(2017\)](#). Turbulent diffusion and subgrid-scale wind fluctuations are simulated based on the Langevin equation, closely following the approach implemented in the Flexible Particle (FLEXPART) model (Stohl et al., 2005). Additional modules allow us to simulate sedimentation and the decay of particle mass, but they were not used here. The model is particularly suited for large-scale simulations on supercomputers due to its efficient Message Passing Interface (MPI) / Open Multi-Processing (OpenMP) hybrid parallelization ([Heng et al., 2016; Rößler et al., 2017](#)).

Trajectory calculations are based on numerical integration of the kinematic equation of motion,

$$\frac{d\mathbf{x}}{dt} = \mathbf{v}(\mathbf{x}, t), \quad (1)$$

where \mathbf{x} denotes the position and \mathbf{v} the velocity of an air parcel at time t . The air parcel position \mathbf{x} is defined by geographic latitude ϕ and longitude λ as horizontal coordinates as well as pressure p as vertical coordinate. The horizontal wind (u, v) and vertical velocity $(\omega = dp/dt)$ at position \mathbf{x} and time t are obtained by [linear-spatial-and-temporal-4-D-linear](#) interpolation of the meteorological data [in space and time](#). The kinematic equation of motion is solved with the explicit midpoint method,

$$\mathbf{x}(t + \Delta t) = \mathbf{x}(t) + \mathbf{v} \left(\mathbf{x}(t) + \frac{\Delta t}{2} \mathbf{v}(\mathbf{x}(t), t), t + \frac{\Delta t}{2} \right) \Delta t. \quad (2)$$

The time step Δt mainly controls the trade-off between accuracy and speed of the calculations. For our simulations we selected $\Delta t = 30$ s, which is sufficiently small so that truncation errors can be neglected ([Rößler et al., 2017](#)). This time step is also consistent with the sampling rate of the balloon data.

The diffusion module of MPTRAC considers two processes. Turbulent diffusion is modelled by means of uncorrelated, Gaussian random displacements of the air parcels with zero mean and standard deviations $\sqrt{D_x \Delta t}$ and $\sqrt{D_z \Delta t}$, where D_x and D_z are the horizontal and vertical diffusion coefficients, respectively. Typical values for the stratosphere are $D_x = 0$ and $D_z = 0.1 \text{ m}^2 \text{ s}^{-1}$, according to choices made for the FLEXPART model (Legras et al., 2003; Stohl et al., 2005). Unresolved subgrid-scale wind fluctuations are most relevant for long-range simulations. These fluctuations are correlated over time and simulated with a Markov model, following the approach of Maryon (1998) and Stohl et al. (2005). For example, the zonal wind fluctuations u' of each air parcel are calculated according to

$$u'(t + \Delta t) = r u'(t) + \sqrt{(1 - r^2) \alpha \sigma_u^2} \xi, \quad (3)$$

with $r = \exp(-2\Delta t / \Delta t_{\text{met}})$ being a correlation coefficient depending on the model time step Δt and the time interval Δt_{met} of the meteorological data (3 or 6 h), α being a scaling factor used for downscaling of [space and time](#) grid-scale variances σ_u^2 to subgrid scales, and ξ being a Gaussian random variate with zero mean and unity variance. The FLEXPART model uses a default value of $\alpha = 0.16$ for downscaling of the grid-scale variances (or 40% in terms of standard deviations). Meridional wind and vertical velocity fluctuations are calculated in the same way.

For this study we implemented a new module in MPTRAC that allows us to simulate the vertical motions of the balloons more realistically. This module is called at each time step and adjusts the pressure of the air parcels so that vertical motions are constrained to either (i) an isobaric surface (constant pressure), (ii) an isopycnic surface (constant density), (iii) an isentropic

surface (constant potential temperature), or (iv) the pressure time series measured by the balloon. In a first approximation the balloons move on isopycnic surfaces, which is represented by option (ii). However, the real dynamics of the balloons are more complex, in particular if they encounter small-scale structures such as gravity waves (Vincent and Hertzog, 2014). On longer time scales it needs to be considered that there are diurnal variations in the balloon density as well as overall mass variations due to the release of dropsondes (Sect. 2.1). These issues are partly circumvented by constraining the vertical motions to the balloon pressure data, which is represented by option (iv).

2.4 ~~Validation~~ Evaluation approach

Although some of the Concordiasi balloon flights can be used to ~~validate~~ evaluate trajectory calculations for time periods as long as three months, we focused on shorter time windows. By splitting the balloon flights into smaller subsets of data, each containing 15 days of observations, we significantly increased the number of samples and improved the statistical accuracy of the results ~~of the short-term validation~~. To further increase the number of samples we also allowed for overlap of the time windows, i. e., we shifted the 15-day windows in steps of 5 days. A shift of 5 days between the windows was selected, because trajectory errors are usually larger than the effective resolution of the meteorological data sets after that time. This means we can consider the results of overlapping windows as being statistically independent. We varied the starting days for the analysis of the different flights to homogenize temporal coverage. As there are data gaps in the GPS and TSEN data of the balloon measurements, we imposed the requirement that each sample should have at least 90% coverage. Based on these criteria we obtained a set of 104 samples of 15-day time windows from the 19 Concordiasi balloon flights.

Absolute horizontal transport deviations (AHTDs) and relative horizontal transport deviations (RHTDs) are standard measures to compare trajectory calculations with observations or to evaluate results for different model configurations (Kuo et al., 1985; Rolph and Draxler, 1990; Stohl et al., 1995; Stohl, 1998). While other measures of trajectory error have also been defined, AHTDs and RHTDs are most often reported because they can be compared easily to other studies. The AHTD at travel time t of the trajectories is calculated as

$$\text{AHTD}(t) = \frac{1}{N_s N_e} \sum_{i=1}^{N_s} \sum_{j=1}^{N_e} \sqrt{[X_{i,j}(t) - x_i(t)]^2 + [Y_{i,j}(t) - y_i(t)]^2}, \quad (4)$$

where N_s refers to the number of reference trajectories and N_e refers to the size of the ensemble of test trajectories that is to be evaluated for each reference trajectory. The coordinates $(X_{i,j}, Y_{i,j})$ and (x_i, y_i) with $i = 1, \dots, N_s$ and $j = 1, \dots, N_e$ refer to the horizontal positions of the test and reference trajectories, respectively. Equation (4) is applied in different ways in this study. For instance, it is used to evaluate transport deviations between a ~~single~~-model trajectory and a balloon trajectory for just one sample ($N_s = 1$ and $N_e = 1$), between model and balloon trajectories for all samples ($N_s = 104$ and $N_e = 1$), or for dispersion simulations ($N_s = 104$ and $N_e = 1000$). Note that we calculated horizontal distances as Euclidean distances of the air parcel positions projected to the Earth's surface. RHTDs are calculated by dividing the AHTD of individual air parcels by the length of the corresponding reference trajectory. Absolute and relative vertical transport deviations (AVTDs and RVTDs) are defined similarly, based on pressure differences converted into vertical distances by means of the barometric formula.

3 Results

3.1 Direct ~~validation-intercomparison~~ of meteorological data

In this section we focus on ~~the validation-a intercomparison~~ of temperatures and horizontal winds directly at the positions of the Concordiasi balloons. For this analysis the meteorological data are interpolated to the balloon positions by means of a
5 4-D linear interpolation in space and time. This interpolation scheme is most commonly applied in state-of-the-art Lagrangian transport models (Bowman et al., 2013). Table 3 presents ~~the~~ summary statistics of ~~low-pass filtered~~ meteorological data minus ~~low-pass filtered~~ Concordiasi balloon observations. ~~Note that we first discuss the results for low-pass filtered data, to exclude the effects of small-scale fluctuations due to gravity waves and turbulence (, which indicates differences in the large-scale state (see Sect. 2.1). These effects will be discussed separately at the end of this section. Our analysis shows that~~
10 ~~all data sets~~ Table 3 shows that the analyses have a positive temperature bias, ~~which is~~ in the range of 0.4 K (ECMWF OA) to 2.1 K (NCEP/NCAR). Zonal wind biases are in the range of -0.3 m s^{-1} (NCEP/NCAR) ~~-0.3~~ to 0.5 m s^{-1} (MERRA). Meridional wind biases are ~~about below~~ 0.1 m s^{-1} for all data sets. Standard deviations vary between 0.5 K (ECMWF OA) and 1.4 K (NCEP/NCAR) for temperature, 0.9 m s^{-1} (ECMWF OA) and 2.3 m s^{-1} (NCEP/NCAR) for the zonal wind, and 0.9 m s^{-1} (ECMWF OA) and 1.9 m s^{-1} (NCEP/NCAR) for the meridional wind. ~~Skewness and excess kurtosis values found here indicate that the distributions are quite symmetric and not affected by many or rather large outliers. We confirmed these findings by calculating additionally more robust and resistant statistical measures of location, spread, and symmetry, namely the median, the interquartile range, and the Yule-Kendall index (Wilks, 2011). These robust measures provided a quite similar picture to the standard measures~~
15 Note that the largest biases and standard deviations were typically found for the NCEP/NCAR data set, which may be attributed to the fact that this data set is independent, whereas the Concordiasi balloon observations
20 have been assimilated into the other analyses. However, the statistics show that there are still significant differences between the data sets with balloon data being assimilated (ECMWF products, MERRA, and MERRA-2), which shows that the analyses are affected also by other observations (e. g. satellite data) and different forecasts models and assimilation procedures.

Figure 5 shows ~~bias large-scale state biases~~ and standard deviations of ~~temperature-temperatures~~ and horizontal winds at different latitudes averaged over the entire time period of the campaign. ~~All meteorological data sets~~ Variations between
25 different months are typically smaller (not shown). All analyses show an increasing temperature bias from mid to high latitudes. The temperature warm bias at $80-85^\circ\text{S}$ is largest for NCEP/NCAR (3.1 K), followed by MERRA (1.4 K), ~~ERA-Interim (1.1 K)~~ MERRA-2 (1.3 K), and ECMWF OA (0.5 K). ~~At $60-65^\circ\text{S}$ the temperature biases range from 0.2 to 1.2 K.~~ ERA-Interim (1.1 K). ~~Temperature standard deviations do not show variation with latitude (ECMWF OA and ERA-Interim) or just a slight decrease from mid to high latitudes (MERRA and NCEP/NCAR). Zonal wind biases are below $\pm 0.7 \text{ m s}^{-1}$ and meridional wind biases are below $\pm 0.3 \text{ m s}^{-1}$ for all latitude bands considered here. Standard deviations of the zonal and meridional wind do not vary significantly with latitude for ECMWF OA and ERA-Interim. MERRA shows latitudinal variation in the range of $1.3-1.9 \text{ m s}^{-1}$ for the zonal wind and $1.2-1.6 \text{ m s}^{-1}$ for the meridional wind. NCEP/NCAR mostly shows increasing standard deviations from high to mid latitudes, ranging from 1.8 to 3.0 m s^{-1} for the zonal wind and from 1.5 to 2.4 m s^{-1} for the meridional wind.~~

30 ~~the meridional wind.~~

Figure ?? shows bias and standard deviations of temperature and horizontal winds for different months averaged over all latitudes. The temperature bias is at a maximum in September (1.3 K for ERA-Interim and 0.8 K for ECMWF OA) or October (2.7 K for NCEP/NCAR and 1.3 K for MERRA). Temperature standard deviations remain rather constant in the range from 0.5 to 0.9 K (ECMWF OA, ERA-Interim, and MERRA) or increase from 0.9 K in September to 1.6 K in December (NCEP/NCAR). Zonal wind biases vary more for MERRA and NCEP/NCAR, with a range of $\pm 0.9 \text{ m s}^{-1}$, and less for ECMWF OA and ERA-Interim, with a range of $\pm 0.3 \text{ m s}^{-1}$. Meridional wind biases remain below $\pm 0.2 \text{ m s}^{-1}$ for all months. The standard deviations of the zonal and meridional winds remain rather constant at about $0.8\text{--}1.0 \text{ m s}^{-1}$ for ECMWF OA and ERA-Interim, but tend to increase from September to December for MERRA and NCEP/NCAR. In December we found maximum standard deviations of 2.7 m s^{-1} (NCEP/NCAR) and 1.7 m s^{-1} (MERRA) for the zonal wind and 2.1 m s^{-1} (NCEP/NCAR) and 1.6 m s^{-1} (MERRA) for the meridional wind.

Overall, the direct validation of the large-scale features of the meteorological analyses in the Antarctic lower stratosphere provides satisfactory results. The summary statistics presented here are similar to those of earlier campaigns using superpressure balloons observations, in particular with respect to results presented by Boccara et al. (2008) for the Voreore campaign in 2005. Temperature biases of meteorological analyses at the southern hemisphere winter pole are well-known phenomena, which were reported also for other winters (Gobiet et al., 2005; Parrondo et al., 2007; Boccara et al., 2008). Using GPS radio occultation measurements in June to August 2003, Gobiet et al. (2005) showed that temperature biases of ECMWF analyses over the southern hemisphere winter pole vary with altitude. They found a warm bias of up to 3.5 K at 18–19 km (close to the altitude of the Concordiasi balloon observations), a cold bias of up to -3 K at 21–22 km, and a warm bias of up to 3.5 K at 26–27 km earlier winters in other studies (Gobiet et al., 2005; Parrondo et al., 2007; Boccara et al., 2008). Gobiet et al. (2005) speculate that the assimilation of microwave radiances from satellite measurements into the ECMWF analyses may be a reason for the temperature bias. We note that a warm bias is still present in Antarctic winter 2010 in all data sets, but its magnitude is significantly reduced for ECMWF OA. The magnitude of the temperature warm bias found here for NCEP/NCAR is comparable with those found in earlier studies. The temperature bias for the other analyses is smaller, which may be attributed to improvements of the forecast model, data assimilation scheme, and observations used to produce this analysis. Although many factors influence the accuracy and precision of meteorological analyses, Table 3 indicates that the spatiotemporal resolution is a rather important factor. Both bias and standard deviations are lowest for ECMWF OA, which has the highest resolution, followed by ERA-Interim, MERRA, and finally the fact the Concordiasi data have been assimilated. The same reason likely explains why wind biases as well as temperature and wind standard deviations shown in Fig. 5 are generally largest for NCEP/NCAR, which has the lowest resolution whereas they are smaller and more similar to each other for both ECMWF data sets and MERRA-2.

Finally, we also analyzed the effects of the low-pass filter that was applied to remove small-scale fluctuations from the data. Table 4 provides standard deviations of unfiltered minus filtered temperatures and horizontal winds. For comparison we also provided standard deviations of unfiltered minus filtered balloon data in Table 4, which are considered as a measure of real small-scale fluctuations in the atmosphere. In fact, high-pass filtered horizontal winds for the analyses and the balloon data. Note that the balloon observations are an excellent source of data to study real small-scale fluctuations in the atmosphere.

which are mostly attributed to gravity waves (e. g., Hertzog et al., 2008, 2012; Plougonven et al., 2013; Vincent and Hertzog, 2014; Jewtoukoff et al., 2015). ~~While large-scale biases are not affected by filtering and therefore not reported here, a~~ A comparison of standard deviations allows us to assess how well small-scale fluctuations are represented in the meteorological data sets. ~~For the high-resolution ECMWF OA data the standard deviations removed by filtering are largest and about the~~ same size as the standard deviations related to the differences of meteorological data minus balloon data analyses. We found that ECMWF OA reproduces about 60% and ERA-Interim ~~and MERRA, MERRA, and MERRA-2~~ MERRA, and MERRA-2 about 30% of the standard deviations of the temperature and wind fluctuations of the balloons. NCEP/NCAR reproduces about 15% for temperature and 30% for the winds. ~~This is in good~~ These differences are associated with the spatial resolution of the analyses (see Table 2), because the forecast models are able to simulate gravity waves patterns more realistically, if they are operating at higher spatial resolution. Our results are in excellent agreement with the studies of Jewtoukoff et al. (2015), which found that ECMWF analyses underestimate gravity wave momentum fluxes derived from the Concordiasi balloon observations by a factor of 5, and Hoffmann et al. (2016b), which found that wave amplitudes in the ECMWF analyses are typically underestimated by a factor of 2–3 compared to Atmospheric InfraRed Sounder (AIRS/Aqua) observations.

3.2 Analysis of vertical motions

In the ~~following sections of this paper remaining sections~~ validation-evaluation of trajectory calculations using the MPTRAC model with Concordiasi superpressure balloon observations. As outlined in Sect. 2.3, we implemented several new options in the MPTRAC model to constrain the vertical motions of ~~the~~ air parcels. We first tried to identify the approach that is best suited to simulate the vertical motions of the superpressure balloons in a realistic manner. Note that previous trajectory studies on tropospheric altitude-controlled balloons used pressure measurements to constrain vertical motions (Baumann and Stohl, 1997; Riddle et al., 2006). Trajectory evaluations with stratospheric superpressure balloons were conducted with the isopycnic approach (Hertzog et al., 2004; Boccarra et al., 2008). In our comparison we considered vertical motions based on prescribed pressure time series as measured by the balloons, isopycnic motions, isentropic motions, and vertical motions prescribed by the vertical velocities of the meteorological data sets (referred to as ‘omega velocities’ below). ~~The comparison was conducted using ERA-Interim data as input for the trajectory calculations.~~

For illustration, Fig. 6 shows examples of trajectories calculated with different types of vertical motions and the corresponding balloon observations. This comparison was conducted using ERA-Interim data as input for the trajectory calculations. Within 15 days the balloon is advected by the polar night jet over a distance of nearly 30 000 km and encircles the south pole more than twice. At the end of the simulations we found horizontal transport deviations of about 30 km (0.1%) using the balloon pressure, 100 km (0.3%) for the isopycnic approach, 350 km (1.2%) for the isentropic approach, and 400 km (1.3%) for the omega velocity. In this particular example the balloon trajectory is reproduced with excellent accuracy by all simulations. We picked this particular example for presentation because the simulations are not strongly affected by any individual, complex meteorological conditions. In ~~the~~ this example the balloon trajectory is best reproduced by constraining vertical movements based on the balloon pressure measurements or by using the isopycnic approach, as expected from the balloon dynamics (Sect. 2.1). Larger transport deviations are found using omega velocities and the isentropic approach. However, note that the trajec-

tories based on omega velocities and the isentropic approach are in good agreement with each other, which was expected as atmospheric motions are isentropic on short time scales.

In order to take into account statistical variations, Fig. 7 shows transport deviations calculated from 104 samples of 15-day balloon trajectories of the Concordiasi campaign, which we selected according to the approach outlined in Sect. 2.4. The AHTDs increase rather steadily to about 1610–1750 km after 15 days. Like ~~As~~ in the example shown in Fig. 6, the results cluster in two groups. Trajectories calculated using the balloon pressure and the isopycnic approach are similar to each other and yield results at the lower end of the AHTD ranges. Trajectories calculated using omega velocities and the isentropic approach are also similar to each other and yield results at the upper end of the AHTD ranges. The corresponding RHTDs are in a range of ~~4.14–5.25~~ % after 2 days and increase to ~~6.8–7.4~~ about 7 % after 15 days. The mean difference between the two groups of simulations is about 0.7 percentage points. Note that RHTDs are quite large during the first 12–24 h, which is not representative, because the calculations are based on rather short reference trajectories. In addition, Fig. 7 also shows vertical transport deviations based on the isopycnic and isentropic approach as well as omega velocities. The AVTDs of the isopycnic approach increase steadily to about 200 m after 15 days. The corresponding RVTDs converge at 6–7% after 4 days. The AVTDs using omega velocities and the isentropic approach increase rapidly during the first 2 days and then increase more slowly up to 560–680 m after 15 days. The corresponding RVTDs converge to 17–21%. A possible reason for larger initial ~~errors~~ deviations using omega velocities and the isentropic approach are uncertainties in the initial pressure values used to define the trajectory seeds. Simulations based on omega velocities or the isentropic approach are more strongly affected by short-term fluctuations of the initial pressure values than simulations based on the isopycnic approach. To mitigate uncertainties caused by short-term fluctuations, we used the mean pressure of the first 3 h of each balloon trajectory for initialization. However, our analysis still indicates that vertical motions are best calculated using either the balloon pressure measurements or the isopycnic approach. For the remaining analyses we decided to calculate the trajectories using the balloon pressure measurements because ~~these take into account any~~ this takes into account changes in the overall mass configuration of the balloon-gondola system (Sect. 2.1).

3.3 Impact of different meteorological analyses on trajectory calculations

In this section we present a comparison of transport deviations obtained with different meteorological data sets. ~~Figures 8 and ??~~ show ~~Figure 8 shows~~ two examples of 15-day trajectory calculations using ECMWF OA, ERA-Interim, MERRA, MERRA-2, and NCEP/NCAR data. The examples mainly serve to illustrate the large range of variability found in different simulations. For flight number 2 the simulated trajectories reproduce the observed balloon trajectory quite well. We found maximum AHTDs in ~~the a~~ range of 650–1050 km and maximum RHTDs in ~~the a~~ range of 3–7% for the different data sets. Note that the maxima occur on different days, i. e., simulated trajectories may first deviate from and then approach the observed trajectories again. Despite being shorter (i. e., 12 700 km versus 29 700 km), the simulated trajectories for flight number 12 deviate much ~~larger~~ stronger from the observations. Here we found maximum AHTDs of 3100–5200 km and maximum RHTDs of 53–70%. The two examples illustrate the large variability between different samples, which is attributed to situation-dependent factors

, including such as the individual meteorological conditions. A large number of independent samples needs to be analyzed in order to obtain statistically significant results.

Figure 9 shows transport deviations for the different meteorological data sets calculated from 104 samples of 15-day trajectories (Sect. 2.4). In contrast to the individual examples, we found that the AHTDs increase rather steadily over time, which suggests that outliers play a minor role and that the statistics are robust. After 15 days the AHTDs are close in a range of 1400 to 2200 km for NCEP/NCAR, 1800 km for MERRA, 1600 km for ERA-Interim, and 1400 km for ECMWF-OA. From Fig. 9 we can also estimate the growth rates of the AHTDs. The growth rate for NCEP/NCAR is close, which are typically within 60 to 170 km day⁻¹ for the first 12 days, but slightly decreases thereafter. For MERRA we found a growth rate of 120 km day⁻¹. The growth rates of both ERA-Interim and ECMWF-OA are close to 60 km day⁻¹ during the first 5 days, close to 110 km day⁻¹ during day 6. The RHTDs are in a range of 4–12, and get more variable during day 13 12% after 2 days, but converge to a smaller range of 6–15. The RHTDs of 9% after 15 days. Although the transport deviations grow rather steadily, the relative differences between the data sets tend to get smaller over time. The largest transport deviations and growth rates were found for NCEP/NCAR decrease from about 12% after day 1 to about 8.5% after day 15. The RHTDs of MERRA decrease from 10% to 7.5%. The RHTDs of both ECMWF-OA and ERA-Interim slightly increase from 4.5% to 6–7%. These results, which may be attributed to the fact that the wind data of this analysis are most uncertain because the Concordiasi balloon observation were not assimilated (Sect. 3.1). However, our results still agree well with those reported by Boccara et al. (2008) for the Vorcore campaign in 2005. 2005, despite the fact that data assimilation of the balloon observations did not play a role in that study. For 15 days' trajectory time Boccara et al. (2008) found mean spherical distances of about 1650 km (with an interquartile range of 800–3600 km) for ECMWF analyses (sampled at 0.5° × 0.5° horizontal resolution and 60 levels vertically) and 2350 km (1400–3800 km) for NCEP/NCAR data. The transport deviations and growth rates found here also compare well with a wider range of results for the troposphere reported by Stohl (1998). Our analysis indicates that the best accuracy of trajectory calculations in the Antarctic lower stratosphere is achieved with ECMWF-OA and ERA-Interim, followed by MERRA and NCEP/NCAR. This is related to the accuracy of the horizontal winds of the meteorological data sets as discussed in Sect. 3.1.

3.4 Impact of subgrid-scale wind fluctuations

In this section we discuss the influence of diffusion on the trajectory calculations. We assessed this by means of dispersion simulations, each consisting of 1000 trajectories for each sample, and by applying the MPTRAC diffusion module described in Sect. 2.3. Note that these simulations consider only horizontal diffusion, because vertical motions have been restricted to the pressure measurements of the balloons. Following Stohl et al. (2005), the turbulent horizontal diffusivity coefficient in the stratosphere was set to zero, $D_x = 0$, i. e., the diffusion in our simulations is related only to horizontal subgrid-scale wind fluctuations. For comparison with diffusion-free simulations, two examples of dispersion simulations are also shown in Fig. 8. For flight number 2 we found only minor spread of the air parcels due to diffusion whereas for flight number 12 it is quite substantial, illustrating that diffusion may vary significantly from case to case. The examples also suggest that the uncertainties of the trajectory calculations are linked to the meteorological situation, as low diffusion goes along with good accuracy of the trajectories for flight number 2 whereas high diffusion goes along with low accuracy for flight number 12.

Kahl (1996) analyzed correlations between trajectory model errors and the complexity of the meteorological situation under study in more detail. He quantified the complexity of the meteorological conditions by means of the so-called ‘meteorological complexity factor’ (MCF), which measures the dispersion of a set of stochastic trajectories generated by random perturbations superimposed upon an observed wind field. Kahl (1996) pointed out that trajectory errors are representative only if they are larger than the corresponding MCF. Similar to Kahl (1996), we estimated the MCF of our simulations by applying Eq. (4) to the trajectory ensemble. However, instead of taking the balloon trajectory as a reference, the MCF was calculated using a simulated trajectory without diffusion as a reference. The simulated reference trajectory is usually close to the ensemble mean because the deviations of the ensemble trajectories are often symmetric around the ensemble mean. The MCFs of the ~~four~~ five meteorological data sets of our study are shown in Fig. 9. The MCFs increase rather steadily over time. After 15 days we found values of about 1300 km for ECMWF OA, 800–900 km for MERRA and NCEP/NCAR, ~~and 600 km~~ for ERA-Interim, and 300 km for MERRA-2. These differences in the MCFs came somewhat unexpected, as the spread of air parcels ideally should be the same in all simulations, independent of the meteorological data set and the diffusion model being applied. The differences are not directly related to the resolution of the meteorological data sets, as can be seen from the ranking of the MCFs of the data sets. The inconsistencies of the MCFs found here might be due to dynamical inconsistencies of the analysis wind fields that are introduced during the data assimilation process. Such dynamical inconsistencies may lead to more rapid dispersion and spurious mixing in Lagrangian transport model simulations (Stohl et al., 2004).

In principle, we may tune the scaling factor α in Eq. (3) of the MPTRAC diffusion module to achieve simulations with more consistent MCFs. However, we refrained from any tuning measures, because appropriate reference data ~~for validation~~ are lacking. We applied a constant scaling factor $\alpha = 0.16$ in all simulations, which is the default value used in the FLEXPART model. However, despite the different levels of MCFs found in the simulations, we conclude that the transport deviations between the simulations and the balloons ~~are~~ can be considered representative, because ~~the are generally~~ they are notably larger than the MCFs. To further confirm this result we also calculated the AHTDs between the trajectory ensembles and the balloon trajectories. We found that the transport deviations with or without diffusion are rather similar (Fig. 9). The AHTDs for ERA-Interim, MERRA, MERRA-2, and NCEP/NCAR differ less than ± 50 km. For ECMWF OA the AHTDs with diffusion are up to 200 km larger than the AHTDs without diffusion. We attribute this to the fact that simulated diffusion is largest for ECMWF OA, as indicated by the corresponding MCFs. This shows that diffusion does not induce any significant uncertainties in our analysis of transport deviations. The results remain meaningful, even if diffusion is not explicitly taken into account.

4 Summary and conclusions

In this study we ~~validated temperature and horizontal wind data of~~ conducted an intercomparison of temperatures and horizontal winds from the ECMWF operational analysis (~~OA~~) and the ERA-Interim, MERRA, MERRA-2, and NCEP/NCAR reanalyses at southern hemisphere mid and high latitudes (~~about in the lower stratosphere. The analyses were compared with Concordiasi superpressure balloon observations in September 2010 to January 2011. Most of the balloon observations took place at 60–85°S~~) in the lower stratosphere (about latitude and 17–18.5 km). ~~The validation was based on Concordiasi superpressure~~

balloon observations in September 2010 to January 2011. We found temperature warm biases altitude. In this comparison we had to consider that 15 min-downsampled Concordiasi data have been assimilated into both ECMWF data sets, MERRA, and MERRA-2, but that they were not considered for the NCEP/NCAR reanalysis. For the direct intercomparison of the temperature and wind data at the balloon positions, a band-pass filter with 15 h cutoff period was applied to separate between the large-scale state and small-scale features.

The most prominent finding regarding the large-scale state was a temperature warm bias of the analyses in the range of 0.4 K at high latitudes. This bias was largest for NCEP/NCAR (up to 3.1 K at 80–85°S), but it was also present in the other analyses (up to 0.5–1.4 K, which are similar to the values found in earlier studies (Gobiet et al., 2005; Parrondo et al., 2007; Boccara et al., 2008) at 80–85°S), despite the fact that the balloon observations have been assimilated. Stratospheric temperature biases for the southern hemisphere polar vortex have already been found in other studies for earlier winters (Gobiet et al., 2005; Parrondo et al., 2007; Boouroum et al., 2014). Our study indicates that they were still present in 2010/2011. Zonal and meridional wind biases of the low-pass filtered data are below $\pm 0.5 \text{ m s}^{-1}$. After applying a low-pass filter to remove small-scale fluctuations due to gravity waves and turbulence, standard deviations of analyses minus observations Standard deviations are in the range of 0.4–1.4 K for temperature and 0.9–2.3 m s^{-1} for the winds. Overall, these are satisfactory validation results that are comparable to other studies using superpressure balloon observations in the Antarctic lower stratosphere (e.g., Boccara et al., 2008). Note that ECMWF OA, ERA-Interim, and MERRA validation results for Antarctica are much better than those found by Podglajen et al. (2014) for the equatorial lower stratosphere. As Podglajen et al. (2014) and this study both used observations gathered in 2010, this provides further evidence that the quality of meteorological analyses tends to degrade from high latitudes towards the Equator. Podglajen et al. (2014) showed that the lower quality of the reanalyses at low latitudes is associated with poor representation of large-scale equatorial waves, which might be improved by more direct observations of stratospheric wind profiles over wide regions along the equatorial belt. Considering four different horizontal wind components. We found significant differences between the meteorological analyses, even with the balloon data being assimilated, which suggests that the analyses are also significantly affected by other observations and the different forecast models and assimilation procedures. Observing system experiments would be required to assess the specific impact of the balloon observations on the analyses.

The five meteorological data sets considered in our study, we found clear indications that differ significantly in spatial and temporal resolution of the data and. The truncation of the models also play plays an important role in determining accuracy and precision of the analyses. Best large-scale accuracy and precision are achieved by ECMWF OA (highest resolution), followed by ERA-Interim, MERRA, and NCEP/NCAR (lowest resolution). Model truncation also affects the representation of how well the analyses are capable of representing small-scale fluctuations. A number of studies already demonstrated that superpressure balloon observations are particularly suited to study gravity waves (e.g., Hertzog et al., 2008, 2012; Plougonven et al., 2013; Vincent and Hegerl, 2004). Standard deviations of unfiltered minus high-pass filtered temperature and wind data of the balloons are reproduced at a level of about 60% by ECMWF OA the ECMWF operational analysis, but only by at a level of 15–30% by the reanalyses. For ECMWF OA operational analysis temperatures this is consistent with recent studies of Jewtoukoff et al. (2015) and Hoffmann et al. (2016b), providing further evidence that the ECMWF operational model explicitly resolves a significant portion of the atmospheric gravity wave spectrum.

We also used the Concordiasi balloon observations to ~~validate trajectory calculations in the Antarctic lower stratosphere~~ evaluate trajectory calculations with our rather new Lagrangian particle dispersion model MPTRAC. Some difficulties are related to the fact that the overall mass configuration of the balloon-gondola system ~~may change during the flight. Our analysis showed changed during some of the flights. The analysis of vertical motions confirmed~~ that balloon trajectories are best reproduced by the isopycnic approach (Hertzog et al., 2004; Boccara et al., 2008) or by nudging vertical motions to the pressure measurements of the balloons (Baumann and Stohl, 1997; Riddle et al., 2006). In this study we analyzed 104 samples of trajectories from 19 balloon flights for time periods of 15 days. ~~Relative-Absolute~~ horizontal transport deviations ~~are in the range of 4.5–7% for ECMWF OA and ERA-Interim, 7.5~~ typically grow at rates of ~~60–10% for MERRA, and 8.5–170–km 12% for NCEP/NCAR. Growth rates of absolute day⁻¹ for all data sets. Relative~~ horizontal transport deviations ~~are in the range of 60~~ showed larger differences at the beginning of the simulations, but converged to a range of ~~6–110 km day⁻¹ for ECMWF OA and ERA-Interim, about 120 km day⁻¹ for MERRA, and about 170 km day⁻¹ 9%~~ after 15 days. The largest transport deviations and growth rates were found for NCEP/NCAR. These results agree, which may be attributed to the fact that the Concordiasi balloon observation were not assimilated into this analysis. In general, our results compare well with those of reported by Boccara et al. (2008) for the Vorcore campaign in 2005. They show a significant improvement compared to early studies of Knudsen and Carver (1994) and Knudsen et al. (1996), which found transport deviations of about 20% between trajectories based on ECMWF analyses and long-duration balloon observations. We conducted the trajectory calculations with the Lagrangian particle dispersion model MPTRAC (Hoffmann et al., 2016a); and our study provides a contribution to the validation of this new model. However, the results will be transferable also to other Lagrangian transport models for the stratosphere. ~~We used the diffusion module of MPTRAC to conduct dispersion simulations 2005, despite the fact that data~~ assimilation of the balloon observations did not play a role in that study.

In order to assess the impact of diffusion we conducted dispersion simulations with MPTRAC. The analysis revealed some difficulties with the modelling approach for subgrid-scale wind fluctuations and the wind data driving ~~the these~~ simulations, as the spread of air parcel trajectories simulated with different meteorological data sets was not consistent. We also did not find correlations between the spread and the spatial and temporal resolution of the data sets. Future work may comprise additional analyses and ~~may~~ focus on tuning of the subgrid-scale parametrization scheme. Selected examples of dispersion simulations indicate that the accuracy of trajectory calculations is linked to meteorological complexity, as suggested by Kahl (1996). In this study we analyzed a rather large number of trajectory samples, though, and the effects of meteorological complexity averaged out and did not alter the results of the analysis of transport deviations significantly. ~~Future studies applying chemistry-transport~~ The evaluation suggests that the MPTRAC model is capable of calculating trajectories in the Antarctic lower stratosphere with an accuracy similar to that obtained in other studies. The methods and results should be transferable to other Lagrangian transport models for the stratosphere and may help to improve future studies using these models to assess the dynamics of the polar vortex or to investigate polar ozone loss ~~may use our validation results as additional guideline for error analysis.~~

5 Code and data availability

The quality-controlled meteorological TSEN data set is available from Laboratoire de Météorologie Dynamique (LMD) from their web site at <http://www.lmd.polytechnique.fr/VORCORE/McMurdoE.htm> (last access: 21 December 2016). The ERA-Interim reanalysis and operational analyses are distributed by the European Centre for Medium-Range Weather Forecasts (ECMWF), see <http://www.ecmwf.int/en/forecasts/datasets> (last access: 21 December 2016). MERRA data are provided by the Global Modeling and Assimilation Office at NASA Goddard Space Flight Center through the NASA GES DISC online archive, see <http://disc.sci.gsfc.nasa.gov/mdisc/data-holdings/merra> (last access: 21 December 2016). NCEP/NCAR reanalysis data were obtained from the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their web site at <http://www.esrl.noaa.gov/psd> (last access: 21 December 2016). The code of the Massive-Parallel Trajectory Calculations (MPTRAC) model is available under the terms and conditions of the GNU General Public License, Version 3 from the repository at <https://github.com/slcs-jsc/mptrac> (last access: 21 December 2016).

Author contributions. All authors contributed to the design of the study and provided input to the manuscript. LH conducted the transport simulations and the scientific analysis. AH provided support regarding the scientific analysis of the Concordiasi superpressure balloon observations. TR and OS were responsible for preprocessing of the meteorological data. XW provided the characterization of the meteorological conditions during the campaign.

Acknowledgements. Concordiasi was built by an international scientific group and is currently supported by the following agencies: Météo-France, CNES, IPEV, PNRA, CNRS/INSU, NSF, NCAR, Concordia consortium, University of Wyoming, and Purdue University. ECMWF also contributes to the project through computer resources and support, and scientific expertise. The two operational polar agencies PNRA and IPEV are thanked for their support at Concordia station. Concordiasi is part of the THORPEX-IPY cluster within the International Polar Year effort. The authors acknowledge the Jülich Supercomputing Centre (JSC) for providing computing time on the supercomputer JURECA.

References

- Baldwin, M. P. and Dunkerton, T. J.: Propagation of the Arctic Oscillation from the stratosphere to the troposphere, *J. Geophys. Res.*, 104, 30937–30946, 1999.
- Baldwin, M. P., Thompson, D. W., Shuckburgh, E. F., Norton, W. A., and Gillett, N. P.: Weather from the stratosphere?, *Science*, 301, 317–319, 2003.
- Baumann, K. and Stohl, A.: Validation of a Long-Range Trajectory Model Using Gas Balloon Tracks from the Gordon Bennett Cup 95, *J. Appl. Met.*, 36, 711–720, doi:10.1175/1520-0450-36.6.711, 1997.
- Boccarra, G., Hertzog, A., Basdevant, C., and Vial, F.: Accuracy of NCEP/NCAR reanalyses and ECMWF analyses in the lower stratosphere over Antarctica in 2005, *J. Geophys. Res.*, 113, doi:10.1029/2008JD010116, 2008.
- 10 Bosilovich, M., Akella, S., Coy, L., Cullather, R., Draper, C., Gelaro, R., Kovach, R., Liu, Q., Molod, A., Norris, P., Wargan, K., Chao, W., Reichle, R., Takacs, L., Vikhliayev, Y., Bloom, S., Collow, A., Firth, S., Labow, G., Partyka, G., Pawson, S., Reale, O., Schubert, S. D., and Suarez, M.: MERRA-2: Initial evaluation of the climate, Tech. rep., NASA, series on Global Modeling and Data Assimilation, NASA/TM–2015-104606, Vol. 43, 2015.
- Bowman, K. P., Lin, J. C., Stohl, A., Draxler, R., Konopka, P., Andrews, A., and Brunner, D.: Input Data Requirements for Lagrangian Trajectory Models, *B. Am. Meteorol. Soc.*, 94, 1051–1058, 2013.
- 15 Chipperfield, M. P.: Multiannual Simulations with a Three-Dimensional Chemical Transport Model, *J. Geophys. Res.*, 104, 1781–1805, 1999.
- De Laat, A. and van Weele, M.: The 2010 Antarctic ozone hole: Observed reduction in ozone destruction by minor sudden stratospheric warmings, *Sci. Rep.*, 1, doi:10.1038/srep00038, 2011.
- 20 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., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, I., 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, *Quart. J. Roy. Meteorol. Soc.*, 137, 553–597, 2011.
- 25 Friedrich, L. S., McDonald, A. J., Bodeker, G. E., Cooper, K. E., Lewis, J., and Paterson, A. J.: A comparison of Loon balloon observations and stratospheric reanalysis products, *Atmos. Chem. Phys.*, 17, 855–866, 2017.
- Fujiwara, M., Wright, J. S., Manney, G. L., Gray, L. J., Anstey, J., Birner, T., Davis, S., Gerber, E. P., Harvey, V. L., Hegglin, M. I., Homeyer, C. R., Knox, J. A., Krüger, K., Lambert, A., Long, C. S., Martineau, P., Monge-Sanz, B. M., Santee, M. L., Tegtmeier, S., Chabrillat, S., Tan, D. G. H., Jackson, D. R., Polavarapu, S., Compo, G. P., Dragani, R., Ebisuzaki, W., Harada, Y., Kobayashi, C., McCarty, W., Onogi, K., Pawson, S., Simmons, A., Wargan, K., Whitaker, J. S., and Zou, C.-Z.: Introduction to the SPARC Reanalysis Intercomparison Project (S-RIP) and overview of the reanalysis systems, *Atmos. Chem. Phys. Discuss.*, 2016, 1–52, doi:10.5194/acp-2016-652, 2016.
- 30 Fujiwara, M., Wright, J. S., Manney, G. L., Gray, L. J., Anstey, J., Birner, T., Davis, S., Gerber, E. P., Harvey, V. L., Hegglin, M. I., Homeyer, C. R., Knox, J. A., Krüger, K., Lambert, A., Long, C. S., Martineau, P., Molod, A., Monge-Sanz, B. M., Santee, M. L., Tegtmeier, S., Chabrillat, S., Tan, D. G. H., Jackson, D. R., Polavarapu, S., Compo, G. P., Dragani, R., Ebisuzaki, W., Harada, Y., Kobayashi, C., McCarty, W., Onogi, K., Pawson, S., Simmons, A., Wargan, K., Whitaker, J. S., and Zou, C.-Z.: Introduction to the SPARC Reanalysis Intercomparison Project (S-RIP) and overview of the reanalysis systems, *Atmos. Chem. Phys.*, 17, 1417–1452, doi:10.5194/acp-17-1417-2017, 2017.
- 35

- Gobiet, A., Foelsche, U., Steiner, A. K., Borsche, M., Kirchengast, G., and Wickert, J.: Climatological validation of stratospheric temperatures in ECMWF operational analyses with CHAMP radio occultation data, *Geophys. Res. Lett.*, 32, doi:10.1029/2005GL022617, 2005.
- Groß, J.-U., Günther, G., Konopka, P., Müller, R., McKenna, D. S., Stroh, F., Vogel, B., Engel, A., Müller, M., Hoppel, K., Bevilacqua, R., Richard, E., Webster, C. R., Elkins, J. W., Hurst, D. F., Romashkin, P. A., and Baumgardner, D. G.: Simulation of ozone depletion in spring 2000 with the Chemical Lagrangian Model of the Stratosphere (CLaMS), *J. Geophys. Res.*, 107, 8295, doi:10.1029/2001JD000456, 2002.
- Groß, J.-U., Konopka, P., and Müller, R.: Ozone Chemistry during the 2002 Antarctic Vortex Split, *J. Atmos. Sci.*, 62, 860–870, 2005.
- Heng, Y., Hoffmann, L., Griessbach, S., Röbler, T., and Stein, O.: Inverse transport modeling of volcanic sulfur dioxide emissions using large-scale simulations, *Geosci. Model Dev.*, 9, 1627–1645, 2016.
- Hertzog, A., Basdevant, C., Vial, F., and Mechoso, C.: The accuracy of stratospheric analyses in the northern hemisphere inferred from long-duration balloon flights, *Quart. J. Roy. Meteorol. Soc.*, 130, 607–626, 2004.
- Hertzog, A., Basdevant, C., and Vial, F.: An Assessment of ECMWF and NCEP–NCAR Reanalyses in the Southern Hemisphere at the End of the Presatellite Era: Results from the EOLE Experiment (1971–72), *Mon. Wea. Rev.*, 134, 3367–3383, 2006.
- Hertzog, A., Boccara, G., Vincent, R. A., Vial, F., and Cocquerez, P.: Estimation of gravity wave momentum flux and phase speeds from quasi-Lagrangian stratospheric balloon flights. Part II: Results from the Vorcore campaign in Antarctica, *J. Atmos. Sci.*, 65, 3056–3070, 2008.
- Hertzog, A., Alexander, M. J., and Plougonven, R.: On the intermittency of gravity wave momentum flux in the stratosphere, *J. Atmos. Sci.*, 69, 3433–3448, 2012.
- Hoffmann, L., Röbler, T., Griessbach, S., Heng, Y., and Stein, O.: Lagrangian transport simulations of volcanic sulfur dioxide emissions: impact of meteorological data products, *J. Geophys. Res.*, pp. 4651–4673, doi:10.1002/2015JD023749, 2016a.
- Hoffmann, L., Spang, R., Orr, A., Alexander, M. J., Holt, L. A., and Stein, O.: A decadal satellite record of gravity wave activity in the lower stratosphere to study polar stratospheric cloud formation, *Atmos. Chem. Phys. Discuss.*, 16, 1–33, doi:10.5194/acp-2016-757, 2016b.
- Jewtoukoff, V., Hertzog, A., Plougonven, R., Cámara, A. d. I., and Lott, F.: Comparison of Gravity Waves in the Southern Hemisphere Derived from Balloon Observations and the ECMWF Analyses, *J. Atmos. Sci.*, 72, 3449–3468, 2015.
- Kahl, J. D.: On the prediction of trajectory model error, *Atmos. Environment*, 30, 2945–2957, 1996.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-year reanalysis project, *B. Am. Meteorol. Soc.*, 77, 437–471, 1996.
- Klekociuk, A., Tully, M., Alexander, S., Dargaville, R., Deschamps, L., Fraser, P., Gies, H., Henderson, S., Javorniczky, J., Krummel, P., Petelina, S., Shanklin, J., Siddaway, J., and Stone, K.: The Antarctic ozone hole during 2010, *Aust. Meteor. Ocean. J.*, 61, 253–267, 2011.
- Knudsen, B. and Carver, G.: Accuracy of the isentropic trajectories calculated for the EASOE campaign, *Geophys. Res. Lett.*, 21, 1199–1202, 1994.
- Knudsen, B., Rosen, J., Kjome, N., and Whitten, A.: Comparison of analyzed stratospheric temperatures and calculated trajectories with long-duration balloon data, *J. Geophys. Res.*, 101, 19 137–19 145, 1996.
- Knudsen, B. M., Pommereau, J.-P., Garnier, A., Nunes-Pinharanda, M., Denis, L., Newman, P., Letrenne, G., and Durand, M.: Accuracy of analyzed stratospheric temperatures in the winter Arctic vortex from infrared Montgolfier long-duration balloon flights – 2. Results, *J. Geophys. Res.*, 107, SOL 4–1–SOL 4–13, doi:10.1029/2001JD001329, 2002.

- Knudsen, B. M., Christensen, T., Hertzog, A., Deme, A., Vial, F., and Pommereau, J.-P.: Accuracy of analyzed temperatures, winds and trajectories in the Southern Hemisphere tropical and midlatitude stratosphere as compared to long-duration balloon flights, *Atmos. Chem. Phys.*, 6, 5391–5397, 2006.
- Kuo, Y.-H., Skumanich, M., Haagensohn, P. L., and Chang, J. S.: The accuracy of trajectory models as revealed by the observing system simulation experiments, *Mon. Wea. Rev.*, 113, 1852–1867, 1985.
- Lawrence, Z. D., Manney, G. L., Minschwaner, K., Santee, M. L., and Lambert, A.: Comparisons of polar processing diagnostics from 34 years of the ERA-Interim and MERRA reanalyses, *Atmos. Chem. Phys.*, 15, 3873–3892, doi:10.5194/acp-15-3873-2015, 2015.
- Legras, B., Joseph, B., and Lefèvre, F.: Vertical diffusivity in the lower stratosphere from Lagrangian back-trajectory reconstructions of ozone profiles, *J. Geophys. Res.*, 108, 4562, doi:10.1029/2002JD003045, 2003.
- 10 Lin, J., Brunner, D., Gerbig, C., Stohl, A., Luhar, A., and Webley, P., eds.: Lagrangian modeling of the atmosphere, vol. 200 of *Geophysical Monograph Series*, American Geophysical Union, Washington DC, 2012.
- Manney, G. L., Swinbank, R., Massie, S. T., Gelman, M. E., Miller, A. J., Nagatani, R., O’Neill, A., and Zurek, R. W.: Comparison of U.K. Meteorological Office and U.S. National Meteorological Center stratospheric analyses during northern and southern winter, *J. Geophys. Res.*, 101, 10 311–10 334, doi:10.1029/95JD03350, 1996.
- 15 Manney, G. L., Allen, D. R., Krüger, K., Naujokat, B., Santee, M. L., Sabutis, J. L., Pawson, S., Swinbank, R., Randall, C. E., Simmons, A. J., et al.: Diagnostic comparison of meteorological analyses during the 2002 Antarctic winter, *Mon. Wea. Rev.*, 133, 1261–1278, 2005.
- Maryon, R.: Determining cross-wind variance for low frequency wind meander, *Atmos. Environment*, 32, 115–121, 1998.
- McKenna, D. S., Groöß, J.-U., Günther, G., Konopka, P., Müller, R., Carver, G., and Sasano, Y.: A new Chemical Lagrangian Model of the Stratosphere (CLaMS) 2. Formulation of chemistry scheme and initialization, *J. Geophys. Res.*, 107, ACH 4–1–ACH 4–14, doi:10.1029/2000JD000113, 2002a.
- 20 McKenna, D. S., Konopka, P., Groöß, J.-U., Günther, G., Müller, R., Spang, R., Offermann, D., and Orsolini, Y.: A new Chemical Lagrangian Model of the Stratosphere (CLaMS) 1. Formulation of advection and mixing, *J. Geophys. Res.*, 107, ACH 15–1–ACH 15–15, doi:10.1029/2000JD000114, 2002b.
- Newman, P. A. and Schoeberl, M. R.: Middle atmosphere: Polar vortex, in: *Encyclopedia of Atmospheric Sciences*, edited by Holton, J. R., Pyle, J., and Curry, J. A., pp. 1321–1328, Academic, San Diego, Calif., 2003.
- 25 Parrondo, M. C., Yela, M., Gil, M., von der Gathen, P., and Ochoa, H.: Mid-winter lower stratosphere temperatures in the Antarctic vortex: comparison between observations and ECMWF and NCEP operational models, *Atmos. Chem. Phys.*, 7, 435–441, 2007.
- Plougonven, R., Hertzog, A., and Guez, L.: Gravity waves over Antarctica and the Southern Ocean: Consistent momentum fluxes in mesoscale simulations and stratospheric balloon observations, *Quart. J. Roy. Meteorol. Soc.*, 139, 101–118, 2013.
- 30 Podglajen, A., Hertzog, A., Plougonven, R., and Žagar, N.: Assessment of the accuracy of (re)analyses in the equatorial lower stratosphere, *J. Geophys. Res.*, 119, 11,166–11,188, 2014.
- Polvani, L. M. and Kushner, P. J.: Tropospheric response to stratospheric perturbations in a relatively simple general circulation model, *Geophys. Res. Lett.*, 29, doi:10.1029/2001GL014284, 2002.
- Rabier, F., Bouchard, A., Brun, E., Doerenbecher, A., Guedj, S., Guidard, V., Karbou, F., Peuch, V.-H., El Amraoui, L., Puech, D., et al.: The CONCORDIASI project in Antarctica, *B. Am. Meteorol. Soc.*, 91, 69–86, 2010.
- 35 Rabier, F., Cohn, S., Cocquerez, P., Hertzog, A., Avallone, L., Deshler, T., Haase, J., Hock, T., Doerenbecher, A., Wang, J., Guidard, V., Thépaut, J.-N., Langland, R., Tangborn, A., Balsamo, G., Brun, E., Parsons, D., Bordereau, J., Cardinali, C., Danis, F., Escarnot, J.-P., Fourrié, N., Gelaro, R., Genthon, C., Ide, K., Kalnajs, L., Martin, C., Meunier, L.-F., Nicot, J.-M., Perttula, T., Potts, N., Ragazzo,

- P., Richardson, D., Sosa-Sesma, S., and Vargas, A.: The Concordiasi Field Experiment over Antarctica: First Results from Innovative Atmospheric Measurements, *B. Am. Meteorol. Soc.*, 94, ES17–ES20, doi:10.1175/BAMS-D-12-00005.1, 2013.
- Rex, M., Harris, N., Von der Gathen, P., Lehmann, R., Braathen, G., Reimer, E., Beck, A., Chipperfield, M., Alfier, R., Allaart, M., O’Connor, F., Dier, H., Dorokhov, V., Fast, H., Gil, M., Kyrö, E., Litynska, Z., Mikkelsen, I., Molyneux, M., Nakane, H., Notholt, J., Rummukainen, M., Viatte, P., and Wenger, J.: Prolonged stratospheric ozone loss in the 1995–96 arctic winter, *Nature*, 389, 835–838, 1997.
- Riddle, E. E., Voss, P. B., Stohl, A., Holcomb, D., Maczka, D., Washburn, K., and Talbot, R. W.: Trajectory model validation using newly developed altitude-controlled balloons during the International Consortium for Atmospheric Research on Transport and Transformations 2004 campaign, *J. Geophys. Res.*, 111, D23S57, doi:10.1029/2006JD007456, 2006.
- Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and Woollen, J.: MERRA: NASA’s Modern-Era Retrospective Analysis for Research and Applications, *J. Clim.*, 24, 3624–3648, 2011.
- Rolph, G. D. and Draxler, R. R.: Sensitivity of three-dimensional trajectories to the spatial and temporal densities of the wind field, *J. Appl. Met.*, 29, 1043–1054, 1990.
- Rößler, T., Stein, O., Heng, Y., and Hoffmann, L.: Regional and seasonal truncation errors of trajectory calculations using ECMWF high-resolution operational analyses and forecasts, *Geosci. Model Dev. Discuss.*, 2017, 1–27, doi:10.5194/gmd-2016-314, 2017.
- Schoeberl, M. R. and Hartmann, D. L.: The Dynamics of the Stratospheric Polar Vortex and Its Relation to Springtime Ozone Depletions, *Science*, 251, 46–52, 1991.
- Schulzweida, U.: CDO User’s Guide, Max-Planck-Institut für Meteorologie, Hamburg, <https://code.zmaw.de/projects/cdo> (last access: 27 May 2015), 2014.
- Solomon, S.: Stratospheric ozone depletion: A review of concepts and history, *Rev. Geophys.*, 37, 275–316, 1999.
- Stohl, A.: Computation, accuracy and applications of trajectories – a review and bibliography, *Atmos. Environment*, 32, 947–966, 1998.
- Stohl, A., Wotawa, G., Seibert, P., and Kromp-Kolb, H.: Interpolation errors in wind fields as a function of spatial and temporal resolution and their impact on different types of kinematic trajectories, *J. Appl. Met.*, 34, 2149–2165, 1995.
- Stohl, A., Cooper, O., and James, P.: A cautionary note on the use of meteorological analysis fields for quantifying atmospheric mixing, *J. Atmos. Sci.*, 61, 1446–1453, 2004.
- Stohl, A., Forster, C., Frank, A., Seibert, P., and Wotawa, G.: Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2, *Atmos. Chem. Phys.*, 5, 2461–2474, 2005.
- Thompson, D. W. J., Baldwin, M. P., and Wallace, J. M.: Stratospheric Connection to Northern Hemisphere Wintertime Weather: Implications for Prediction, *J. Clim.*, 15, 1421–1428, 2002.
- Vincent, R. A. and Hertzog, A.: The response of superpressure balloons to gravity wave motions, *Atmos. Meas. Tech.*, 7, 1043–1055, 2014.
- von der Gathen, P., Rex, M., Harris, N. R. P., Lucic, D., Knudsen, B. M., Braathen, G. O., De Backer, H., Fabian, R., Fast, H., Gil, M., Kyrö, E., Mikkelsen, I. S., Rummukainen, M., Stahelin, J., and Varotsos, C.: Observational evidence for chemical ozone depletion over the Arctic in winter 1991–92, *Nature*, 375, 131–134, 1995.
- Waugh, D. W. and Polvani, L. M.: Stratospheric polar vortices, *The Stratosphere: Dynamics, Transport, and Chemistry*, Geophys. Monogr. Ser., 190, 43–57, 2010.
- Wilks, D. S.: Statistical methods in the atmospheric sciences, vol. 100, Academic Press, 2011.

Wohltmann, I. and Rex, M.: The Lagrangian chemistry and transport model ATLAS: validation of advective transport and mixing, *Geosci. Model Dev.*, 2, 153–173, 2009.

Wohltmann, I., Wegner, T., Müller, R., Lehmann, R., Rex, M., Manney, G. L., Santee, M. L., Bernath, P., Sumińska-Ebersoldt, O., Stroh, F., von Hobe, M., Volk, C. M., Hösen, E., Ravegnani, F., Ulanovsky, A., and Yushkov, V.: Uncertainties in modelling heterogeneous chemistry

5 and Arctic ozone depletion in the winter 2009/2010, *Atmos. Chem. Phys.*, 13, 3909–3929, 2013.

Table 1. Concordiasi Balloon Flights over Antarctica in September 2010 to January 2011

Flight Number	Flight Code	Gondola ID	Flight Start	Flight End
1	MSD01	10V01N46	2010/09/23	2010/12/11
2	MSD02	10V02N48	2010/09/23	2010/11/18
3	MSD03	10V03N39	2010/10/15	2010/11/04
4	MSD04	10V04N40	2010/09/24	2010/12/27
5	MSD05	10V05N44	2010/09/25	2010/12/22
6	MSD06	10V06N37	2010/09/28	2010/12/09
7	MSD07	10V07N41	2010/09/30	2010/12/09
8	MSD08	10V08N49	2010/10/26	2011/01/19
9	MSD09	10V09N22	2010/10/07	2011/01/04
10	MSD10	10V10N25	2010/10/14	2010/12/24
11	MSD11	10V11N56	2010/10/19	2010/12/29
12	MSD12	10V12N66	2010/10/20	2011/01/23
13	MSD13	10V13N65	2010/10/19	2010/11/30
14	PSC14	10V14N42	2010/09/15	2010/12/21
15	PSC15	10V15N32	2010/09/08	2010/09/16
16	PSC16	10V16N35	2010/09/11	2010/10/11
17	PSC17	10V17N31	2010/09/14	2010/12/10
18	PSC18	10V18N43	2010/09/29	2010/12/16
19	PSC19	10V19N27	2010/10/08	2010/12/24

Table 2. Temporal and ~~Spatial~~spatial Resolution of Meteorological Data Sets as considered in this Study

Data Product	Temporal Resolution	Top Level	Vertical Levels	Horizontal Resolution
ECMWF OA	3 h	0.01 hPa	91	$0.125^\circ \times 0.125^\circ$
ERA-Interim	6 h	0.1 hPa	60	$1.000^\circ \times 1.000^\circ$
<u>MERRA-2</u>	<u>3 h</u>	<u>0.01 hPa</u>	<u>72</u>	<u>$0.500^\circ \times 0.667^\circ$</u>
MERRA	3 h	0.1 hPa	42	$1.250^\circ \times 1.250^\circ$
NCEP/NCAR	6 h	10 hPa	17	$2.500^\circ \times 2.500^\circ$

Table 3. Statistics of Low-Pass Filtered Meteorological Analyses Minus-minus Concordiasi Balloon Observations (Based on $N \approx 2.52 \times 10^6$ Measurements)

	ECMWF OA	ERA-Interim	<u>MERRA-2</u>
Temperature [K]			
Bias	0.4	0.8	<u>1.0</u>
Standard Deviation	0.5	0.6	<u>0.7</u>
Skewness 0.5 0.4 -0.2 -0.8 Excess Kurtosis 2.2 1.0 -0.1 1.0 Zonal Wind [m s^{-1}]			
Bias	0.1	0.3	<u>0.3</u>
Standard Deviation	0.9	1.0	1.6 2.3 Skewness 0.1 0.1 -0.4 -0.4 Ex
Meridional Wind [m s^{-1}]			
Bias	0.1	0.1	<u>0.0</u>
Standard Deviation	0.9	0.9	<u>1.1</u>
Skewness 0.0 0.0 -0.3 0.0 Excess Kurtosis 1.8 3.3 2.9 1.5 height			

Table 4. Standard Deviations of ~~Unfiltered-Minus-Filtered~~ High-Pass filtered Meteorological ~~Data~~ Analyses and Concordiasi Balloon Observations

	Balloons	ECMWF OA	ERA-Interim	<u>MERRA-2</u>	MERRA	NCEP/NCAR
Temperature [K]	0.7	0.4	0.2	0.2	<u>0.2</u>	0.1
Zonal Wind [m s^{-1}]	1.5	0.9	0.4	0.4	0.4	<u>0.4</u>
Meridional Wind [m s^{-1}]	1.6	1.0	0.5	0.5	0.5	<u>0.5</u>

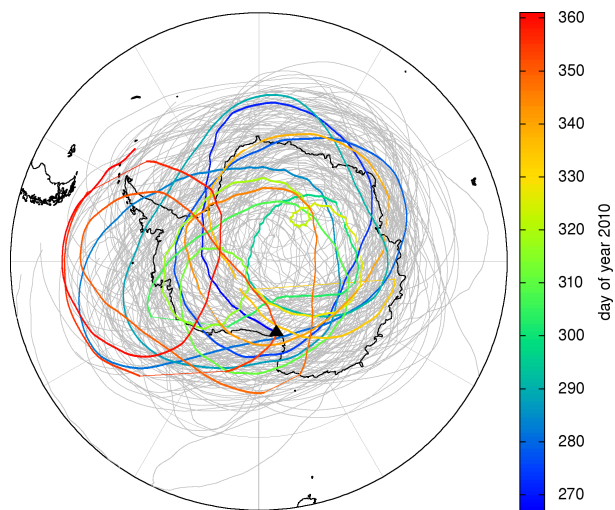


Figure 1. Map of superpressure balloon trajectories (gray curves) during the Concordiasi field campaign in Antarctica in September 2010 to January 2011. The colored curve highlights the trajectory of flight number 4, with colors from blue to red indicating measurement time. The black triangle shows the location of McMurdo station.

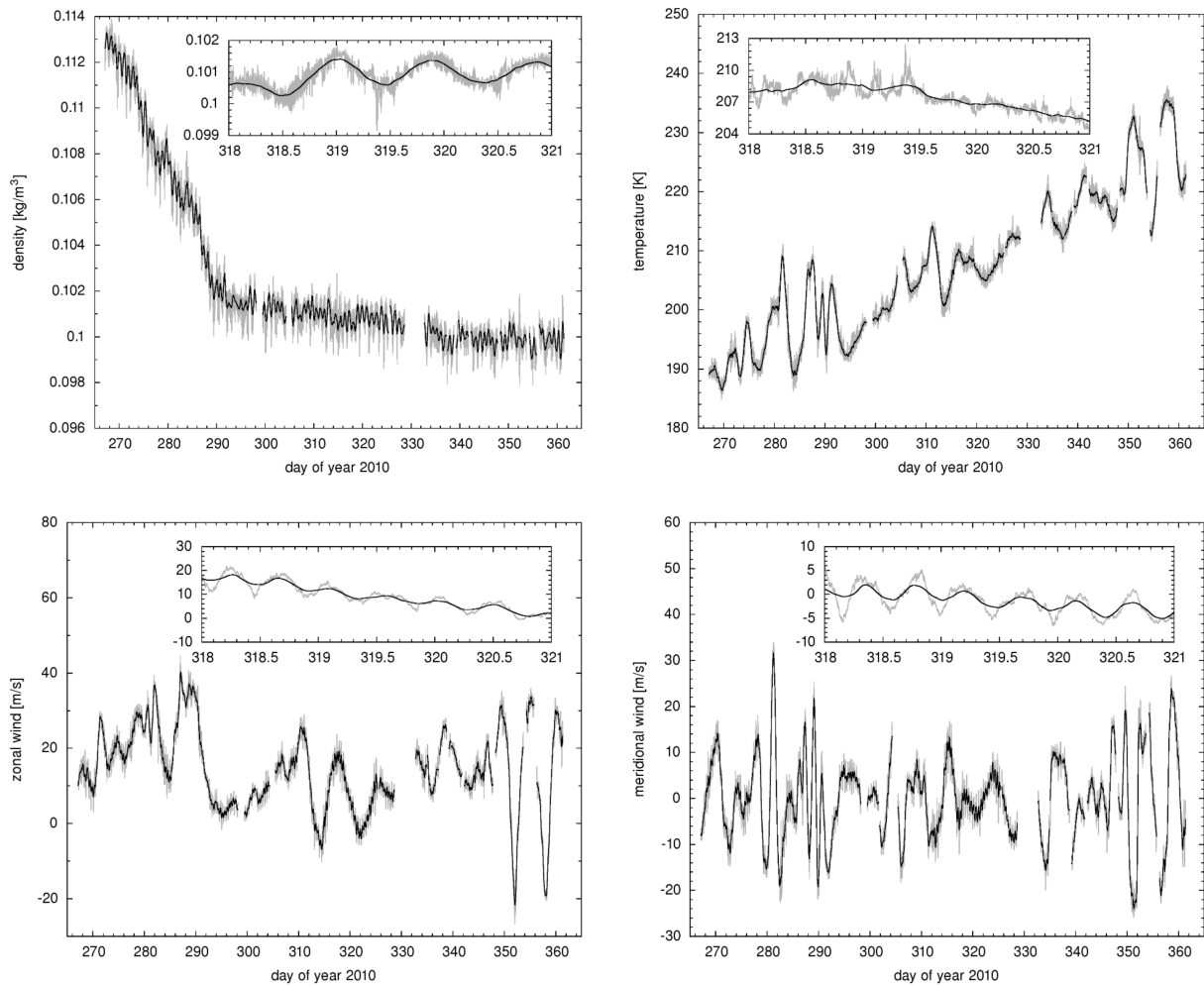


Figure 2. Time series of meteorological data of flight number 4 of the Concordiasi campaign (see Fig. 1). Grey curves show unfiltered data from GPS and TSEN measurements at 30 s time intervals. Black curves show results of a low-pass filter with 15 h cut-off frequency. Inset plots show data for 14–17 November 2010.

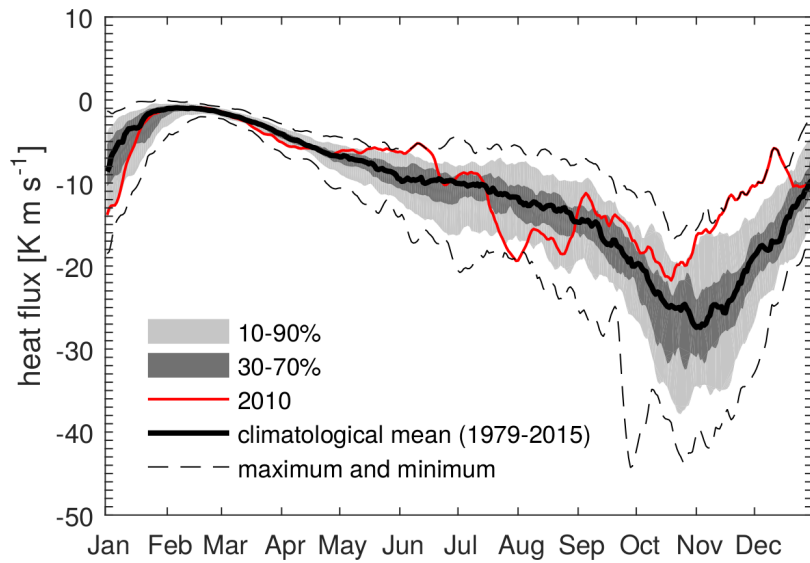


Figure 3. Activity of the polar vortex at 50 hPa as represented by the 45-day running mean of the eddy heat flux between 45 and 75°S. The red curve shows results for the year 2010. Black and gray curves illustrate statistics of the long-term mean (1979–2015). [Data were This analysis was obtained from NASA Ozone Watch from their web site at https://ozonewatch.gsfc.nasa.gov](https://ozonewatch.gsfc.nasa.gov) (last access: 16 December 2016) and is based on MERRA-2 data.

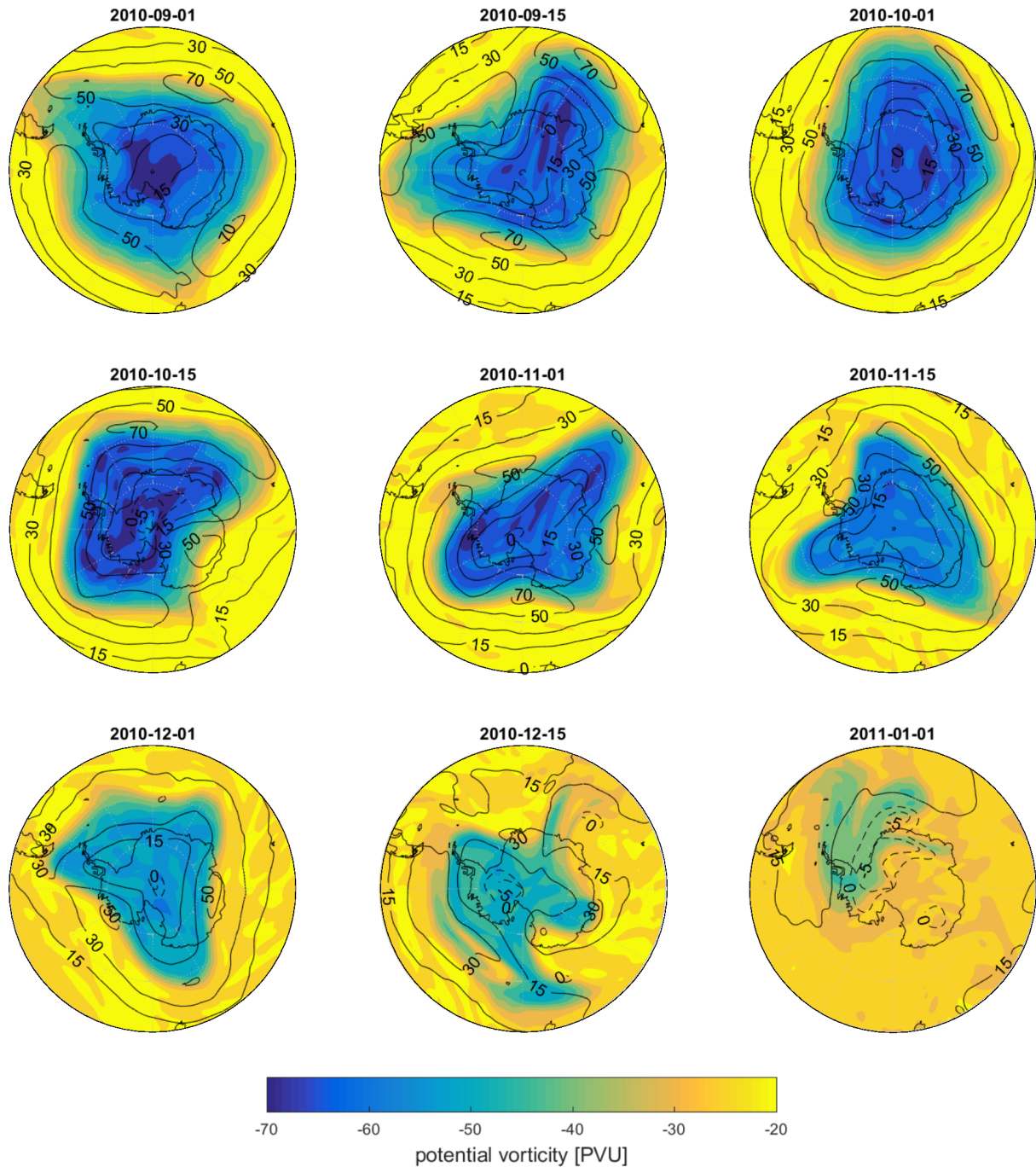


Figure 4. ERA-Interim potential vorticity ($1\text{PVU} = 10^{-6} \text{K m}^2 \text{s}^{-1} \text{kg}^{-1}$; shaded) and zonal wind contours (m s^{-1} ; black curves) on the 475 K isentropic surface. Data are shown for 0 UTC on selected days. [Outer circles of the polar maps indicate a latitude of 45°S. The prime meridian is oriented towards the top of the maps.](#)

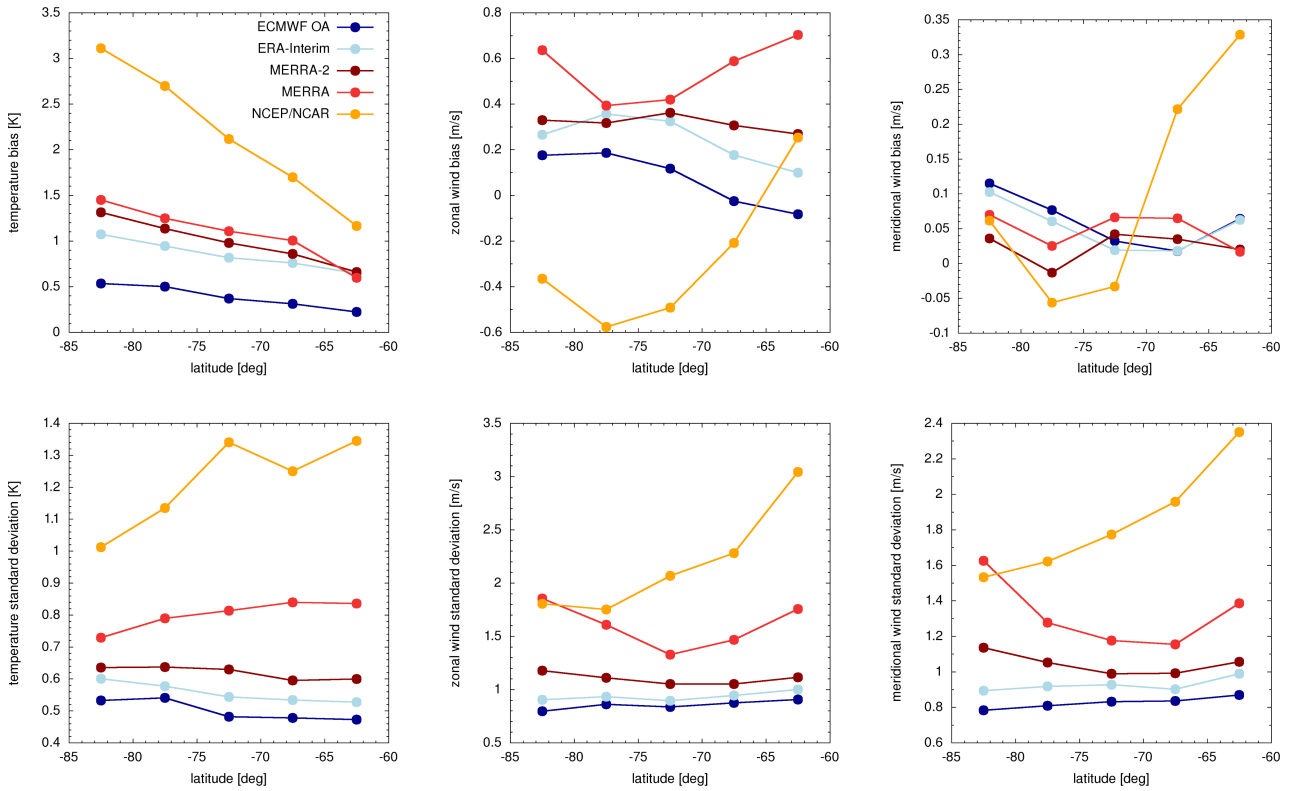


Figure 5. Bias and standard deviations of temperature and horizontal winds of meteorological analyses minus Concordiasi balloon data for at different latitudes.

CONCORDIASI (10V01N46) | 2010-09-23, 02:27 UTC

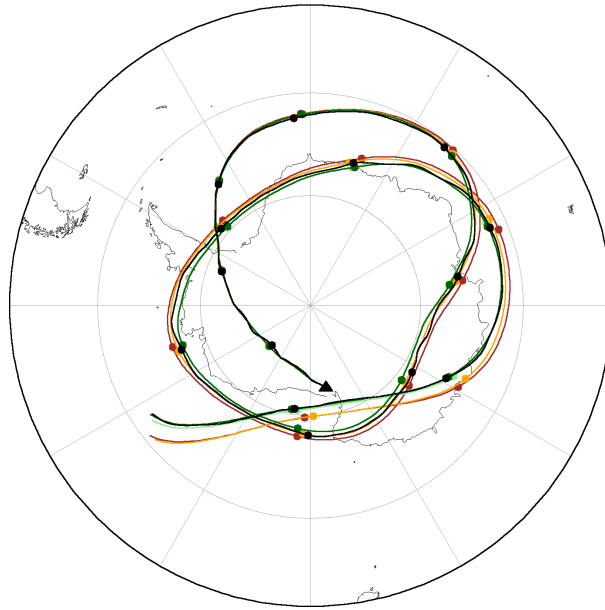


Figure 6. Same as Fig. 5, but for Comparison of 15-day trajectories calculated with different monthstypes of vertical motion (dark green: balloon pressure, light green: isopycnic, orange: isentropic, red: omega velocity) and corresponding Concordiasi balloon trajectory (black). The plot title provides the gondola ID and the starting time. The triangle indicates the starting position of the trajectories. Circles indicate trajectory positions at 0 UTC each day.

Comparison of 15-day trajectories calculated with different types of vertical motion (dark green: balloon pressure, light green: isopycnic, orange: isentropic, red: omega velocity) and corresponding Concordiasi balloon trajectory (black). The plot title provides the starting time and the triangle indicates the starting position of the trajectories. Circles indicate trajectory positions at 0 UTC each day.

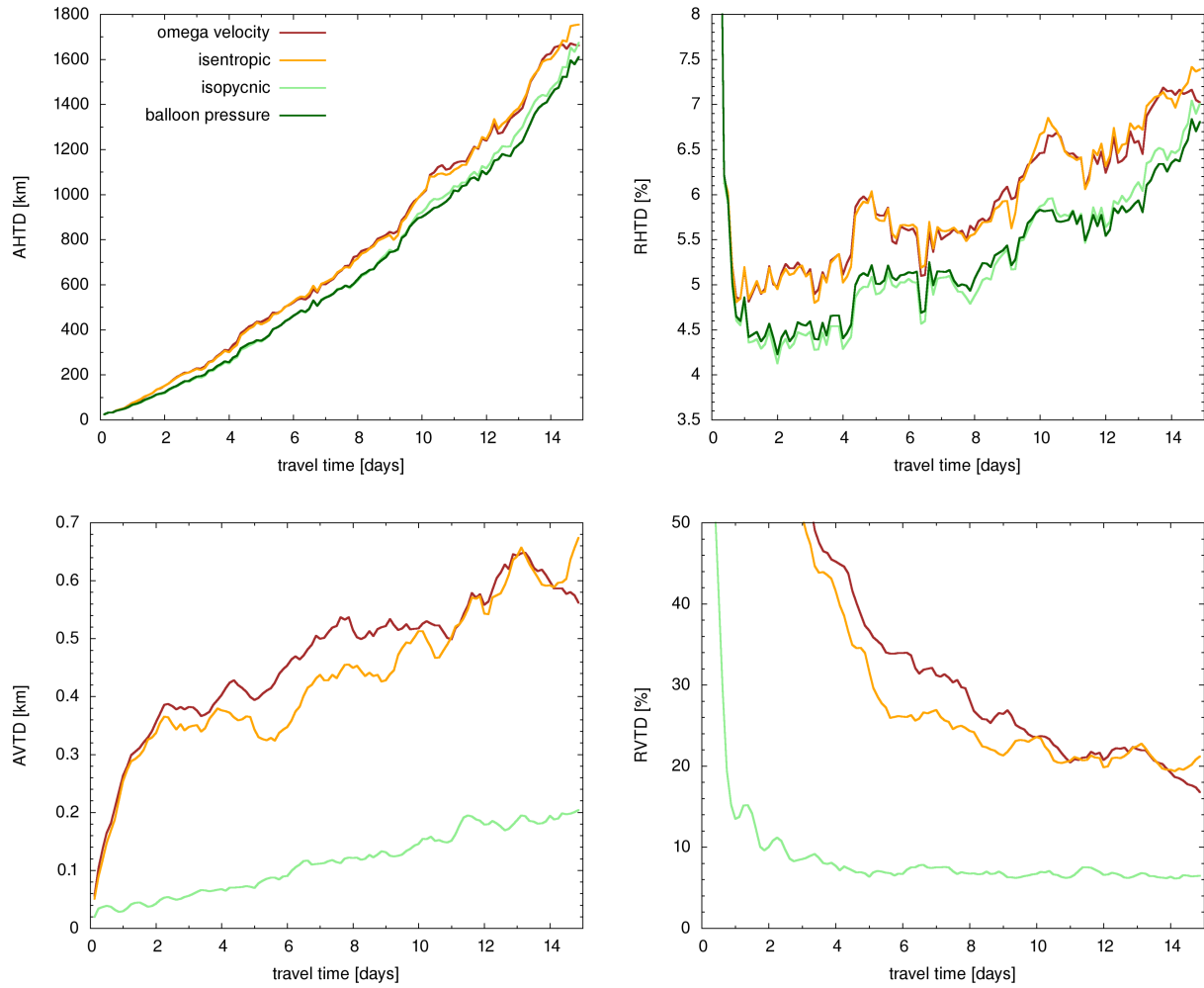


Figure 7. Transport deviations of simulated and observed balloon trajectories for different types of vertical motion. Trajectories were calculated with ERA-Interim horizontal winds. The analysis is based on 104 samples of 15-day trajectories from the Concordiasi campaign.

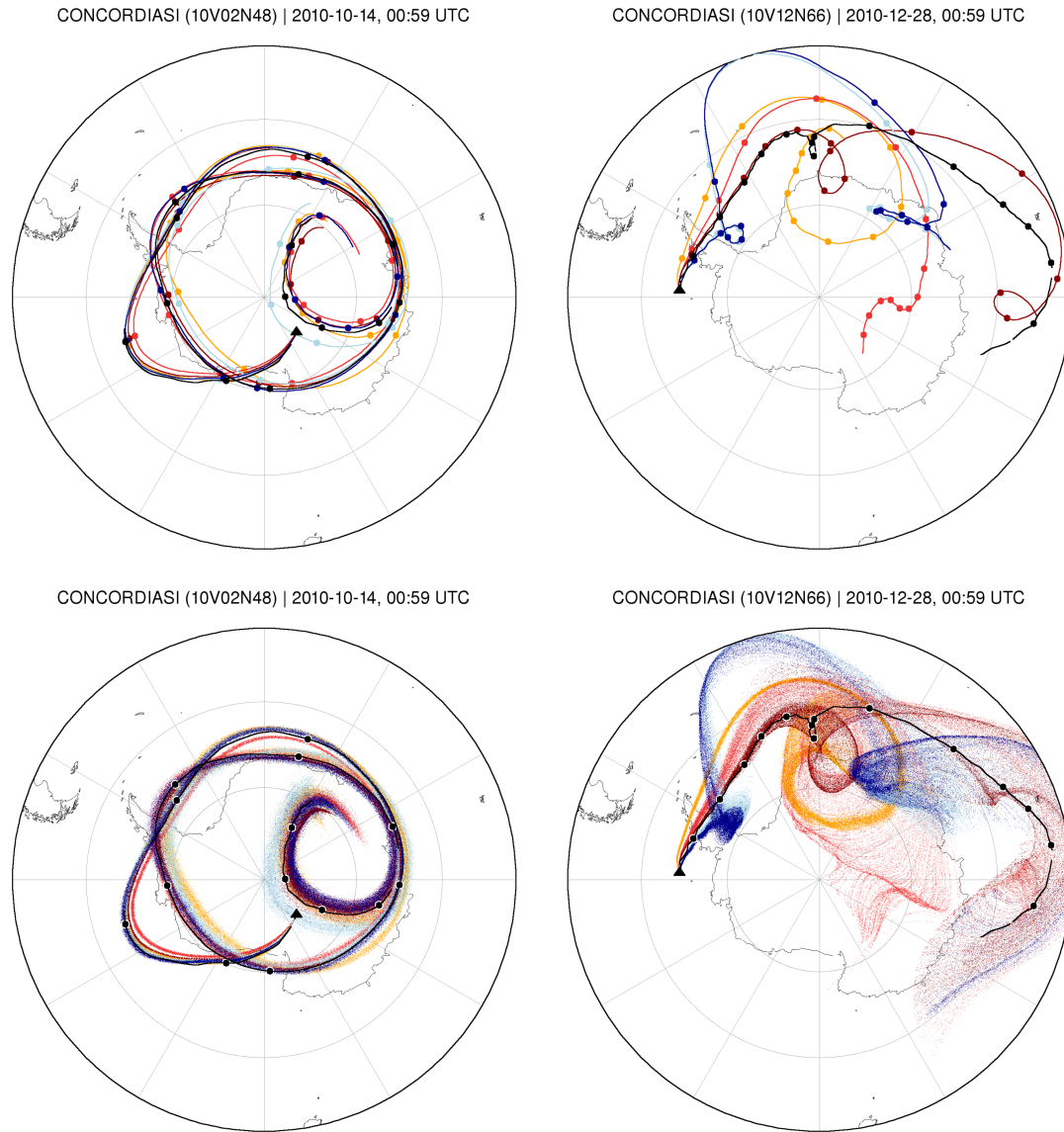


Figure 8. Examples of trajectories calculated with different meteorological analyses (dark blue: ECMWF OA, light blue: ERA-Interim, dark red: MERRA-MERRA-2, light red: MERRA, orange: NCEP/NCAR) and corresponding Concordiasi balloon trajectory (black). Plot titles provide the starting times and triangles indicate the starting positions of the trajectories. Circles indicate trajectory positions at 0 UTC each day. Plots at the top show individual trajectories calculated without diffusion. Plots at the bottom illustrate dispersion simulations with diffusion being considered.

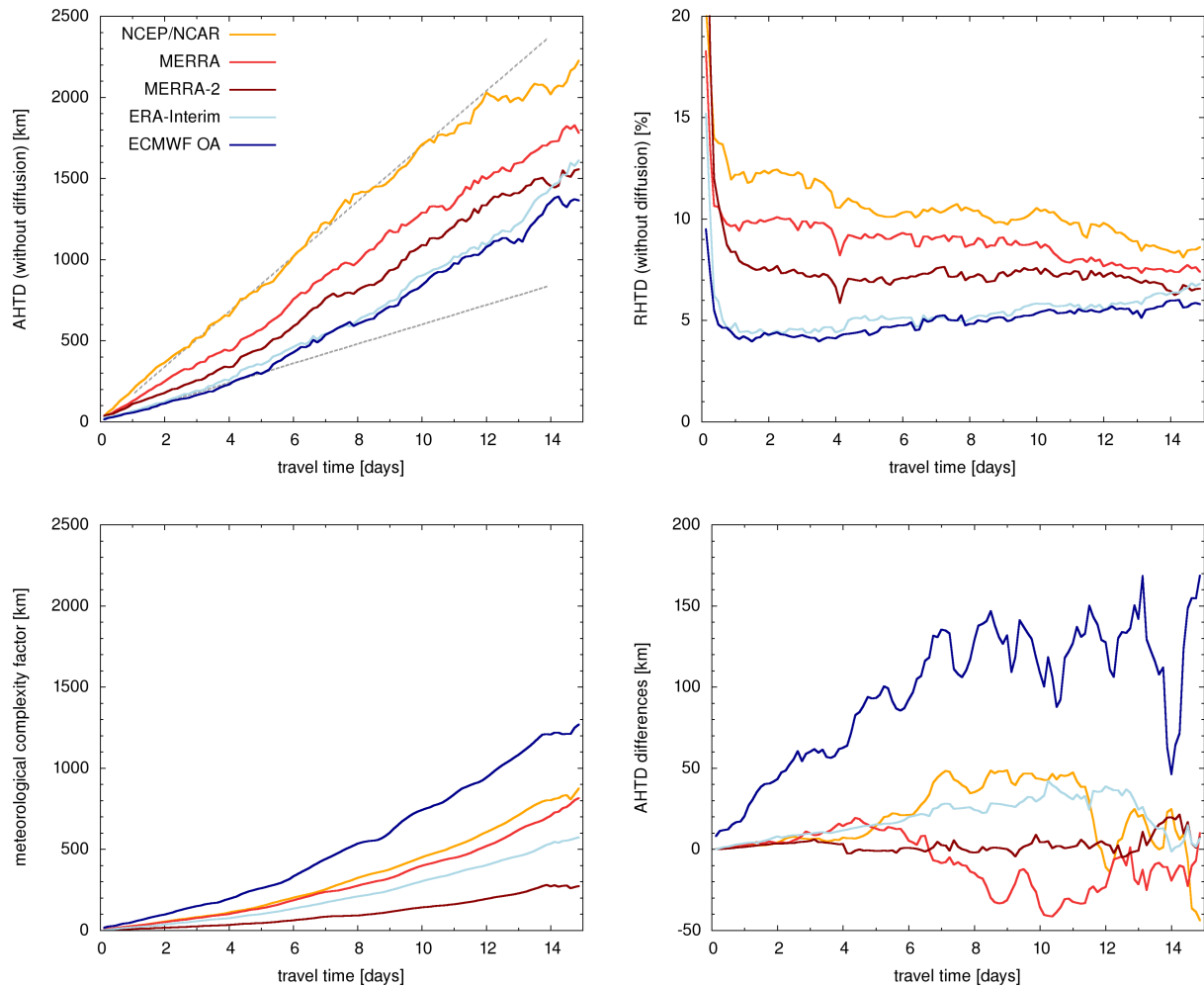


Figure 9. Absolute Horizontal transport deviations of simulated and observed balloon trajectories for different meteorological analyses (top). Dotted gray lines represent AHTD growth rates of 60 and relative 170 km day⁻¹. Also shown are the meteorological complexity factor for dispersion simulations (bottom, left) horizontal transport deviations of and the trajectories shown in Fig. 8 AHTD differences that are introduced by adding diffusion (top bottom, right).

Horizontal transport deviations of simulated and observed balloon trajectories for different meteorological analyses (top). Dotted gray lines represent AHTD growth rates of 60, 120, and 170 km day⁻¹. Also shown are the meteorological complexity factor for dispersion simulations (bottom, left) and the AHTD differences due to diffusion (bottom, right).