Validation of eight atmospheric reanalyses in the Antarctic Peninsula region

Tiina Nygård¹, Timo Vihma¹, Gerit Birnbaum², Jörg Hartmann², John King³, Tom Lachlan-Cope³, Russell Ladkin³, Christof Lüpkes², Alexandra Weiss³

¹Finnish Meteorological Institute, Helsinki, Finland

²Alfred Wegener Institute Helmholz Centre for Polar and Marine Research, Bremerhaven,

Germany

³British Antarctic Survey, Cambridge, UK

Corresponding author: Tiina Nygård, Meteorological Research, Finnish Meteorological Institute,

P.O. Box 503, FI-00101 Helsinki, Finland. (tiina.nygard@fmi.fi)

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/qj.2691

Keywords

Reanalysis, validation, Antarctic Peninsula, vertical profile, research aircraft, spatial resolution

ABSTRACT

Eight atmospheric reanalyses were compared against observed vertical profiles of temperature, specific humidity and wind speed collected by two research aircraft in February–March 2010 in the Antarctic Peninsula region. These data offered a rare possibility to validate reanalyses against independent *in-situ* data which have not been assimilated into the reanalyses. The reanalyses had generally too moist profiles with too low wind speeds, but otherwise the errors in the reanalyses had large spatial differences. On the eastern side of the peninsula, the near-surface temperatures were largely overestimated. None of the reanalysis outperformed the others in all variables, at all altitudes and on both sides of the peninsula. Generally, NCEP-CFSR and MERRA had the smallest errors in temperature profiles, JRA-55 had marginally the most accurate specific humidity profiles and NCEP-CFSR had the best wind profiles. The reanalyses were coherent, although biased, on the western side of the Antarctic Peninsula, but on the eastern side the spread was large. All the reanalyses underestimated the variability between the individual profiles of temperature and wind speed. The modern reanalyses with a sufficient spatial resolution and an adequate data assimilation method outperformed the others, especially on the eastern side.

1. Introduction

A dramatic increase of near-surface air temperature in the Antarctic Peninsula area during the last decades has generated considerable recent research interest in the region. The temperature increase exceeds by several times the global mean temperature rise (Turner et al., 2005), and has also

caused a serious concern about increased melt, loss of ice shelves, acceleration of inland glaciers and increased ice discharge into the ocean. Significant positive trends have also been detected in mid-tropospheric temperatures (Screen and Simmonds, 2012, Turner et al., 2006). The regional climate of the Antarctic Peninsula is sharply divided into two regimes by the mountainous spine, i.e., unbroken mountain chain with peaks 1400–2000 m above the sea level. On the west coast, the climate is relatively mild and maritime, and most of sea ice melts in summer, whereas on the east coast the average temperature is 5–10°C lower, sea ice remains throughout the year, and the climate is more continental (Morris and Vaughan, 1994, Reynolds, 1981). Consequently, the magnitude and significance of observed temperature trends, as well as their seasonality, have also a sharp contrast on either side of the spine; on the western side the warming has been strongest in winter while on the eastern side the maximum has occurred in summer (Ding and Steig, 2013).

Because the observational data from the Antarctic Peninsula, as from the rest of the Antarctic region, are discontinuous, sparse and spatially incomplete, many of the climate studies have at least partly relied on global atmospheric reanalyses (Miles et al., 2008, Nygård et al., 2013, Tietäväinen and Vihma, 2008, Välisuo et al., 2014) or recently also on reconstructions utilizing reanalyses (Monaghan et al., 2008, Nicolas and Bromwich, 2014, Screen and Simmonds, 2012). Global reanalyses combine past observations from stations and satellites with a numerical weather prediction model using a single consistent assimilation scheme. They provide a gridded and dynamically coherent representation of the atmospheric state, which is often also used as boundary forcing for sea ice-ocean models, land surface models and atmospheric limited-area models and as first-order validation for climate models. Despite the noticeable progress made in the representation of the Antarctic in reanalyses, spurious trends have been reported in many variables such as precipitation and evaporation/sublimation (Bromwich and Fogt, 2004, Bromwich et al., 2007, Bromwich et al., 2011, Cullather and Bosilovich, 2011).

Previous validation studies have mainly focused on long-term and large-scale features in reanalyses, and only a few validation studies have addressed regional atmospheric conditions in the Antarctic and short-term variability. Välisuo et al. (2014) validated near-surface variables in the Antarctic Peninsula by comparing three reanalyses against automatic weather station data. They found a consistent warm bias and significant biases in wind components. Välisuo et al. (2014) supported the conclusion of Stössel et al. (2011) that the coarse resolution of reanalyses did not represent the complex orography of the Antarctic Peninsula accurately enough from the point of view of modelling of the wind field. Due to the inadequate resolution, the westerlies were overestimated and the northward wind component was underestimated. Välisuo et al. (2014) also found that many of the biases in the surface energy balance components partly compensated each other, as also reported by Vancoppenolle et al. (2011) for sea ice on both sides of the Antarctic Peninsula and found that diurnal cycles of near-surface variables were largely erroneous in all of the reanalyses.

In this paper, we present validation of reanalyses with respect to vertical profiles of air temperature, specific humidity and wind speed in the Antarctic Peninsula region. Eight reanalyses, including older generation and modern ones, are compared against vertical profiles measured by two research aircraft in February–March 2010. These measurements offered a rare possibility to evaluate the reanalyses using independent *in-situ* data which had not been assimilated into any of the reanalyses. We quantify the accuracy of each reanalysis with a special focus on the spatial variability of errors around the Antarctic Peninsula. We also discuss progress made since the first reanalyses, probable reasons for the errors as well as consequences of the errors.

2. Observations and reanalyses

A research aircraft campaign was carried out in February–March 2010, using the research station Rothera (67.6°S, 68.1°W) of British Antarctic Survey as a base. During the campaign in the region of the Antarctic Peninsula, a DHC6 Twin Otter operated by the British Antarctic Survey (BAS) and a Basler BT-67 ("Polar 5") operated by the Alfred Wegener Institute (AWI) flew 95 and 49 vertical profiles, respectively. The vertical profiles contained detailed information on temperature, specific humidity and wind speed. Temperature on Polar 5 was measured with a Rosemount Pt100 sensor. The wind speed data of Polar 5 were based on a Rosemount 5-hole probe at the noseboom with Rosemount pressure transducers for flow measurements in the aircraft system and a Honeywell Lasernav INS and TrimbleGPS for aircraft attitude and speed with respect to the Earth. The humidity was measured using a dew point mirror. Temperature on BAS Twin Otter was measured with a Goodrich Rosemount sensor 102, atmospheric humidity with a cooled-mirror dew point hygrometer and wind speed with a nine-hole turbulence probe (BAT probe). The instrumentation of the BAS Twin Otter is described in more detail in Weiss et al. (2011) and the instrumentation of AWI Polar 5 in Tetzlaff et al. (2015). A close formation comparison flight between the two aircraft revealed a temperature difference of 0.7 K and a wind speed difference of 0.5 m s⁻¹ on average (Figure 1), and these can also be considered as estimations for measurement errors.

The flights were made both on the western and eastern side of the Antarctic Peninsula and over various surface types (Figure 2). In the data analysis, the flights were divided into four categories depending on their location with respect to the peninsula and surface type (Figure 2). Data of the sub-region W1 were collected over the open sea west of the peninsula or in the vicinity of Wilkins Ice Shelf. W2 data were from the vicinity of complex orography either over land or over the George VI Ice Shelf west of the peninsula. On the eastern side, E1 data were from the ice shelves and E2 from sea ice. All the four categories contained roughly the same number of profiles and measurements with the both aircraft. The bottom of the profiles was typically at 40–1000 m above the surface and the top at 1000–4000 m. All observations were finally averaged to 100 m intervals in vertical. In this study, we present average profiles from 100 to 3000 m.

The aircraft observations represented independent data which were not assimilated into any reanalysis. To evaluate impacts of data assimilation of vertical profiles and near-surface data, we also utilized radio soundings from the research stations Rothera and Marambio (64.2°S, 56.7°W) (Figure 2) from the period of measurement campaign. The number of radiosoundings used in the comparison with the reanalyses was 7 from Rothera and 10 from Marambio. Only temperature and specific humidity were analysed based on the radio soundings because the vertical resolution of wind speed profiles was not adequate.

Assimilated fields of eight reanalyses were investigated in this study. An overview of the reanalyses is given in Table 1. National Center for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) (Kalnay et al., 1996) was the first widely used global reanalysis and is also so far the most frequently used reanalysis for the Antarctic. Later, its improved version by NCEP and the Department of Energy (NCEP-DOE) was released (Kanamitsu et al., 2002). The third older generation reanalysis used in this study was the Japan Meteorological Agency Climate Data Assimilation System (JCDAS) (Onogi et al., 2007). The newer reanalyses included in the study were NCEP Climate Forecast System Reanalysis (NCEP-CFSR) (Saha et al., 2010), Japanese 55-year reanalysis (JRA-55) (Kobayashi et al., 2015), ERA-Interim of the European Centre for Medium-Range Weather Forecasts (Dee et al., 2011), Modern Era Retrospective Analysis for Research and Applications (MERRA) (Rienecker et al., 2011) and the 20th Century Reanalysis version 2 (20CR) (Compo et al., 2011). 20CR is the only of the reanalyses made for century-scale climate studies as it spans from 1871 to 2012, and was taken into this study as an example of century-scale reanalyses. A wide range of observations have been assimilated into the reanalyses by applying either the three-dimensional variational data assimilation (3D-VAR) (used in NCEP/NCAR, NCEP-DOE, NCEP-CFSR, JCDAS, MERRA) or its extended system 4D-VAR (used in ERA-Interim and JRA-55). However, 20CR assimilates only the surface pressure, monthly sea surface temperature and sea ice distribution. NCEP-CFSR, MERRA and ERA-Interim have the

highest spatial grid resolutions $(0.5-0.75^{\circ} \times 0.5-0.75^{\circ})$, whereas NCEP/NCAR, NCEP-DOE and 20CR have the lowest ones $(2-2.5^{\circ} \times 2-2.5^{\circ})$. The effective spatial resolution of reanalyses, which tells the scale of phenomena resolved, is roughly 5–7 times the grid resolution (Skamarock, 2004). The number of vertical model levels varied between the reanalyses, but all the data of this study were taken from the pressure levels available (12-13) pressure levels in the lowest 3000 m). To facilitate comparison, the reanalysis data were linearly interpolated to each observation location, observation time and to 100 m vertical intervals. Spatial interpolation was made using the nearest grid box around the measurement site. Treatment of sea ice in each reanalysis is listed in Table 1. The surface temperature over sea ice, ice sheets, and ice shelves in all reanalyses is based on a prognostic scheme, with a few model levels in ice/snow.

3. Results

3.1 Temperature

The observed mean temperature in the lowest hundreds of meters was nearly 10 K lower on the eastern side of the Antarctic Peninsula (E1 and E2) than on the western side (W1 and W2) (Figure 3). The near-surface temperature was lowest over the eastern ice shelves, where a remarkably stable layer reached the height of approximately 800 m. Hence, the observed near-surface temperatures reflected the mean spatial distribution of surface temperature for 1965–1980 presented by Reynolds (1981) despite our relatively little data amount and short study period. The temperature profiles on the western side, especially in W1, were generally well captured by all the reanalyses, having only small variations of errors with the altitude (Figure 3). The temperature bias was mostly negative, except in the lowest 1500 m in NCEP/NCAR and NCEP-DOE. On the eastern side, in turn, the reanalyses had a temperature spread of 7 K in the lowest hundreds of meters. Over the ice shelves (E1) in the east, all the reanalyses represented the temperature in the lowest 1000 m poorly having 3–8 K too high values and too weakly stable stratification. Above 1000 m, the root-mean-square-errors (RMSE) were much smaller than at lower altitudes and biases were mainly negative also on

the eastern side. Over the sea ice (E2), further away from the immediate influence of the peninsula, NCEP/NCAR, NCEP-DOE and 20CR stood out with the largest bias and RMSE. These reanalyses had very similar temperature profiles on both sides of the peninsula, although the mean observed temperature profile was different on each side. The layer of large errors, particularly of coarse-resolution reanalyses, reached from the surface to approximately 500 m in E1 and 1000 m in E2. The variability of biases and RMSEs was notably larger than the observed temperature difference between the two aircraft during the formation comparison flight (Section 2). This suggests that the results for reanalysis errors were not largely affected by measurement errors.

Temperature inversions, i.e., layers where temperature increases with height, were identified within the lowermost 1000 m in 42 of the observed profiles. In MERRA, all of these 42 temperature inversions were qualitatively captured. MERRA, however, also included temperature inversions approximately in 30 % cases in which a temperature inversion did not exist in the observed profile. NCEP-CFSR was the second best to capture temperature inversions (captured 25 temperature inversions out of 42), JRA-55 and ERA-Interim captured less than half of them (15–17) and NCEP-DOE, 20CR, NCEP/NCAR and JCDAS captured only a fraction of them (6–11).

With respect to the RMSE of temperature, MERRA and NCEP-CFSR were the best reanalyses on the both sides of the peninsula, when the whole vertical profile was considered (Table 2). They had also the smallest bias in the eastern side of the peninsula, but on the western side the bias was smallest in NCEP/NCAR and NCEP-DOE. Below 1000 m altitude on the western side, ERA-Interim was also equally good with NCEP/NCAR and NCEP-DOE. On the eastern side, JRA-55 and MERRA had the most realistic temperature near the surface. However, MERRA had a distinct shape of temperature profile on both sides as it was typically colder than the other reanalyses in the lowest 500 m and warmer above 1000 m (Figure 3). Progress made from NCEP/NCAR reanalysis to NCEP-CFSR was clearest in the lowest hundreds of meters on the eastern side. In case of Japanese reanalyses, the temperature RMSE was notably reduced in JRA-55 in the eastern side of

the peninsula compared to earlier JCDAS, but the temperature bias was even more negative in JRA-55.

Accepted Article

Standard deviations of the observed profiles demonstrated relatively large variability between the individual temperature profiles (Figure 4). Practically all the reanalyses underestimated the variability compared to the observations, especially in the lowermost 1000 m. Near the surface, NCEP-DOE and MERRA had the standard deviation closest to the observed one, whereas 20CR was clearly the most unrealistic with respect to variability. Above 1000 m, the standard deviations of the reanalyses were relatively well in accordance with the observations.

In comparison against the radiosounding data from Rothera and Marambio, the temperature bias in the reanalyses (between -4 and 0 K) was found to be approximately at the same level as in comparison against the independent aircraft data in W1 and W2 (between -3 and +1 K). We also compared reanalyses against the aircraft observations in the vicinity of Rothera on days when radio soundings were made and on days with no radio soundings. The differences were not systematic; in MERRA, near-surface temperature bias and RMSE were reduced to half on the days with radio soundings, in NCEP/CFSR temperature bias and RMSE were 1 K smaller on the days with no radio soundings, whereas in many reanalyses differences were very small.

3.2 Humidity

The observed specific humidity was over 1 g kg⁻¹ higher in W1 than over the ice shelves of E1 (Figure 5). The shape of the observed specific humidity profiles differed notably between the four sub-regions, although above 1000 m the specific humidity was nearly the same, 2 g kg⁻¹, in W1, W2 and E1. The reanalyses had clearly too humid profiles (Figure 5), except that MERRA was approximately 0.5 g kg⁻¹ too dry in the lowest hundreds of meters. The largest problems were found in E1 in the lowest 100 m where there was a substantial spread between the reanalyses. There, the specific humidity in different reanalyses ranged from 1.2 to 3.0 g kg⁻¹. Nearly similar

discrepancy between the reanalyses was found in the lowest hundreds of meters in E2. In the vicinity of the peninsula (W2 and E1), the RMSE increased towards the surface, whereas further away from the peninsula the errors were rather independent of the altitude. The observed mean profiles included layers where the humidity increased with height, i.e., humidity inversions. These humidity inversions, particularly near the surface, were essentially captured by the reanalyses (not shown).

With respect to specific humidity profiles, none of the reanalyses was clearly better than the others. The total RMSE was smallest in JRA-55, but it was also practically on an equal level in MERRA, NCEP-CFSR, ERA-Interim, and JCDAS (Table 2). The specific humidity bias, in turn, was smallest in JCDAS, JRA-55 and MERRA. The positive bias was generally largest in NCEP/NCAR and NCEP-DOE throughout the profile from the surface to 3000 m (Figure 5). 20CR, in turn, had the largest errors in the profile shape because it was clearly too humid in the lowest 1000 m and in W1, W2 and E1 too dry above. The humidity bias profiles of the old-generation models (JCDAS, NCEP-DOE, NCEP/NCAR) and 20CR resembled the shape of their temperature bias profiles (Figures 3 and 5). Temperature and specific humidity errors of those reanalyses had a positive (from 0.23 to 0.40) and statistically significant Spearman correlation (which denotes the strength of the monotonic relationship which is not necessarily linear) whereas in the other reanalysis this correlation was weaker and significantly negative (the correlation coefficient -0.18--0.20). NCEP-CFSR had nearly 0.5 g kg⁻¹ smaller bias than its predecessors NCEP/NCAR and NCEP-DOE in the lowest 1000 m. The standard deviation of specific humidity had little variability with the height both in the observations and reanalyses (Figure 4). Particularly NCEP-DOE and NCEP/NCAR overestimated the standard deviation of specific humidity whereas 20CR and JCDAS clearly underestimated the standard deviation.

The bias and RMSE of specific humidity were notably smaller for the assimilated radiosounding data from Rothera compared to the independent aircraft data from W2 (not shown); the near-surface

RMSE was approximately 0.5 g kg⁻¹ at Rothera and nearly 1 g kg⁻¹ in W2. At Marambio, the bias and RMSE of specific humidity were at the same level as for the aircraft data in E1 and E2, showing no detectable impacts of data assimilation. When aircraft observations around Rothera were compared on days with and without radio soundings, no systematic differences were found.

3.3 Wind speed

The shapes of observed wind speed profiles were rather similar in sub-regions W1, W2 and E1, but in E2 a low-level jet was found in the mean profile (Figure 6). Wind speed variability with height was common in all the sub-regions. The reanalysis profiles were much smoother than the observed profiles and had no clear sign of a low-level jet in E2. They also underestimated the wind speed on both sides of the peninsula. The reanalysis wind speeds were similar on both sides of the peninsula, although below 2000 m the observed wind speed was higher on the eastern side. The wind speed was generally better captured near the surface than between 300 m and 1500 m. The reanalyses were surprisingly similar to each other in E1, although they clearly deviated from the observed profile. The large wind speed bias in E1 was mainly due to an underestimated westerly wind component (not shown). In E2, the spread between the reanalyses was as large as 2-3 m s⁻¹ and especially the v wind component was largely underestimated in the coarse-resolution reanalyses. In the other reanalyses, contributions of the u and v wind components to the bias in E2 were nearly equal. Compared to the estimated measurement errors (Section 2), differences in RMSE of wind speed between the reanalyses were larger.

Considering the RMSE, NCEP-CFSR had the most realistic wind speed profiles, and JRA-55 was nearly as good (Table 2). The bias was smallest, and practically on an equal level, in NCEP-CFSR, MERRA, NCEP-DOE, and JRA-55. 20CR largely underestimated the wind speed and had the largest total RMSE (Table 2). The other coarse-resolution reanalyses NCEP/NCAR and NCEP-DOE had also poor wind speed profiles, but their total wind speed RMSE was on the same level as

in MERRA and JCDAS. Reanalyses were coherent with respect to the standard deviation of wind speed, but they all clearly underestimated the standard deviation (Fig. 4). 20CR had the smallest, and most unrealistic, standard deviation and NCEP-CFSR had the largest standard deviation, except within the lowest 500 m. Wind speed biases and RMSEs were not systematically different in the vicinity of Rothera on days when radio soundings were made compared to days with no radio soundings.

Representation of orography varied largely between the reanalyses, and it had an influence on the wind field around the Antarctic Peninsula. As an example, surface geopotential height and 925 hPa mean wind components at of ERA-INTERIM, JCDAS and NCEP/NCAR are shown in Figure 7. It is noteworthy that wind components in Figure 7 are interpolated values in areas where the surface geopotential of reanalysis was above the level of 925 hPa. ERA-INTERIM had the most realistic orography of these three, whereas representation of the peninsula was much simpler in JCDAS and NCEP/NCAR. For example, the Larsen Ice Shelf was on a mountain slope in JCDAS and NCEP/NCAR. The main differences in the wind components were (i) the strength of westerly winds in Bellingshausen Sea, (ii) the sign of the u wind component on the eastern slope of the peninsula, (iii) the strength of northerly winds on the west coast of the peninsula and (iv) the sign of the v wind component over the Larsen Ice Shelf. Some of these features were obviously influenced by the surface geopotential field (Figure 7). This was particularly true for the v wind component. Nevertheless, orographic impacts could not be quantified and separated from other differences between the reanalyses.

4. Discussion and conclusions

The reanalyses had generally too moist profiles with too low wind speeds in the Antarctic Peninsula region. In addition, the near-surface temperatures were largely overestimated on the eastern side of the peninsula and the near-surface stability was underestimated over the Larsen ice shelf (E1). Our results were qualitatively similar to previous reanalysis validation results from the Central Arctic.

For example, Lüpkes et al. (2010) and Jakobson et al. (2012) reported a warm and moist bias. In Jakobson et al. (2012) this bias was combined with too low wind speeds in the lowermost hundreds of meters in spring and summer. The errors in the vertical structure in reanalyses on the western side of the Antarctic Peninsula, which were reported in this study, were smaller than in the Central Arctic. On the eastern side of the peninsula, the RMSEs of temperature and specific humidity were approximately equal to those in the Arctic but the RMSE of wind speed was higher.

Our analysis revealed that it is a very challenging task to make an objective ranking of the reanalysis around the Antarctic Peninsula. None of the reanalysis outperformed the others in all variables, at all altitudes and on both sides of the peninsula. It is also noteworthy that our results were only from late summer and it is likely that the performance of each reanalysis varies between the seasons. On the western side of the peninsula, differences between the reanalyses were relatively small. On the contrary, on the eastern side the spread between the reanalyses was large, and a user of a reanalysis should pay attention in choosing the most accurate one. On the eastern side, the newer reanalyses with a sufficient spatial resolution and an adequate data assimilation method (3D- or 4D-VAR) (i.e., NCEP-CFSR, ERA-Interim, MERRA and JRA-55) had the smallest errors in the vertical profiles. Generally, MERRA and NCEP-CFSR had the smallest vertical and spatial mean errors in the temperature profiles. It is, however, interesting that despite the relatively small vertical mean RMSEs in MERRA, its profiles often had a distinct shape. JRA-55 had the most accurate specific humidity profiles with a very small marginal, and the vertical mean errors in wind speed profiles were smallest in NCEP-CFSR. Progress made since the first reanalyses was clearest on the eastern side. For example, NCEP-CFSR had considerably smaller specific humidity RMSE throughout the profile and temperature RMSE in the lowermost 1000 m than its predecessors NCEP/NCAR and NCEP-DOE. Part of these improvements most probably stem from the better spatial resolution of NCEP-CFSR. The good performance of NCEP-CFSR

over sea ice (E2) may be connected to the fact that NCEP-CFSR is the only one of the reanalyses which is coupled with a sea-ice and ocean model.

An adequate spatial resolution is primarily needed to resolve the orography of the Antarctic Peninsula, which interrupts the otherwise zonal wind field. Especially temperature and wind speed profiles of the coarse-resolution reanalyses (NCEP/NCAR, NCEP-DOE and 20CR) were very similar on both sides of the peninsula and were clearly little influenced by the orography of the peninsula, although differences between the sides were seen in the observations. Based on our results, impacts of orography seem to reach from the surface to at least 1000 m altitude, and these impacts are mainly found on the eastern side of the peninsula due to prevailing westerly winds. Stössel et al. (2011) concluded that, in addition to spatial resolution of reanalysis, the underlying representation of orography in the reanalyses has a large influence on wind fields. Stössel et al. (2011) discussed that in the operational ECMWF analyses barrier winds on the western side Antarctic Peninsula are captured, but in the NCEP/NCAR reanalysis the orography is poorly resolved which leads to a lack of barrier winds and spurious offshore winds on the eastern side. Our results demonstrated that the reanalyses have large differences in the representation of the orography and it has an evident impact on the flow in the Antarctic Peninsula region. Unfortunately, the impacts of horizontal resolution cannot be separated from other differences between the reanalyses. Quantitative investigation on impacts of resolution would require controlled experiments with a reanalysis applying several different spatial resolutions.

Another probable reason for the differences in reanalysis performance between the western and eastern sides of the peninsula is related to the strong stable stratification over the Larsen Ice Shelf and adjacent sea ice. Stable stratification in this region can result from (i) northward advection of cold, continental air in a barrier jet, (ii) surface cooling of warm air originating in a foehn wind blowing down from the Peninsula mountains (King et al., 2008) or (iii) nocturnal cooling. The stable stratification is typically challenging for numerical models (Steeneveld, 2014), and the

results suggest that also reanalyses are not able to realistically reproduce the strong stable stratification at low levels in E1 and E2.

The amount of observational data assimilated into reanalyses may also contribute to spatial differences in errors. Regular radiosoundings were available from the research station Rothera on the western side and the research station Marambio on the tip of the peninsula. Errors in specific humidity were notably smaller for the assimilated radio-sounding data from Rothera compared to the independent aircraft data, but at Marambio clear differences were not found. On the other hand, our results suggested that temperature profiles were little influenced by the assimilation of Rothera and Marambio radio soundings. Small impacts of Rothera and Marambio data might, for example, indicate that these observations were simply not utilized in the assimilation due to quality issues or that the offset between first-guess and observation was typically small.

Since observations are scarcely available in the Antarctic, satellite data are the main observational data sources in the region. The cold bias (0–2 K) seen on the western side of the Antarctic Peninsula and above 1000 m on the eastern side is consistent with the cold bias found in the satellite data (Boylan et al., 2015, Wang et al., 2013). This can be an indication that the cold bias in the satellite data contributes to the cold bias of the reanalyses. Furthermore, assimilation of satellite data over open ocean in the west had probably less challenges than assimilation of satellite data over ice-covered surfaces in the east, which, in turn, also increased the errors in the east compared to the west. However, it is remarkable that 20CR, which is a century-scale reanalysis and did not assimilate radiosondes or satellite data, performed almost as well as NCEP/NCAR and JCDAS. The temperature bias of 20CR was even smaller than in many of the other reanalyses, although the RMSEs of temperature, specific humidity and wind speed were high. On the other hand, 20CR largely underestimated the variability between individual profiles and was clearly the most unrealistic reanalysis with respect to variability.

In the Antarctic, reanalyses have earlier been applied to estimate, for instance, the atmospheric moisture budget (Tietäväinen and Vihma, 2008) and moisture advection (Nygård et al., 2013). These kinds of calculations are dependent on accurate wind and humidity profiles, and also require that errors do not vary much with space and height. Our results indicated that the humidity difference, and especially the temperature difference, between the western and eastern side of the Antarctic Peninsula were underestimated by the reanalyses, and wind speed was simultaneously underestimated. These errors suggest that zonal advection of heat and moisture might be generally underestimated by the reanalyses in the region. Despite the deficiencies, the-state-of-the-art reanalyses, however, provide the best available information on the three-dimensional structure of the atmosphere, and can be utilized in climatological studies and as forcing for sea ice – ocean models when attention is paid to the uncertainties. In the future, better spatial resolution and representation of orography will probably yield significantly smaller errors in the vertical profiles and less spatial variability of the errors in the Antarctic Peninsula region.

Acknowledgements

This work was supported by the Academy of Finland (grant 263918) and the European Facility for Airborne Research (EUFAR). The Twin Otter flights were funded by the UK Natural Environment Research Council through the British Antarctic Survey's core programme "Polar Science for Planet Earth". We acknowledge ECMWF for providing the ERA-Interim data, NCAR for providing NCEP/NCAR, NCEP/DOE, NCEP/CFSR, JCDAS, JRA-55 and 20CR data (http://rda.ucar.edu/), and NASA Goddard Earth Sciences Data and Information Services Center for MERRA data (<u>http://disc.sci.gsfc.nasa.gov/mdisc/data-holdings</u>). We are grateful to the anonymous reviewers for valuable comments on the manuscript.

- Boylan P, Wang J, Cohn SA, Fetzer E, Maddy ES, Wong S. 2015. Validation of AIRS version 6 temperature profiles and surface-based inversions over Antarctica using Concordiasi dropsonde data. J. Geophys. Res. Atmos., 120: 3, 992-1007, DOI: 10.1002/2014JD022551
- Bromwich DH, Fogt RL. 2004. Strong trends in the skill of the ERA-40 and NCEP-NCAR reanalyses in the high and midlatitudes of the southern hemisphere, 1958-2001. *J. Clim.*, 17: 23, 4603-4619, DOI: 10.1175/3241.1
- 10.1175/3241.1 Bromwich DH, Fogt RL, Hodges KI, Walsh JE. 2007. A tropospheric assessment of the ERA-40, NCEP, and JRA-25 global reanalyses in the polar regions. *J. Geophys. Res. Atmos.*, 112: D10, D10111, DOI: 10.1029/2006JD007859
 - Bromwich DH, Nicolas JP, Monaghan AJ. 2011. An Assessment of Precipitation Changes over Antarctica and the Southern Ocean since 1989 in Contemporary Global Reanalyses. J. Clim., 24: 16, 4189–4209, DOI: 10.1175/2011jcli4074.1
 - Cavalieri DJ, Gloersen P, Campbell WJ. 1984. Determination of sea ice parameters with the NIMBUS 7 SMMR. J. Geophys. Res. Atmos., 89: D4, 5355-5369, DOI: 10.1029/JD089iD04p05355
 - Compo GP, Whitaker JS, Sardeshmukh PD, Matsui N, Allan RJ, Yin X, Gleason BE, Vose RS, Rutledge
 G, Bessemoulin P, Brönnimann S, Brunet M, Crouthamel RI, Grant AN, Groisman PY, Jones PD,
 Kruk MC, Kruger AC, Marshall GJ, Maugeri M, Mok HY, Nordli Ø, Ross TF, Trigo RM, Wang
 XL, Woodruff SD, Worley SJ. 2011. The Twentieth Century Reanalysis Project. *Q. J. R. Meteorol.*Soc., 137: 654, 1-28, DOI: 10.1002/qj.776
- Cullather RI, Bosilovich MG. 2011. The Moisture Budget of the Polar Atmosphere in MERRA. *J. Clim.*, 24: 11, 2861-2879, DOI: 10.1175/2010jcli4090.1
 - Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Hólm EV, Isaksen L,

Kållberg P, Köhler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette JJ, Park BK, Peubey C, de Rosnay P, Tavolato C, Thépaut JN, Vitart F. 2011. The ERA-Interim reanalysis:

- Ding Q, Steig EJ. 2013. Temperature Change on the Antarctic Peninsula Linked to the Tropical Pacific. J. Clim., 26: 19, 7570-7585, DOI: 10.1175/JCLI-D-12-00729.1 Ishii M, Shouji A, Sugimoto S, Matsumoto T. 2005. Objective analyses of sea-surface temperature and marine meteorological variables for the 20th century using ICOADS and the Kobe Collection J. Climatol., 25: 7, 865-879, DOI: 10.1002/joc.1169
 - reanalyses over the central Arctic Ocean. Geophys. Res. Lett., 39: 10, DOI: 10.1029/2012gl051591
 - Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Leetmaa A, Reynolds R, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Jenne R, Joseph D. 1996. The NCEP/NCAR 40-Year Reanalysis Project. Bull. Am. Meteorol. Soc., 77: 3, 437-471, DOI: 10.1175/1520-0477
 - **JI**C Kanamitsu M, Ebisuzaki W, Woollen J, Yang S-K, Hnilo JJ, Fiorino M, Potter GL. 2002. NCEP-DOE AMIP-II Reanalysis (R-2). Bull. Am. Meteorol. Soc., 83: 11, 1631-1643, DOI: 10.1175/BAMS-83-11-1631
 - King JC, Lachlan-Cope TA, Ladkin RS, Weiss A. 2008. Airborne Measurements in the Stable Boundary Layer over the Larsen Ice Shelf, Antarctica. Boundary-Layer Meteorol., 127: 3, 413-428, DOI: 10.1007/s10546-008-9271-4
 - Kobayashi S, Ota Y, Harada Y, Ebita A, Moriya M, Onoda H, Onogi K, Kamahori H, Kobayashi C, Endo H, Miyaoka K, K. T. 2015. The JRA-55 Reanalysis: General Specifications and Basic Characteristics. J. Meteor. Soc. Japan, 93: 1, DOI: 10.2151/jmsj.2015-001

Lüpkes C, Vihma T, Jakobson E, König-Langlo G, Tetzlaff A. 2010. Meteorological observations from ship cruises during summer to the central Arctic: A comparison with reanalysis data. Geophys. Res.

- for Climate Monitoring. in *AMS 14th Conference on Satellite Meteorology and Oceanography*, edited, Atlanta Miles GM, Marshall GJ, McConnell JR, Aristarain AJ. 2008. Recent accumulation variability and change on the Antarctic Peninsula from the ERA40 reanalysis. *Int. J. Climatol.*, 28: 11, 1400 + 477-10.1002/joc.1642 Monaghan AJ, Brown

 - near-surface temperature. J. Geophys. Res. Atmos., 113: D4, D04105, DOI: 10.1029/2007JD009094 Morris EM, Vaughan DG. 1994. Snow surface temperatures in West Antarctica. Antarct. Sci., 6: 4, 529–

535

- Nicolas JP, Bromwich DH. 2014. New Reconstruction of Antarctic Near-Surface Temperatures: Multidecadal Trends and Reliability of Global Reanalyses. J. Clim., 27: 21, 8070-8093, DOI: 10.1175/jcli-d-13-00733.1
- Nygård T, Valkonen T, Vihma T. 2013. Antarctic Low-Tropospheric Humidity Inversions: 10-yr Climatology. J. Clim., 26, 5205-5219, DOI: 10.1175/JCLI-D-12-00446.1
 - Onogi K, Tsutsui J, Koide H, Sakamoto M, Kobayashi S, Hatsushika H, Matsumoto T, Yamazaki N, Kamahori H, Takahashi K, Kadokura S, Wada K, Kato K, Oyama R, Ose T, Mannoji N, Taira R. 2007. The JRA-25 Reanalysis. J. Meteor. Soc. Japan, 85: 3, 369-432, DOI: 10.2151/jmsj.85.369
 - Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP, Kent EC, Kaplan A. 2003. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. J. Geophys. Res. Atmos., 108: D14, DOI: 10.1029/2002JD002670

Reynolds JM. 1981. The Distribution of Mean Annual Temperatures in the Antarctic Peninsula. Br. Antarct. Surv. Bull., 54, 123–133

- Reynolds RW, Rayner NA, Smith TM, Stokes DC, Wang W. 2002. An improved In Situ and Satellite SST Analysis for Climate. *J. Clim.*, 15, 1609–1625
- Rienecker MM, Suarez MJ, Gelaro R, Todling R, Bacmeister J, Liu E, Bosilovich MG, Schubert SD, Takacs L, Kim G-K, Bloom S, Chen J, Collins D, Conaty A, da Silva A, Gu W, Joiner J, Koster RD, Lucchesi R, Molod A, Owens T, Pawson S, Pegion P, Redder CR, Reichle R, Robertson FR, Ruddick AG, Sienkiewicz M, Woollen J. 2011. MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Clim.*, 24: 14, 3624-3648, DOI: 10.1175/JCLI-D-11-00015.1 Saha S, Moorthi S, Pan H-L, Wu X, Wang J, Nadiga S, Tripp P, Kistler R, Woollen J, Behringer D, Liu
- Saha S, Moorthi S, Pan H-L, Wu X, Wang J, Nadiga S, Tripp P, Kistler R, Woollen J, Behringer D, Liu H, Stokes D, Grumbine R, Gayno G, Wang J, Hou Y-T, Chuang H-Y, Juang H-MH, Sela J, Iredell M, Treadon R, Kleist D, Van Delst P, Keyser D, Derber J, Ek M, Meng J, Wei H, Yang R, Lord S, Van Den Dool H, Kumar A, Wang W, Long C, Chelliah M, Xue Y, Huang B, Schemm J-K, Ebisuzaki W, Lin R, Xie P, Chen M, Zhou S, Higgins W, Zou C-Z, Liu Q, Chen Y, Han Y, Cucurull L, Reynolds RW, Rutledge G, Goldberg M. 2010. The NCEP Climate Forecast System Reanalysis. *Bull. Am. Meteorol. Soc.*, 91: 8, 1015-1057, DOI: 10.1175/2010BAMS3001.1
 - Screen JA, Simmonds I. 2012. Half-century air temperature change above Antarctica: Observed trends and spatial reconstructions. J. Geophys. Res. Atmos., 117: D16, D16108, DOI: 10.1029/2012JD017885
 - Skamarock WC. 2004. Evaluating Mesoscale NWP Models Using Kinetic Energy Spectra. *Mon. Weather Rev.*, 132: 12, 3019-3032, DOI: 10.1175/MWR2830.1
 - Stark JD, Donlon CJ, Martin MJ, McCulloch ME. 2007. OSTIA: An operational, high resolution, real time, global sea surface temperature analysis system. in *Oceans '07 IEEE Conference, 'Marine Challenges: Coastline to Deep Sea' 18–21 June 2007*, edited, Aberdeen, UK

Steeneveld GJ. 2014. Current challenges in understanding and forecasting stable boundary layers over land and ice. *Frontiers in Environmental Science*, DOI: 10.3389/fenvs.2014.00041

Stössel A, Zhang Z, Vihma T. 2011. The effect of alternative real-time wind forcing on Southern Ocean sea ice simulations. *J. Geophys. Res. Oceans*, 116: C11, C11021, DOI: 10.1029/2011JC007328

- Tastula E-M, Vihma T, Andreas EL, Galperin B. 2013. Validation of the diurnal cycles in atmospheric reanalyses over Antarctic sea ice. J. Geophys. Res. Atmos., 118: 10, 4194-4204, DOI: 10.1002/jgrd.50336
- 10.1002/jgrd.50336 Taylor KE, Williamson D, Zwiers F. 1997. AMIP Sea Surface Temperature and Sea Ice Concentration Boundary Conditions. <u>http://www-pcmdi.llnl.gov/projects/amip/AMIP2EXPDSN/BCS/index.php:</u> revised 22 June 2015
 - Tetzlaff A, Lüpkes C, Hartmann J. 2015. Aircraft-based observations of atmospheric boundary layer modification over arctic leads. *Q. J. R. Meteorol. Soc.*, DOI: 10.1002/qj.2568
 - Tietäväinen H, Vihma T. 2008. Atmospheric moisture budget over Antarctica and the Southern Ocean based on the ERA-40 reanalysis. *Int. J. Climatol.*, 28: 15, 1977-1995, DOI: 10.1002/joc.1684
 - Turner J, Colwell SR, Marshall GJ, Lachlan-Cope TA, Carleton AM, Jones PD, Lagun V, Reid PA, Iagovkina S. 2005. Antarctic climate change during the last 50 years. *Int. J. Climatol.*, 25: 3, 279-294, DOI: 10.1002/joc.1130
 - Turner J, Lachlan-Cope TA, Colwell S, Marshall GJ, Connolley WM. 2006. Significant Warming of the Antarctic Winter Troposphere. *Science*, 311: 5769, 1914-1917, DOI: 10.1126/science.1121652
 - Vancoppenolle M, Timmermann R, Ackley SF, Fichefet T, Goosse H, Heil P, Leonard KC, Lieser J, Nicolaus M, Papakyriakou T, Tison J-L. 2011. Assessment of radiation forcing data sets for largescale sea ice models in the Southern Ocean. *Deep Sea Res. II Top. Stud. Oceanogr.*, 58: 9–10, 1237-1249, DOI: 10.1016/j.dsr2.2010.10.039
 - Wang J, Hock T, Cohn SA, Martin C, Potts N, Reale T, Sun B, Tilley F. 2013. Unprecedented upper-air dropsonde observations over Antarctica from the 2010 Concordiasi Experiment: Validation of

satellite-retrieved temperature profiles. *Geophys. Res. Lett.*, 40: 6, 1231-1236, DOI: 10.1002/grl.50246

Weiss AI, King J, Lachlan-Cope T, Ladkin R. 2011. On the effective aerodynamic and scalar roughness length of Weddell Sea ice. *J. Geophys. Res. Atmos.*, 116: D19, DOI: 10.1029/2011JD015949
Välisuo I, Vihma T, King JC. 2014. Surface energy budget on Larsen and Wilkins ice shelves in the Antarctic Peninsula: results based on reanalyses in 1989–2010. *The Cryosphere*, 8: 4, 1519-1538, DOI: 10.5194/tc-8-1519-2014

Tables

Table 1. Characteristics of reanalyses

Reanalysis	Model resolution	Dataset resolution	Data assimilation	Treatment of sea ice Sea ice concentration (SIC) based on OSTIA analysis (Stark et al., 2007)		
ERA- INTERIM	T255, L60	0.75 x 0.75	4D-VAR			
JCDAS	T106, L40	1.25 x 1.25	3D-VAR	SIC based on passive microwave (PM) data (Cavalieri et al., 1984, Matsumoto et al., 2006)		
JRA-55	T319, L60	1.25 x 1.25	4D-VAR	SIC based on COBE SST – analysis (Ishii et al., 2005)		
MERRA	2/3 lon x 1/2 lat deg, L72	2/3 lon x 1/2 lat deg	3D-VAR	SIC linearly interpolated in time from weekly Reynolds fields (Reynolds et al., 2002)		
NCEP- CFSR	T382, L64	0.5 x 0.5	3D-VAR	SIC analysis based on PM data. An interactive prognostic model for sea ice thermodynamics and dynamics (Saha et al., 2010)		
NCEP- DOE	T62, L28	2.5 x 2.5	3D-VAR	SIC based on PM data (Cavalieri et al., 1984) with corrections (Taylor et al., 1997)		
NCEP/ NCAR	T62, L28	2.5 x 2.5	3D-VAR	SIC based on PM data (Cavalieri et al., 1984); values exceeding 55% set as 100%		
20CR	T62, L28	2 x 2	Ensemble Kalman Filter	Monthly SIC from the Hadley Center data set (Rayner et al., 2003)		

Table 2. Vertically and spatially averaged bias and RMSE of temperature (T), specific humidity (q) and wind speed (U) for the eastern side (east) and western side (west) of the peninsula and the whole peninsula region (total). The smallest biases and RMSEs are marked as bold .

			ERA-				NCEP-	NCEP-	NCEP/	
			INTERIM	JCDAS	JRA-55	MERRA	CFSR	DOE	NCAR	20CR
Τ(Κ)	bias	west	-1.35	-1.61	-1.78	-0.73	-0.88	-0.57	-0.49	-1.58
		east	-0.50	-0.56	-0.98	-0.13	-0.18	0.43	1.15	0.50
		total	-0.92	-1.09	-1.38	-0.43	-0.53	-0.07	0.33	-0.54
	RMSE	west	1.56	1.82	1.92	1.29	1.28	1.54	1.55	2.05
		east	2.00	2.42	1.89	1.59	1.63	2.47	2.66	3.21
		total	1.78	2.12	1.91	1.44	1.46	2.00	2.11	2.63
q (g kg ⁻¹)	bias	west	0.12	0.01	0.01	0.08	0.22	0.46	0.42	0.01
		east	0.13	0.08	-0.08	-0.14	0.12	0.23	0.33	0.26
		total	0.13	0.05	-0.03	-0.03	0.17	0.34	0.38	0.14
	RMSE	west	0.46	0.50	0.45	0.46	0.50	0.62	0.64	0.68
		east	0.35	0.39	0.32	0.34	0.28	0.59	0.61	0.56
		total	0.40	0.44	0.38	0.40	0.39	0.61	0.62	0.62
U (m s ⁻¹)	bias	west	-0.96	-0.67	-0.75	-0.96	-0.75	-0.89	-0.88	-1.68
		east	-1.52	-2.02	-1.59	-1.26	-1.38	-1.42	-2.16	-2.52
		total	-1.24	-1.34	-1.17	-1.11	-1.07	-1.16	-1.52	-2.10
	RMSE	west	2.85	3.14	2.78	2.90	2.81	3.07	3.08	3.68
		east	2.96	3.32	2.80	3.58	2.60	3.42	3.40	3.78
		total	2.91	3.23	2.79	3.24	2.70	3.25	3.24	3.73

List of Figures

Figure 1. Comparisons of wind (arrows) along the flight path and vertical wind speed profile measured by BAS Twin Otter and AWI Polar 5 –aircraft.

Figure 2. Locations of vertical profiles measured by BAS Twin Otter and AWI Polar 5 –aircraft. The study region was divided into four sub-regions, marked with dashed lines.

Figure 3. Profiles of potential temperature (top), temperature bias (middle), and temperature RMSE (bottom) in the four sub-regions.

Figure 4. Standard deviations of temperature (T), specific humidity (q) and total wind speed (U) based on all the data.

Figure 5. Profiles of specific humidity (top), specific humidity bias (middle), and specific humidity RMSE (bottom) in the four sub-regions.

Figure 6. Profiles of total wind speed (top), wind speed bias (middle), and wind speed RMSE (bottom) in the four sub-regions.

Figure 7. Surface geopotential height for the Antarctic Peninsula (top), mean u-component of wind at 925 hPa (middle), and mean v-component of wind at 925 hPa (bottom) of ERA-INTERIM, JCDAS and NCEP/NCAR. The mean values have been calculated for the measurement campaign period in 2010.





Figure 1. Comparisons of wind (arrows) along the flight path and vertical wind speed profile measured by BAS Twin Otter and AWI Polar 5 –aircraft.



Figure 2. Locations of vertical profiles measured by BAS Twin Otter and AWI Polar 5 –aircraft. The study region was divided into four sub-regions, marked with dashed lines.



Figure 3. Profiles of potential temperature (top), temperature bias (middle), and temperature RMSE (bottom) in the four sub-regions.



Figure 4. Standard deviations of temperature (T), specific humidity (q) and total wind speed (U) based on all the data.



Figure 5. Profiles of specific humidity (top), specific humidity bias (middle), and specific humidity RMSE (bottom) in the four sub-regions.



Figure 6. Profiles of total wind speed (top), wind speed bias (middle), and wind speed RMSE (bottom) in the four sub-regions.



TT1C

1600

1400

1200

1000

800

600

400

200

-2

m s

-6

m s^{−8}

m

Figure 7. Surface geopotential height for the Antarctic Peninsula (top), mean u-component of wind at 925 hPa (middle), and mean v-component of wind at 925 hPa (bottom) of ERA-INTERIM, JCDAS and NCEP/NCAR. The mean values have been calculated for the measurement campaign period in 2010.