



The Arctic summer atmosphere: an evaluation of reanalyses using ASCOS data

C. Wesslén^{1,2}, M. Tjernström^{1,2}, D. H. Bromwich^{3,4}, G. de Boer^{5,6}, A. M. L. Ekman^{1,2}, L.-S. Bai⁴, and S.-H. Wang⁴

¹Department of Meteorology, Stockholm University, Stockholm, Sweden

²Bert Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden

³Atmospheric Sciences Program, Department of Geography, The Ohio State University, Columbus, Ohio, USA

⁴Polar Meteorology Group, Byrd Polar Research Center, The Ohio State University, Columbus, Ohio, USA

⁵Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado, USA

⁶NOAA/Earth System Research Laboratory, Boulder, Colorado, USA

Correspondence to: C. Wesslén (cecilia@misu.su.se)

Received: 16 May 2013 – Published in Atmos. Chem. Phys. Discuss.: 20 June 2013

Revised: 11 December 2013 – Accepted: 28 January 2014 – Published: 14 March 2014

Abstract. The Arctic has experienced large climate changes over recent decades, the largest for any region on Earth. To understand the underlying reasons for this climate sensitivity, reanalysis is an invaluable tool. The Arctic System Reanalysis (ASR) is a regional reanalysis, forced by ERA-Interim at the lateral boundaries and incorporating model physics adapted to Arctic conditions, developed to serve as a state-of-the-art, high-resolution synthesis tool for assessing Arctic climate variability and monitoring Arctic climate change.

We use data from Arctic Summer Cloud-Ocean Study (ASCOS) to evaluate the performance of ASR and ERA-Interim for the Arctic Ocean. The ASCOS field experiment was deployed on the Swedish icebreaker *Oden* north of 87° N in the Atlantic sector of the Arctic during August and early September 2008. Data were collected during the transits from and to Longyearbyen and the 3-week ice drift with *Oden* moored to a drifting multiyear ice floe. These data are independent and detailed enough to evaluate process descriptions.

The reanalyses captures basic meteorological variations coupled to the synoptic-scale systems, but have difficulties in estimating clouds and atmospheric moisture. While ERA-Interim has a systematic warm bias in the lowest troposphere, ASR has a cold bias of about the same magnitude on average. The results also indicate that more sophisticated descriptions of cloud microphysics in ASR did not significantly improve the modeling of cloud properties compared to ERA-Interim. This has consequences for the radiation balance, and hence

the surface temperature, and illustrate how a modeling problem in one aspect of the atmosphere, here the clouds, feeds back to other parameters, especially near the surface and in the boundary layer.

1 Introduction

The Arctic climate has changed substantially over the recent decades, more than anywhere else on Earth (ACIA, 2005; IPCC, 2007). Arctic warming has been more than twice as large as the global average (Serreze et al., 2009; Richter-Menge and Jeffries, 2011). The impacts are pan-Arctic and manifold, including changes in sea-ice thickness and extent, permafrost, vegetation and ecosystems. These changes, in turn, affect life and nature in both the Arctic and possibly at lower latitudes (e.g., Murray et al., 2010). Unfortunately, the processes responsible for this enhanced temperature increase, sometimes referred to as Arctic amplification, as well as associated feedback mechanisms and Arctic climate sensitivity are poorly understood (ACIA, 2005; Tjernström et al., 2012). Consequently, Arctic climate projections generally feature larger uncertainty than those for other regions (Holland and Bitz, 2003; Karlsson and Svensson, 2013; Liu et al., 2013). Global climate models have been shown to struggle with the simulation of present conditions in the Arctic (Walsh et al., 2002; Chapman and Walsh 2007; Karlsson and Svensson 2011; Svensson and Karlsson 2011; de Boer et

al., 2012). While many hypotheses exist regarding the reason behind these problems, it is commonly believed that a lack of Arctic observations inhibits a more thorough evaluation and improvement of model parameterizations. Specific examples of missing observations include limited spatial coverage of surface and upper air measurements. Additionally, there are only limited observations available that provide sufficient detail to advance our understanding of small-scale processes. A consequence of this is that parameterizations of sub-grid-scale processes in climate models generally stem from empirical evidence obtained in lower-latitude regions, where the climate system may work very differently than in the Arctic.

Much of the present progress in understanding a changing Arctic climate relies on global reanalyses. These reanalyses are dependent on the accuracy of an underlying atmospheric model, which generally provides a spatial resolution that is too coarse to directly study numerous aspects of the climate system. Additionally, the reanalyses rely on the availability of observations with high temporal and spatial resolution. The lack of these types of measurements in Arctic locations impacts reanalyses negatively in three ways: (i) fewer data are available for data assimilation to constrain the models; (ii) all available data are incorporated through data assimilation in the reanalyses, leaving no or very few independent data for evaluation; and (iii) parameterizations contained in the underlying global models are not necessarily capable of simulating processes unique to the Arctic environment. Global reanalyses also differ in important technical aspects, for example in data assimilation techniques, or the descriptions of boundary layer, clouds and other sub-grid-scale processes.

The Arctic System Reanalysis (ASR) is a regional reanalysis developed to serve as a state-of-the-art synthesis tool for assessing Arctic climate variability and monitoring Arctic climate change. The ASR is designed to provide a complete high-resolution (15 km) atmospheric and land-surface data set; however, the development versions used in this study have a resolution of 30 km. As the ASR is a regional product, forcing from the ERA-Interim global reanalysis is used at the lateral boundaries (Fig. 1). In addition, ASR incorporates model physics adapted to Arctic conditions (Bromwich et al., 2010), especially dealing with descriptions of Arctic sea ice and Arctic land.

To date, several evaluations of the forecast performance of the atmospheric model (Polar WRF) used as an underlying model to produce ASR have been performed utilizing both routine observations and data from Arctic field campaigns (Hines and Bromwich, 2008; Bromwich et al., 2009; Hines et al., 2011; Wilson et al., 2011; Wilson et al., 2012). However, the present study is the first evaluation of ASR over the Arctic Ocean. Two developmental versions of the ASR (30 km) are compared with observations from the Arctic Summer Cloud Ocean Study (ASCOS; Tjernström et al., 2012, 2013). ASCOS was an icebreaker-based expedition to the central Arctic Ocean, around 87° N in the Atlantic sector of the Arc-

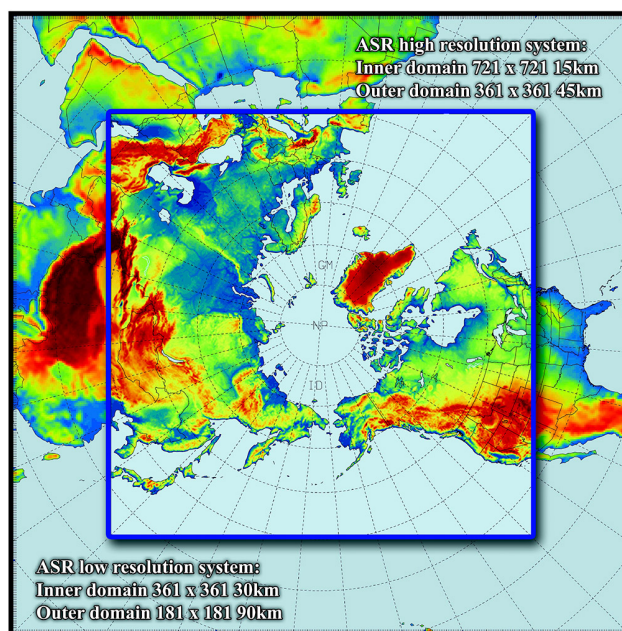


Fig. 1. The inner and outer domains of the Arctic System Reanalysis (ASR), at 30 km and 90 km resolution, respectively (<http://polarmet.osu.edu/ASR/>).

tic (Fig. 2). The expedition took place in summer 2008 and was specifically designed to study processes related to the formation and lifecycle of Arctic low-level clouds. ASCOS provides a source of central Arctic Ocean summer data with two important features: (1) the data are sufficiently detailed so that variables and processes that are usually not evaluated in models can be examined, and (2) most of the data (except surface pressure and 10 m wind) were not assimilated into reanalyses, and therefore provide an essentially independent validation data set.

In addition to the ASR, the ERA-Interim reanalysis (Dee et al., 2011) is included in the evaluation: ERA-Interim is the latest global reanalysis from ECMWF. As mentioned before, in the two versions of ASR analyzed here, the lateral boundaries are forced by ERA-Interim data. Therefore, by including ERA-Interim in the evaluation, it is possible to analyze the added value of a high-resolution regional reanalysis.

Although many meteorological variables are considered in this study, the central areas of interest include low-level clouds, the vertical structure of the troposphere and the surface energy balance, areas where ASCOS provides detailed information. The paper is organized with a discussion of the data and methods used for the comparison in Sect. 2, main results are presented in Sect. 3, a short summary is presented in Sect. 4, and the main conclusions are found in Sect. 5.

Table 1. Summary of the participating reanalyses.

Reanalysis name	Horizontal resolution	Total no. vertical levels	Time step (h)	Surface layer scheme	Micro-physics scheme	PBL scheme	Radiation scheme	Sea ice–SST	Main reference
Arctic System Reanalysis 1 (ASR1)	30 km	71*	3	Mellor–Yamada–Janjic	five-class two-moment; Morrison et al. (2005)	Mellor–Yamada–Janjic	RRTMG	ASR optimized SST and sea ice	Bromwich al. (2010)
Arctic System Reanalysis 2 (ASR2)	30 km	71*	3	Mellor–Yamada–Nakanishi Niino	WRF five-class single moment	Mellor–Yamada–Nakanishi Niino	RRTMG	ASR optimized SST and sea ice	Bromwich al. (2010)
ERA-Interim	T255 (80 km)	60*	3	Louis, MO	three-class two-moment (IFS Cy3Ir2)	Eddy diffusivity mass flux	RRTM, long-wave Fouquart and Bonnel (1980), shortwave	NCEP-RTG	Dee et al. (2011)

* In the database, the results are interpolated to a fixed number of isobaric levels: 35 for AS1, 34 for ASR2 and 37 for ERA-Interim.

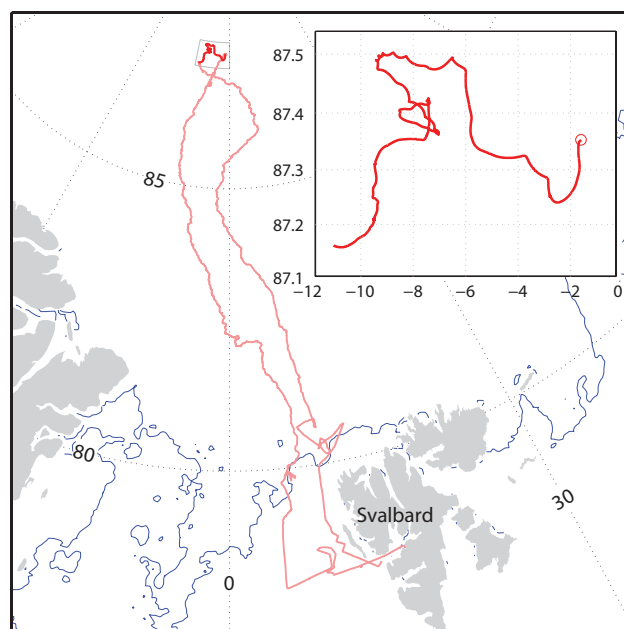


Fig. 2. The ship track from the ASCOS field campaign during 40 days from 2 August and to 8 September 2008. The insert shows the track of the ice drift for three weeks starting 12 August and ending 2 September.

2 Data and methods

2.1 Reanalysis

An atmospheric reanalysis generates a dynamically consistent atmospheric data set by using observational data to constrain successive short-term model forecasts. The observational constraints are applied to the model in a data assimilation cycle, taking into account uncertainties and errors in both the model and observations. The result is a three-dimensional gridded analysis of variables, such as temperature, moisture and winds, as well as the time evolution of these variables. Other variables that are not generally observed can also be extracted from the modeling system, such as fluxes of radiation, turbulent energy and mass fluxes; these variables are very dependent on the model physics. Reanalysis output can therefore be used to study processes that are difficult or even impossible to observe, but underpin the results in the regular variables. Without observations of variables related to these processes, it is difficult to assess the quality of the parameters produced in the reanalyses. However, since they are a result of a dynamically consistent system, this lends some credibility in the representation of the unconstrained variables, although compensating errors certainly remain a problem. Results from reanalyses can also be used to drive other models, such as ice–ocean, land-surface or hydrological models. The quality of a reanalysis is partially dependent on the density and type of observations available to constrain the model.

2.1.1 Arctic System Reanalysis

The regional approach of the ASR results in a smaller model domain, allowing for higher spatial resolution than that used in global reanalyses. The ASR boundary layer and microphysical parameterizations have also been adapted specifically for the Arctic region. As mentioned previously, lateral boundary conditions for the regional model are provided from a global reanalysis (ERA-Interim). As displayed in Fig. 1, the lateral boundaries of the ASR domain are located relatively far south, meaning that ASR benefits from regions with dense observational coverage. It is worth noting that the boundary forcing provided by ERA-Interim only affects the lateral boundaries of the outer ASR domain, and that ASR then applies its own data assimilation of conventional and satellite data inside the regional model domain. Since ASR does not simulate stratospheric processes, it is nudged toward ERA-Interim in the stratosphere throughout the modeling domain.

ASR is developed in collaboration between several institutions led by the Polar Meteorology Group (PMG) of Byrd Polar Research Center (BPRC) at The Ohio State University. ASR is based on a version of the nonhydrostatic Weather Research and Forecasting (WRF) model (Skamarock et al., 2008). As mentioned above, the version of WRF used in the ASR is specially adapted to polar conditions, the so-called “Polar WRF” (e.g., Bromwich et al., 2009). WRF has numerous parameterization options, and in this study we evaluate two developmental versions of ASR with different microphysics and planetary boundary layer schemes. The specifics of these, henceforth called ASR1 and ASR2, are summarized in Table 1.

WRF allows the user to nest multiple domains of various resolutions, and in this study the model was used with an inner and outer domain (Fig. 1). While the target resolution of the final ASR products is 10–15 km, the inner domain applied in this evaluation has a resolution of 30 km, while the resolution of the outer domain is 90 km. The data assimilation (WRF-DA; Barker et al., 2011) was developed at the National Center for Atmospheric Research (NCAR) and uses a cycle of 3 h.

2.1.2 ERA-Interim

ERA-Interim (Dee et al., 2011) is based on a version of the ECMWF Integrated Forecasting System (IFS). Relevant model physics are briefly summarized in Table 1. Four-dimensional variational (4Dvar) data assimilation is applied within IFS. The backbone of the observations includes traditional surface observations and radiosoundings, but has evolved over time to include measurements from an increasing number of satellite sensors. The assimilation of satellite data is of particular importance in the Arctic, where there is a general lack of traditional observations but an abundance of data from polar-orbiting satellites. Because of its global

nature and the amount of computation necessary for 4DVar data assimilation, ERA-Interim has lower horizontal resolution compared to the ASR (Table 1). In addition, since ERA-Interim is global, the ERA-Interim model physics cannot be specifically tailored for the Arctic.

2.2 ASCOS

ASCOS was a Swedish-led field experiment carried out during the International Polar Year (IPY), 2007–2008, onboard the icebreaker *Oden*, from 1 August to 9 September 2008. One main activity of the campaign was a 3-week ice drift with *Oden* moored to an ice floe, drifting with the ice, near 87° N (Fig. 2). The expedition, its targets, instrumentation and observed conditions are described in detail in Tjernström et al. (2012, 2013).

A portion of the data used in this study was collected from ship-borne sensors in operation during ASCOS, including the transit to, through and from the ice, while other observations were obtained using ice-deployed instruments and are therefore only available from the ice drift period. For evaluation of basic meteorological quantities, continuous observations from an automated weather station onboard *Oden* are used in the present study, along with profile data from the 145 6-hourly radiosoundings launched during ASCOS. Note that these radiosoundings were not assimilated into the reanalyses. However, 6-hourly surface pressure and wind observations from *Oden* were assimilated into the ERA-Interim, but not in ASR.

Cloud observations are available for the entire campaign from a combination of an onboard millimeter cloud radar (MMCR) and a laser ceilometer. Estimates of atmospheric liquid water path (LWP) and precipitable water (PWV) were obtained using measurements from a dual-channel microwave radiometer. Ice water path (IWP) was estimated using a combination of different sensors, including the MMCR. Uncertainty in estimates of LWP is roughly 25 g m^{-2} (Westwater et al., 2001), while IWP uncertainty is larger, about a factor of 2 (Shupe et al., 2008).

Observations only available during the ice drift period include all four components of the surface radiation budget, as measured using pairs of pyranometers and pyrgeometers placed above the ice and snow surface, as well as near-ice air temperature, humidity and wind speed, as measured from a mast on the ice. Additionally, turbulent fluxes of sensible and latent heat were estimated through an eddy covariance method, using a combination of sonic anemometers and fast open-path gas analyzers mounted on three masts on the ice. See Tjernström et al. (2013) for details on all instruments.

2.3 Analysis method: interpolation in time and space

The interpolation of the reanalysis data was carried out differently for the data from ASR and ERA-Interim, due to the fact that the data were obtained with a different output

format. Both methods will be described below, starting with the ASR.

The first analysis method is based on a linear interpolation. The ASR data had a spatial resolution of 30 km and a time resolution of 3 h. Using the observations from ASCOS, the reanalysis surface data were interpolated to *Oden's* position in two-dimensional space and time. The ship track from the ASCOS campaign (Fig. 2) was used for the latitude and longitude for interpolation.

When comparing ASR with the ASCOS data from the ice drift, the interpolation was performed using the coordinates obtained from the instruments on the sea ice. For the radiosoundings that were released from the ship, a four-dimensional interpolation of ASR was performed in time and three-dimensional space using the change in longitude and latitude given by the balloon-borne instruments.

The four closest grid points to the observations were used in the linear interpolation. The interpolation is a linear approximation, giving the interpolated values along the ship track a corresponding estimate from the reanalysis. Finally, the results from the spatial interpolation were interpolated in time using the time interval from the ASCOS observations.

When analyzing the vertical data, the reanalysis data were again interpolated to the position of the observations, in this case the radiosoundings from ASCOS. While the observations have a high vertical resolution, the vertical resolution of the ASR was 71 levels. These were interpolated to a fixed number of isobaric levels (before making the comparison with ASCOS), 35 for ASR1 and 34 for ASR2. The vertical resolution of ERA-Interim is 60 levels, but interpolated into 37 isobaric levels (cf. Table 1). These were used for the interpolation from the pressure levels measured by the radiosoundings.

The interpolation method used for the ERA-Interim data was based on evenly distributed points in time obtained from a first interpolation of the reanalysis, along the ASCOS ship track. A result of this is that the location of the measurement point might not coincide completely with the chosen point from ERA-Interim, especially considering the release of the radiosoundings. Thus, for the ERA-Interim data, a temporal interpolation between the points nearest in time to the ASCOS measurements was performed.

2.4 Analysis method: statistics

As the frequency of the observations and reanalysis data differs, and since the reanalysis data strictly are grid-area averages, time averaging of the observed data was performed for all the statistical calculations. Considering the grid size and the average wind speed over an hour, for the observational time series from ASCOS a 1 h running average was chosen for all near-surface variables, and for the radiation and turbulent fluxes. The surface was to 90–100 % covered by sea ice, and therefore the difference will not have significant impact, when averaging over the area; an example of the area is



Fig. 3. During the ASCOS field campaign, the surface was to 90–100 % covered by sea ice. The red circle shows the position of the icebreaker *Oden* during the ice drift.

shown in Fig. 3. For the radiometer estimates, extreme peaks were occasionally noted (cf. Fig. 8a, b), most likely caused by interference from condensate on the instrument window. Therefore, before performing the statistical analysis of these variables, the obviously erroneous peaks were “cut off” at maximum values chosen from the time series (22 kg m^{-2} for PWV, 1 kg m^{-2} for LWP and 0.3 kg m^{-2} for IWP). For these integrated properties (PWV, LWP and IWP), a 3-hourly running average was derived from the high-resolution time series. A different time average interval was chosen for these variables due to the fact that they show higher variability compared to the surface observations.

Objective errors were calculated, defined as the difference between the results from the reanalysis and the observations. These differences are presented both as objective scores and as histograms. The root-mean-square error (RMSE), standard deviation and correlation coefficient between each reanalysis and associated ASCOS observations were also calculated. When interpreting the reanalysis performance, these estimates were used according to the principles described by Hanna et al. (1994). Following their recommendations, a “good result” is taken to be one where there is a small bias (depending on the variable), the standard deviation of the reanalysis is similar to that from the observations and the RMSE is smaller than either of the standard deviations, while the correlation coefficient is high (at least > 0.5). These statistics are also summarized in Taylor diagrams (Taylor, 2001).

3 Results

In this section we present results, starting with near-surface variables in Sect. 3.1 and vertical atmospheric structure in Sect. 3.2. The atmospheric moisture content and the clouds

Table 2. Summary of model errors: mean bias, standard deviation within each data set (SD), root-mean-square error (RMSE) and correlation coefficient between each reanalysis and the ASCOS observations.

	Mean bias			SD				RMSE			Correlation coefficient		
	ASR1	ASR2	ERA-Interim	ASR1	ASR2	ERA-Interim	ASCOS	ASR1	ASR2	ERA-Interim	ASR1	ASR2	ERA-Interim
LWD (W m ⁻²)	-27.1	-43.5	5.7	36.1	32.1	10.9	16.3	41.3	51.7	18.7	0.49	0.49	0.32
LWU (W m ⁻²)	-12.7	-16.5	6.8	14.0	13.9	4.4	8.3	16.5	19.3	11.0	0.57	0.60	0.26
SWD (W m ⁻²)	11.4	35.1	-20.8	36.0	38.9	21.3	25.9	34.2	45.6	43.5	0.51	0.66	0.64
SWU (W m ⁻²)	0.1	14.9	-27.7	26.3	26.2	13.3	18.0	21.7	24.2	38.2	0.55	0.70	0.68
LH (W m ⁻²)	-4.7	-5.9	4.8	9.9	7.3	5.6	1.1	6.6	7.5	7.6	-0.69	-0.69	-0.79
L*QFX (W m ⁻²)	-0.9	-1.6	4.7	7.2	5.1	5.6	1.1	2.2	1.9	7.6	-0.40	-0.41	-0.79
SH (W m ⁻²)	-2.8	-2.9	-0.3	8.3	6.1	2.4	0.9	7.4	6.5	3.4	0.26	0.08	0.18
u* (m s ⁻¹)	0.1	0.07	0.03	0.1	0.1	0.08	0.07	0.1	0.1	0.1	0.84	0.80	0.82
T _{2m} (°C)	-0.8	-1.3	1.3	3.1	3.2	1.9	2.5	2.3	2.5	1.9	0.73	0.76	0.85
RH (%)	-2.8	-3.2	1.0	8.8	8.2	2.5	3.6	8.7	8.4	3.7	0.35	0.35	0.31
U _s (m s ⁻¹)	-1.4	-1.6	-0.4	2.0	2.1	2.7	3.0	2.2	2.3	1.6	0.84	0.86	0.98
P _s (hPa)	-0.6	-0.4	-1.5	10.7	10.7	11.1	13.4	8.0	8.0	8.7	0.80	0.86	0.86
LWP (kg m ⁻²)	-0.08	-0.07	-0.02	0.1	0.04	0.07	0.1	0.2	0.1	0.1	0.13	0.16	0.28
IWP (kg m ⁻²)	-0.05	-0.05	-0.01	0.006	0.007	0.03	0.07	0.09	0.08	0.06	0.24	0.14	0.57
PWV (kg m ⁻²)	-0.8	-2.2	-1.6	2.7	2.5	2.8	3.1	2.8	7.2	5.9	0.44	0.28	0.64

are examined in Sect. 3.3 and surface fluxes are evaluated in Sect. 3.4.

3.1 Near-surface variables

Basic meteorological variables from all reanalyses – i.e., 2 m temperature (T_{2m}), relative humidity (RH), surface pressure (P_s) and scalar wind speed (U_s) – show reasonable agreement with ASCOS observations (Figs. 4, 5). This means that they generally follow the observed temporal variability. To provide a complete view of the statistics and the different representations of the variables, all results are summarized in Table 2 and in a Taylor diagram (Fig. 6).

For T_{2m} (Fig. 4a), ERA-Interim is continuously too warm compared to ASCOS observations (cf. also Table 2) and it fails to represent the two sharp cold periods around DoY¹ 235 and DoY 245. However, ERA-Interim captures the general features of the temperature trends during the observational period, with decreasing temperatures until DoY 245 and a relatively sharp temperature increase at the end as ASCOS transits out of the pack ice. The ASR data sets display similar results and represent the temperature trend at the be-

ginning of the time period (until DoY 237) and at the end (after DoY 245) well. In contrast to ERA-Interim, they also capture the first cold period, around DoY 235, reasonably well. The ASR data sets show, however, poorer agreement (bias -2.7°C) with the observations between DoY 237 and 245. While the observed temperatures rebound to approximately 0°C after the first cold period, and only become slightly colder until DoY 245, the ASR drops again to low values, -5 to -10°C , from \sim DoY 239 and do not rebound until *Oden* is on the transit back out through the pack ice (DoY 247).

The observed relative humidity (RH, Fig. 3b) is consistently very high in the observations, rarely dropping below 90 % and essentially never below 80 % (Tjernström et al., 2012). Several factors contribute to keeping RH high. Moist and warm marine air advected in over the ice from lower latitudes cools down when subjected to the melting ice and snow surfaces. The surface is also quite wet: during the melting season, there is consistently liquid water on the surface. In addition, the absolute moisture across the boundary-layer inversion often increases with height (e.g., Tjernström et al., 2004, 2012), and thus entrainment from above may be an additional source of moisture. Andreas et al. (2002) showed that the near-surface atmosphere over sea ice is always close to saturation, with respect to liquid for near-zero temperatures

¹Throughout the rest of this analysis we will use decimal Day of the Year (DoY) for time reference, defined as DoY=1.0 at 00:00 UTC, 1 January.

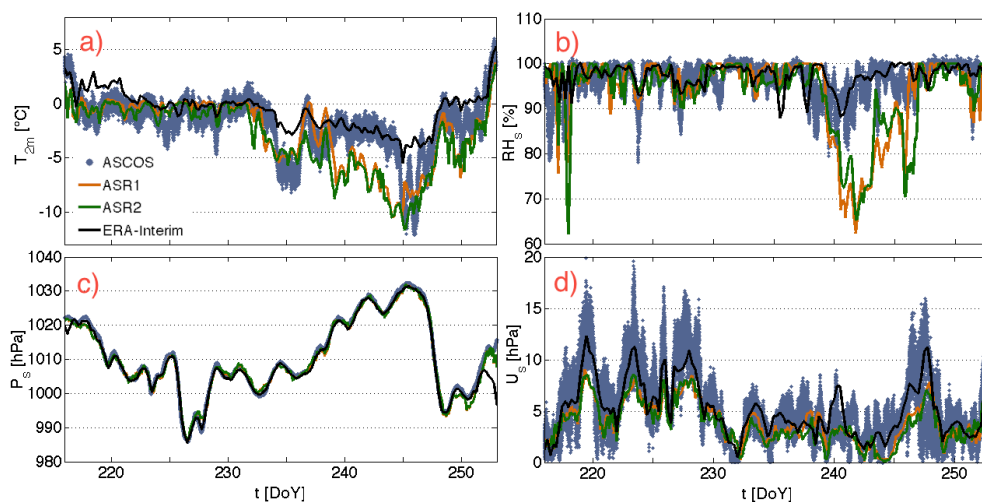


Fig. 4. Time series of near-surface variables for the entire ASCOS expedition showing (a) 2 m temperature in °C, (b) relative humidity (RH) in %, (c) surface pressure in hPa, and (d) scalar wind speed in ms^{-1} , all of which as a function of the time in day of year (DoY) 2008. Each panel shows observations (grey) and interpolated model results from ASR1 (orange), ASR2 (green) and ERA-Interim (back). This color scheme applies throughout all figures.

and for freezing temperatures close to saturation with respect to ice. For most of the time series all three reanalyses also show equally moist conditions as the observations. An exception to this is the ASR during time periods when ASR temperatures are too low compared to observations (DoY ~ 240 to ~ 246).

There is a decrease in RH around DoY 239 according to the observations and all the reanalyses capture the onset of this event. However, the drop in RH is delayed in the ASR (and ERA-Interim), and the magnitude of the drop is also too large. RH then remains too low in ASR until DoY 246, i.e., during the whole time period when the temperature is also underestimated. Although RH is calculated with regard to liquid – and for temperatures well below zero, this RH should drop – the difference in RH at saturation with regard to liquid and ice at these temperatures is only $\sim 10\%$. The ASR bias in RH is much larger and indicates that the model also misrepresents the atmospheric moisture, along with the temperature, during this time period.

Surface pressure (Fig. 4c) is in general well captured (correlation coefficient 0.80–0.86) by all three reanalyses, with two exceptions: one right at the beginning of the measurement period, where all three reanalyses have lower pressure, and one at the very end of the time series, where ERA-Interim is significantly lower than the observations, while ASR remains closer to what was observed.

All reanalysis data sets underestimate scalar wind speed (bias around -1.5 ms^{-1} for ASR and -0.4 ms^{-1} for ERA-Interim) but follow the temporal variability of the observations quite well (Fig. 4d), and ERA-Interim displays higher wind speeds than both versions of the ASR, which are very similar and continuously too low; the absolute difference is larger for higher winds. It should be noted, however, that

the weather station onboard *Oden* was located around 25 m above the surface. Theoretically, the measured winds should be about 10% higher than at 10 m. Comparisons from the ice drift, when wind speed observations were also taken on the ice, indicate that the wind speed from the weather station was about 0.5 ms^{-1} higher than at 8 m. The bias for ERA-Interim is therefore well within this uncertainty, while the ASR biases are still substantial, which could be expected since the surface wind speed observations were assimilated in ERA-Interim.

Figure 5 displays histograms of the differences between each reanalysis and the ASCOS data. The positive bias in ERA-Interim $T_{2\text{m}}$ is clearly visible in the figure, although the variability in the error is smaller than for the two versions of ASR; note that the temperature was *not* assimilated in either reanalysis. The histograms of the error for the ASR has the main maximum at zero, but with long negative tails – a manifestation of the episodes where the ASR temperatures are much too low (cf., e.g., Fig. 4). The histograms of the error in RH are similar for all three reanalyses, with a peak at -1% , which is well within the measurement accuracy for the RH observations. There are more positive errors in ERA-Interim than in ASR, while a long negative tail in ASR again comes from the period with the significant temperature error. All the three reanalysis results have a negative wind speed bias with a similar spread of the error; ERA-Interim is closer to the observations than both the ASR versions, which may be because of the assimilation of observed wind speed as previously discussed. This is even more obvious in the histogram for the surface pressure; all reanalyses have a peak at a roughly -1 hPa bias, but the error distribution is significantly narrower for ERA-Interim than for either ASR.

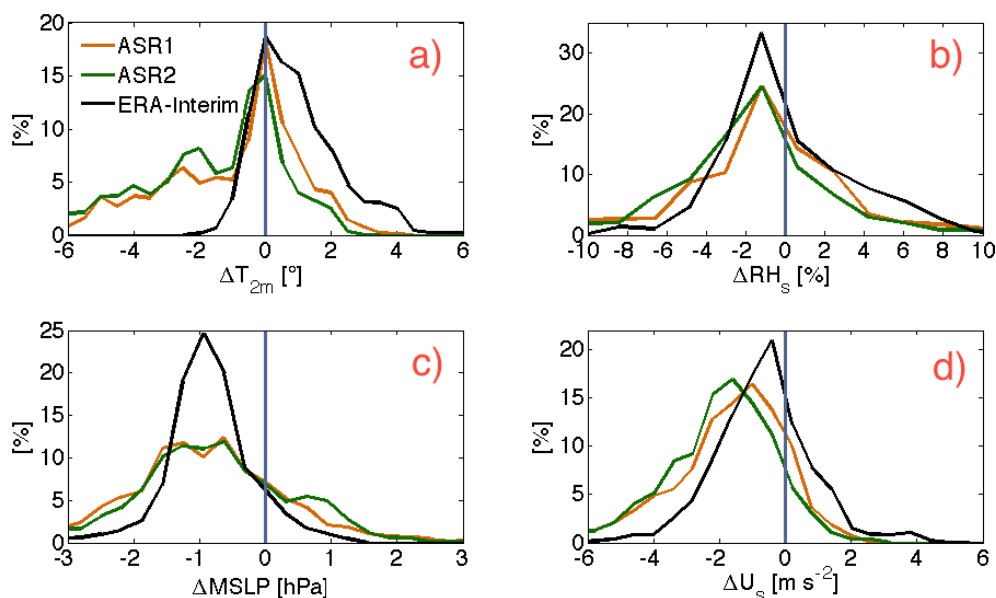


Fig. 5. Histograms (%) of the difference between each reanalysis and the ASCOS observations for the variables in Fig. 4: (a) 2 m temperature in °C, (b) relative humidity (RH) in %, (c) surface pressure in hPa, and (d) scalar wind speed in ms^{-1} .

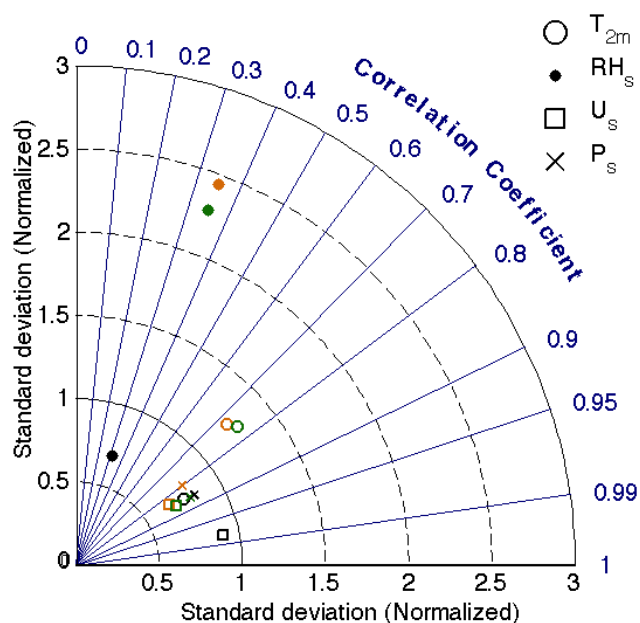


Fig. 6. Taylor diagram summarizing the normalized standard deviation and the correlation coefficient for each of the variables in Fig. 4. Note that the distance to unity at any given point is directly related to the root-mean-square error (RMSE).

Figure 6 summarizes the statistics of the surface meteorological variables, where the majority of the variables display a normalized standard deviation close to unity, except for RH for both ASR1 and ASR2, which displays a larger variability with over a factor of 2. All reanalysis estimates of RH also display a relatively low correlation with the ob-

served RH: correlation coefficients 0.31–0.35. Surface pressures, 2 m temperatures and the scalar wind speeds all have a high correlation coefficient in all reanalyses (0.73–0.98, Table 2). ASR temperatures are too variable, but all other variables have too little variability. The excessive overall variability in ASR RH and $T_{2\text{m}}$ is likely caused by the period when both temperature and RH are too low; otherwise the difference in observed and simulated variability is likely an artifact from the arbitrary choice of averaging time for the observations – a longer running mean would have reduced the observed variability. Hence, Fig. 6 corroborates the conclusion that all surface variables except RH are generally well represented in all reanalyses. The exception is the RH, in particular in ASR.

3.2 Vertical structure

Using the measured temperature (T), relative humidity (RH) and scalar wind (U) from the radiosoundings, it is possible to examine the ability of each reanalysis to capture the vertical structure of the atmosphere. All available radiosoundings (145) taken during the ASCOS field campaign were used in the analysis. It is worth pointing out that wind speed derived from soundings is a vertically integrated property, using the motion of the sonde, and hence that low-level wind speed from soundings is affected by surface wind observations and associated assumptions due to sampling constraints close to the surface. There are also sources of uncertainty in the high-altitude temperature and moisture measurements due to potential radiation errors at higher altitudes and the difficulty in measuring relative humidity at low absolute humidity and temperature.

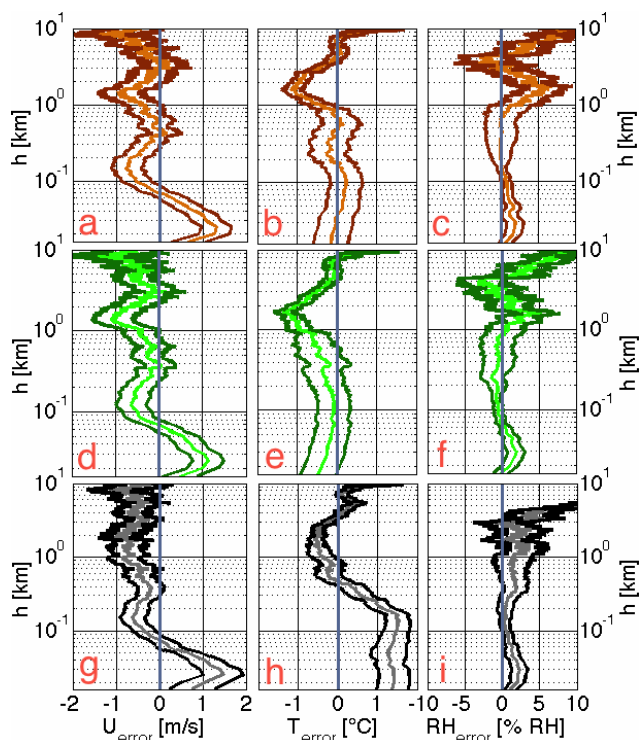


Fig. 7. The mean error of the vertical structure as a function of altitude (km) comparing reanalyses to soundings from (a–c) ASR1, (d–f) ASR2 and (g–i) ERA-Interim. The different panels show (a, d, g) scalar wind speed in m s^{-1} , (b, e, h) temperature in $^{\circ}\text{C}$, and (c, f, i) relative humidity in percent. The lighter-colored middle line is the median difference between reanalysis and observation, while the darker lines are the \pm the 95 % significance interval. Note the logarithmic vertical scale.

The outer, darker lines in Fig. 7 show the 95 % significance interval on either side of the median difference calculated using a double-sided Student t test. This can be interpreted such that if there is a bias but the zero difference line (the null hypothesis) falls within this significance interval, that bias is not significantly different from zero at this level of confidence. It is important to remember that the statistical significance based on a simple Student t test assumes that the differences have a Gaussian distribution; this is not always the case, especially not for RH, where there is a natural threshold at $\text{RH}=100\%$, as well as for the scalar wind speed, which has a lower limit at 0 m s^{-1} . However, a Wilcoxon–Mann–Whitney rank-sum test (assuming no particular statistical distribution of the data) was also performed and no significant differences in the results were seen.

The errors in the vertical wind profiles are similar for all three reanalyses (Fig. 7). In general, the winds close to the surface are higher in the reanalyses than in the observations, which is the opposite compared to the results when using direct anemometer observations, but are too low above 100 m. This is in contrast to results derived using direct anemometer

observations. The differences in the ASR are close to zero in the $\sim 400\text{--}800\text{ m}$ interval and between 2 and 5 km, while the bias in the ERA-Interim is more negative. While ERA-Interim differences are generally statistically significant, differences above 500 m in the ASR evaluation are generally not significant. There are some smaller improvements from ASR1 to ASR2 in the lowest part of the atmosphere (below 100 m). At higher altitudes, ASR1 performs better than ASR2, but the difference between the two ASR versions is in general very small.

For the temperature profiles, the vertical structure is very similar for the two versions of ASR, with a slight indication that the newer version (ASR2) is too cold near the surface; differences below $\sim 1\text{ km}$ are, however, nonsignificant. Both ASR versions are too cold compared to the observations in the 1–5 km interval, and this difference is statistically significant. From $\sim 5\text{ km}$ up to the tropopause the bias is close to zero. The increasing positive bias approaching 10 km, which is seen in all reanalyses, is most likely due to a systematic difference in tropopause height. ERA-Interim has a similar structure to the ASR data sets above 1 km, although it displays a larger bias and is generally too warm in the upper troposphere and below 500 m. The lower-atmospheric bias increases with decreasing altitude and below 200 m ERA-Interim is approximately 1.2°C too warm. This is larger than, but consistent with, the results found for $T_{2\text{m}}$ as discussed above.

The variable displaying the best agreement with the radiosoundings is the relative humidity (RH); that is, the figure with RH shows the narrowest confidence interval compared to both temperatures and winds. All reanalyses are slightly too moist in the lowest 100 m and much too moist above 6–7 km, but the difference is otherwise very close to zero; ERA-Interim, however, has a moister free troposphere compared to ASR, but only errors below 100 m and above 6–7 km are significant. It is worth noting that the specific humidity is quite low at high altitudes, and thus even a large error in RH does not necessarily imply a large error in specific humidity.

3.3 Integrated moisture and clouds

During ASCOS, PWV and LWP were measured using a dual-wave length microwave radiometer, while IWP was estimated using a combination of remote sensors, including the MMCR.

PWV indicates the availability of water for clouds to form, and this variable is tightly linked to weather systems bringing warm and moist air. For PWV (Fig. 8a), agreement between all reanalysis data sets and observations is reasonable in the sense that the reanalyses follow the overall temporal evolution of the observations. The correlation coefficients between the two versions of the ASR and the observations are between 0.28 and 0.44, while for the ERA-Interim it is 0.64. The observed PWV is, however, in general higher than the reanalyses, up to 1 kg m^{-2} . The largest bias seems to appear

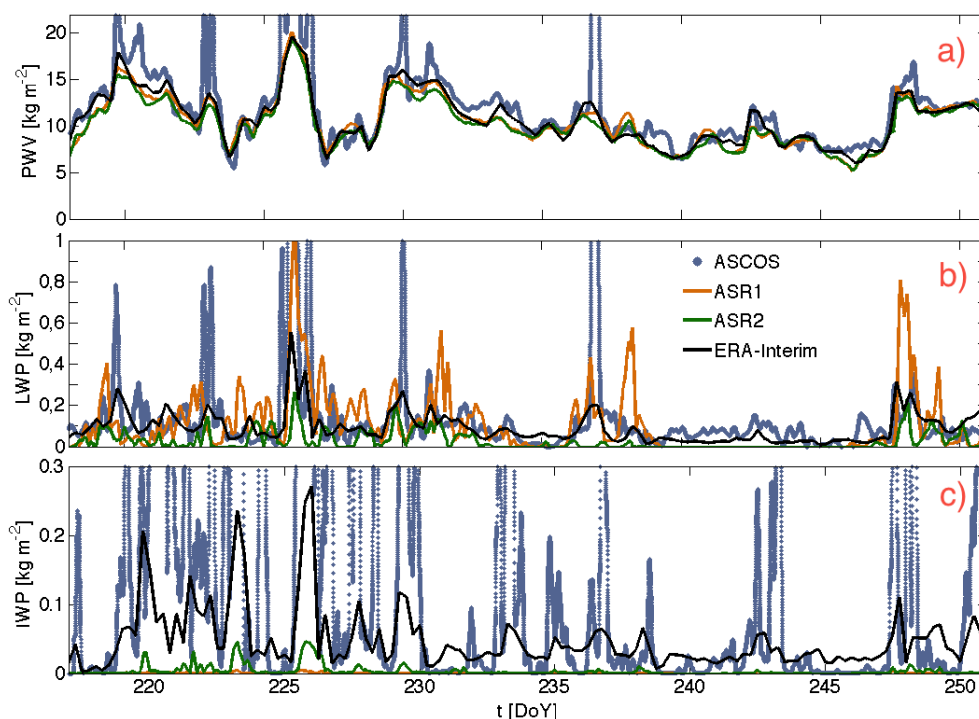


Fig. 8. Time series of integrated water properties showing (a) precipitable water vapor (PWV), (b) liquid water path (LWP) and (c) ice water path (IWP), all in kg m^{-2} , as a function of the time in day of year (DoY) 2008. Note that the observations in (a–b) are from the dual-channel microwave radiometer, while the IWP (c) observations are estimated using the MMCR.

for the northernmost portion of ASCOS, when the ship was located well away from the open ocean (DoY 230–231) and the PWV was generally lower.

The amount of cloud water in the atmosphere is very important for the surface energy balance. Large values in the observed PWV are accompanied by large values in the LWP (Fig. 8b), illustrating the fact that higher PWV and LWP are associated with passing weather systems. The reanalyses captures the timing of several events, even though the magnitude of LWP is not the same as the observations. For LWP, the difference between the reanalyses and the observations is large during the high LWP events. In general, ASR1 shows significantly higher values than the other two reanalysis data sets. These periods of higher values in ASR1 sometimes agree with the observations, such as around DoY 226, but there are also a few occurrences where there is no indication of higher values in the observations, such as around DoY 239.

When only examining the time series in Fig. 8b, there is a difference in LWP between ASR1 and ASR2, where the latter has continuously lower values and is less variable. However, the mean biases for both ASR versions are similar (Table 2) and the correlation coefficients are in the same range (0.13 vs. 0.16, Table 2), compared to the correlation coefficient for ERA-Interim, which is higher (0.28).

A striking difference between both ASR data sets and the observations appears during a weeklong period (DoY 239–

247) when the ASR LWP is virtually zero, whereas the observations indicate values that are low but still significantly larger than zero. Comparing with the analysis of $T_{2\text{m}}$, this period coincides with the ASR cold bias period. Interestingly, although the LWP is too low, ERA-Interim retains a cloud layer during DoY 239–247, despite, or perhaps because of, the simplified parameterization applied in ERA-Interim, where the separation of cloud water between liquid and ice is only a function of temperature. In ASR, liquid cloud water can be transformed to cloud ice, which in turn may precipitate out more easily (e.g., Prenni et al., 2007; Liu et al., 2011).

Ice water path (IWP) is a difficult variable to observe, and the estimates are very noisy (Fig. 8c). In models, IWP is in general badly constrained by observations, for this very reason. The IWP from ASR1 is very low (close to zero) compared to the observations. The IWP from ASR2 is also too low ($<0.05 \text{ kg m}^{-2}$), but the estimate is closer to the observations compared to ASR1. ERA-Interim (Fig. 8c), on the other hand, displays IWP values that are consistently much higher than both the ASR versions. It seems to have a structure where the highest peaks are underestimated but where the values in between the main weather systems are too high, possibly also an effect of the simple cloud phase partitioning. One could question the quality of these comparisons when the observations have such a high uncertainty, but one

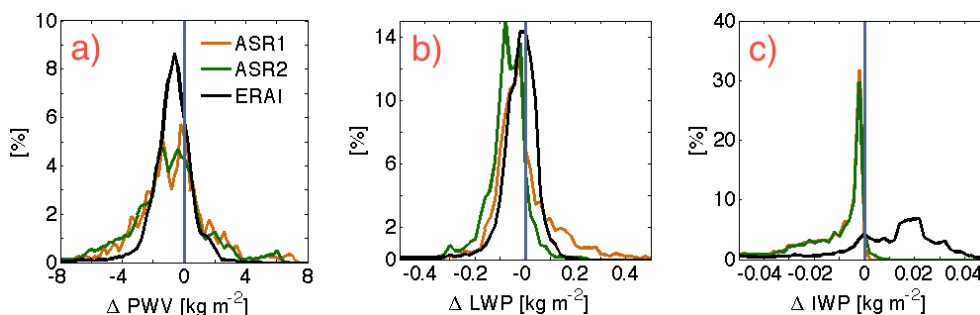


Fig. 9. Histograms (%) of the difference between each reanalysis and the ASCOS observations for the variables in Fig. 8: (a) precipitable water vapor (PWV), (b) liquid water path (LWP) and (c) ice water path (IWP), all in kg m^{-2}

important issue to point out is the large difference between the reanalyses.

The histograms of PWV (Fig. 9a) peak at about the same value in all three reanalyses, and the bias is only slightly negative. This small negative bias is consistent with the time series comparison (Fig. 8a). The histogram for ERA-Interim is narrower than those obtained from both ASR versions, while the ASR histograms feature an unexplained double peak. The corresponding histograms for LWP (Fig. 9b) illustrate that ERA-Interim is closer to the observations, while both ASR versions peak at a lower values; all reanalyses underestimate LWP on average. ASR1 has a long positive tail, which is a result of the few occasions in which the reanalysis has peaks in LWP and the observations do not; otherwise, ASR1 and ASR2 are relatively similar. For IWP (Fig. 9c), the two ASR versions are also very similar, both having peaks close to zero error and relatively narrow distribution except for pronounced negative tails, indicating infrequent occasions with a very large underestimation. ERA-Interim has a large positive bias on average and very rarely displays IWP values lower than in the observations. This is a result from the long periods with low observed IWP, when ERA-Interim has a positive bias despite underestimating the observed high IWP events.

In summary, all reanalyses reproduce the PWV well, with only a slight negative bias. Surprisingly, ERA-Interim, with a relatively simple cloud parameterization, reproduces an LWP that is closest to the observations, indicating a very well-tuned system. Both ASR versions have too low LWP and both versions also fail to produce any clouds during the weeklong stratocumulus period towards the end of ASCOS (DoY 239–247). On the other hand, ERA-Interim has a large positive bias in IWP, while both ASR versions are closer to the observations on average, most of the time within the observational uncertainty. ASR2 displays higher IWP than ASR1, which is compensated by a lower LWP. However, it should be noted that ASR and ERA-Interim behave differently. While ASR shows long periods with no cloud ice (even when this is observed), ERA-Interim always show cloud ice, albeit overestimated for long periods.

Figure 10 displays the standard deviation and the correlation coefficient for the moisture content variables (LWP, IWP and PWV) in all three reanalyses. While all variables display a relatively low normalized standard deviation, below 1 for all reanalyses, the correlation coefficient is relatively high for IWP and PWV in ERA-Interim (0.57–0.64). In both versions of the ASR, the correlation coefficient is low for all three variables (below 0.5, Table 2). Note again that the variability of the observed values are, with the possible exception of PWV, high and that some of this is likely due to instrument problems. While removing obviously erroneous peak values helps to reduce variability in the observations, it does not fully solve the problem; rising and falling values associated with such erroneous peaks will still remain since it is close to impossible to be absolutely certain as to when erroneous peaks occur. This is due to the likely high correlation between properly high values in weather systems and the fact that it is in those systems that precipitation falling on the sensors may cause problems.

The previous analyses in Sect. 3.1 and earlier in Sect. 3.3 indicate that the presence of clouds is crucial for the T_{2m} . This conclusion can be made when comparing the errors in T_{2m} (Fig. 5a) and LWP (Fig. 9b). The characteristics of the clouds are important for the surface energy balance (Sedlar et al., 2011). In general, the reanalyses had more difficulties in simulating the later part of the ASCOS time period. The cloud water content in the ASR was found to be very low (Fig. 8b) during this time period, which means that the clouds in ASR virtually disappear for several days between DoY 240 and 245. Only some thin ice clouds were present, as seen in the IWP (Fig. 8c). ERA-Interim, on the other hand, captures the low stratocumulus clouds seen in the observations to some extent, but the LWP is too low.

In the reanalysis, the vertical cloud boundaries were analyzed using the liquid and ice water contents. This was done by selecting the highest and lowest altitude layer containing cloud water, regardless of phase. The highest cloud-top and the lowest cloud-base heights (CTH and CBH, respectively) in the ASR are compared with the cloud radar for CTH and ceilometer for the CBH. Experiences from the ASCOS field

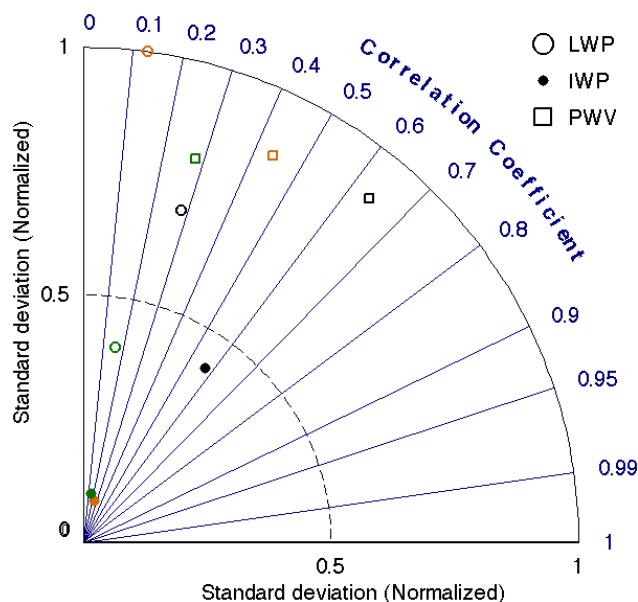


Fig. 10. Taylor diagram summarizing the normalized standard deviation and the correlation coefficient for each of the variables in Fig. 8.

campaign show that the instruments are sensitive enough, since they also recorded cloud particles during events when no clouds were visible to the eye (e.g., Mauritsen et al., 2011). For ERA-Interim, the cloud tops were also separated into the lowest and highest values for cases where more than one cloud layer could be identified; a similar distinction is difficult from observations since when precipitation falls out of an upper cloud layer, the radar will sense the precipitation particles and it will not be possible to distinguish clear layers between clouds. The histograms of the cloud boundaries are found in Fig. 11. Note that this analysis shows the statistical distribution of the clouds whenever they occurred, while the analysis of cloud water above compares the actual occurrences of clouds as a function of time.

Analyzing the height of the cloud base (CBH), ERA-Interim has generally higher CBH than both the ASR versions and the observations (Fig. 11a). Also, distributions of CTH are shown in Fig. 11b. ERA-Interim lowest CTH distribution (solid line) is generally very similar to the observed values, with the majority of the CTH below 1 km. ERA-Interim high CTHs (dotted line) show higher cloud layer that was not present in any of the other data sets. In both versions of the ASR, the CTH is in general too high compared to the observations, with significantly fewer cloud tops below 1 km. Since both versions of the ASR have higher CTH and lower CBH, this indicates that the clouds are generally too thick, which is also seen in Fig. 11c. ASR2, which is the latest version of the ASR, displays even lower cloud bases than ASR1, and hence even thicker clouds. ERA-Interim, on the

other hand, has geometrically too thin clouds, the majority less than 1 km (Fig. 11c).

In summary, while the clouds in both ASR versions had a larger vertical extent and sometimes did not occur at all when they should, they also contained less cloud liquid water compared to the observations. ERA-Interim more frequently provides better estimates of the vertical extent of the cloud layer, and also correctly featured clouds when ASR did not.

3.4 Surface energy fluxes

The observed and reanalyzed energy fluxes for the 3-week period of the ASCOS ice drift are shown in Fig. 12. Both shortwave and longwave radiation were measured separately for the upward and downward fluxes. There are several significant differences between the observations and all reanalyses. The downwelling longwave radiation (LWD, Fig. 12a) is in general underestimated by ASR. Around DoY 233, the LWD in ASR drops significantly, which is in agreement with the observations and was caused by a breakup of low-level clouds (e.g., Tjernström et al., 2012). However, when the observations show increasing values again, ASR1 follows, while ASR2 continues to fluctuate between low and high values. These fluctuations in ASR2 indicate cloudy and clear conditions, respectively. However, clear conditions may consist of thin ice clouds that do not perturb longwave radiation (Cesana et al., 2012). After DoY ~ 239 both versions of ASR mostly stay at lower values, around $220\text{--}240\text{ Wm}^{-2}$ (with a few occasions around 280 Wm^{-2}), whereas the observations stay around 300 Wm^{-2} . ERA-Interim is in general closer to the observations than both ASR versions, except during the time when the observations of LWD drop, around DoY 234–236. It can be recalled from Sect. 3.3 that both ASR versions displayed near-zero values of LWP during this time period (DoY 239–245), whereas ERA-Interim displayed a more realistic LWP behavior compared to the observations. A too low T_{2m} was also noted in ASR1 and ASR2 during the same time period, but not in ERA-Interim. Thus, the error in temperature is most likely due to a failure in ASR to form liquid water clouds during this time period, which results in too low LWD. On the other hand, during the 2-day drop in temperature displayed in Fig. 4 (\sim DoY 236), ERA-Interim fails to dissipate the low clouds, which results in a positive temperature bias, while this event is well captured by both versions of the ASR. This clearly illustrates the importance in accurately modeling the cloud properties.

The upwelling longwave radiation (LWU, Fig. 12b) in ASR is too low compared to the observations except during DoY 234–239 (ASR1) and DoY 234–236 (ASR2). ERA-Interim does not fully capture the temperature drop during DoY 235–237, and hence misses the decrease in LWU. Multiple observations of the T_{2m} and the surface temperatures were conducted during the ASCOS campaign, and the T_{2m} follow the surface temperatures closely, always within 2 standard deviations (Tjernström et al., 2012). The decrease in

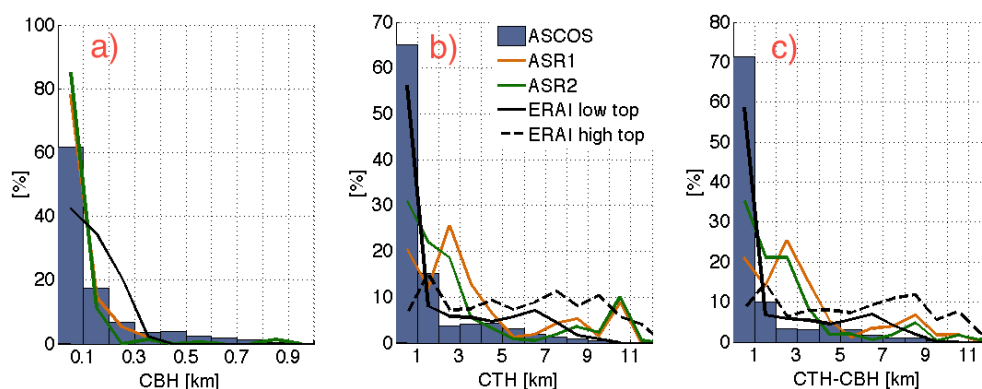


Fig. 11. Histograms (%) of the cloud boundaries showing (a) cloud-base height (CBH), (b) cloud-top height (CTH) and (c) cloud thickness (CTH–CBH) in kilometers. The bars (grey) show results from the ASCOS observations and each of the lines represent results from ASR1 (orange), ASR2 (green) and ERA-Interim (black).

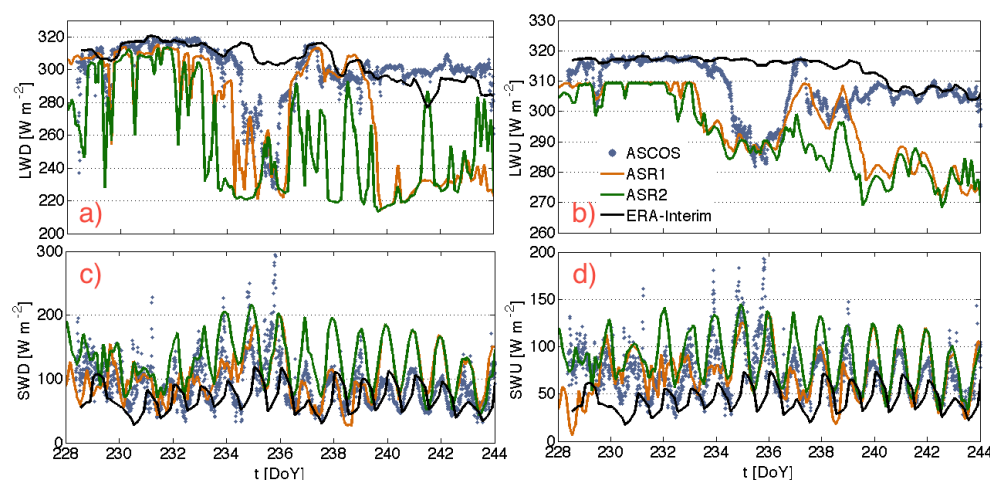


Fig. 12. Time series of the surface radiation components from the ASCOS ice drift showing (a) downward longwave (LWD), (b) upward longwave (LWU), (c) downward shortwave (SWD) and (d) upward shortwave radiation (SWU), all in W m^{-2} , as a function of the time in day of year (DoY) 2008.

LWU provides a very clear example of the important linkages between cloud cover, longwave radiation and near-surface temperature. Another noticeable feature in Fig. 12b is that the ASR LWU also appears to have an upward cap: values above $\sim 310 \text{ W m}^{-2}$ do not occur. This upper bound is likely caused by the upper limit in the surface temperature over melting sea ice in combination with a constant surface emissivity set to $\varepsilon = 0.98$. It is clear that the temperature limit over sea ice in combination with the constant emissivity results in a lower LWU compared to the observations. Observations of surface temperature from ASCOS (not shown; Tjernström et al., 2012) show that the surface temperature rarely reaches above 0°C ; consequently the emissivity must in reality be higher than assumed in ASR. This could possibly be related to the occurrence of a constant weak precipitation and the formation of frost, which keep the top layer of the snow surface relatively fresh. The emissivity in ERA-Interim is also

set to $\varepsilon = 0.98$; however, the temperature is biased slightly high and hence the LWU becomes more realistic.

The diurnal cycle in shortwave radiation is clearly visible in Fig. 12c and d. In general, both versions of ASR (in particular ASR2) display more incoming solar radiation (SWD, Fig. 12c) than the observations. The SWD discrepancy in ASR is particularly clear for the same period as discussed earlier, where ASR is lacking liquid clouds (i.e., DoY 239–247). On the other hand, ASR2 also displays slightly too high SWD when clouds are present, indicating that the clouds are too optically thin in this reanalysis. The upward shortwave radiation (SWU, Fig 12d) displays a similar temporal variability as for SWD, but the values are in general closer to the observations. ERA-Interim seems to have more clouds present, especially for DoY 234–236; hence both SWD and SWU are too low, except for at the end.

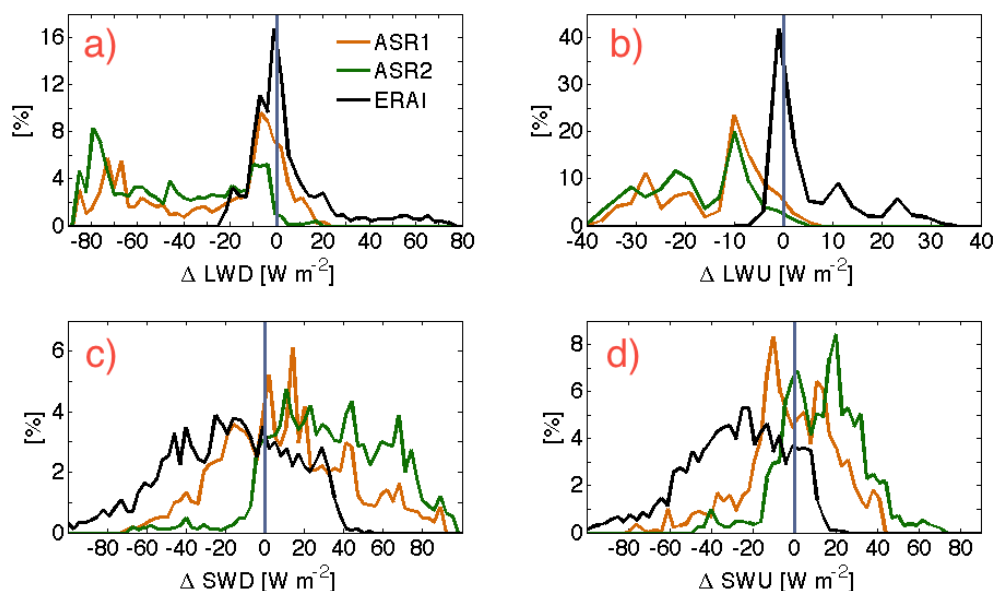


Fig. 13. Histograms (%) of the difference between each reanalysis and the ASCOS observations for the variables in Fig. 12: (a) downward longwave (LWD), (b) upward longwave (LWU), (c) downward shortwave (SWD) and (d) upward shortwave radiation (SWU), all in W m^{-2} .

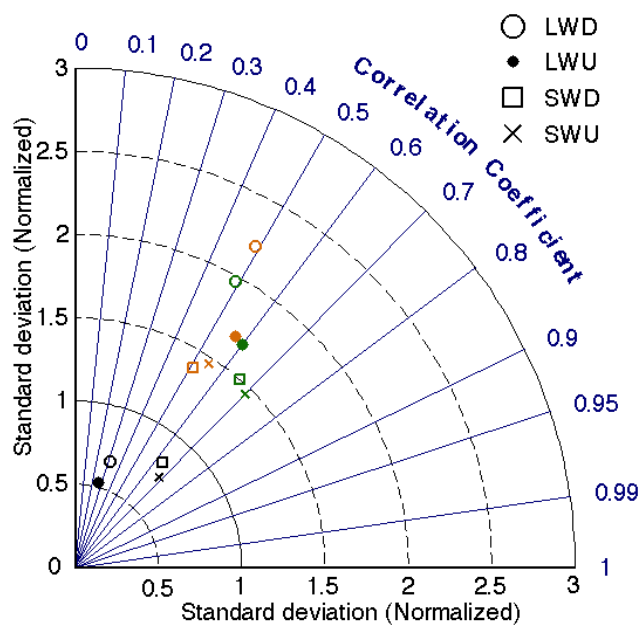


Fig. 14. Taylor diagram summarizing the normalized standard deviation and the correlation coefficient for each of the variables in Fig. 12.

Figure 13 shows the histogram of the reanalysis errors. For both LWD and LWU, the ERA-Interim histograms peak near zero and the spread is reasonable, except for a long positive tail, coming from missing the colder period when the clouds dissipated. For both ASR versions, the histograms of the error are very broad. The LWD error displays two peaks, one around -10 W m^{-2} and another around -70 to -80 W m^{-2} ,

the latter likely due to the anomalous cloudless period. The LWU error also displays a broad histogram, but with only one main peak, around -10 W m^{-2} , a result from the capping due to the constant emissivity.

The shortwave radiation error displays broad histograms: $\pm 100 \text{ W m}^{-2}$. ASR1 is closest to the observations, with an SWD peak that is positive and an SWU peak close to zero. The difference between SWD and SWU must be the surface albedo, and the fact that the error in ASR is reduced in SWU compared to SWD is an indication that the ASR albedo is too low. This is another indication that the observed snow surface might have been “fresher” than assumed in the model. For ERA-Interim the histograms are broad with maximums at negative values. It is worth noticing that the observations were taken in a localized area, and while there was a melt pond within the view of the sensors, the observations cannot realistically depict albedo changes due to melt pond development. The SWU and SWD histograms indicate that ERA-Interim also has a too low surface albedo. It is worth noting here that while ASR also has an implicit melt pond treatment, in ERA-Interim the surface albedo is prescribed using a climatological annual cycle.

For the surface radiation fluxes, summarized in the Taylor diagram in Fig. 14, all variables from ERA-Interim are gathered around a normalized standard deviation between 0.5 and 1.0 and have a low correlation coefficient, around 0.3, for longwave radiation. Hence, the temporal agreement with observations is poor and the variability is underestimated, even though the average errors (e.g., as borne out by the histograms) are reasonable. All radiative flux variables from the ASR display a higher correlation coefficient than ERA-Interim, ~ 0.5 – 0.7 , except SWD and SWU in ASR1. In

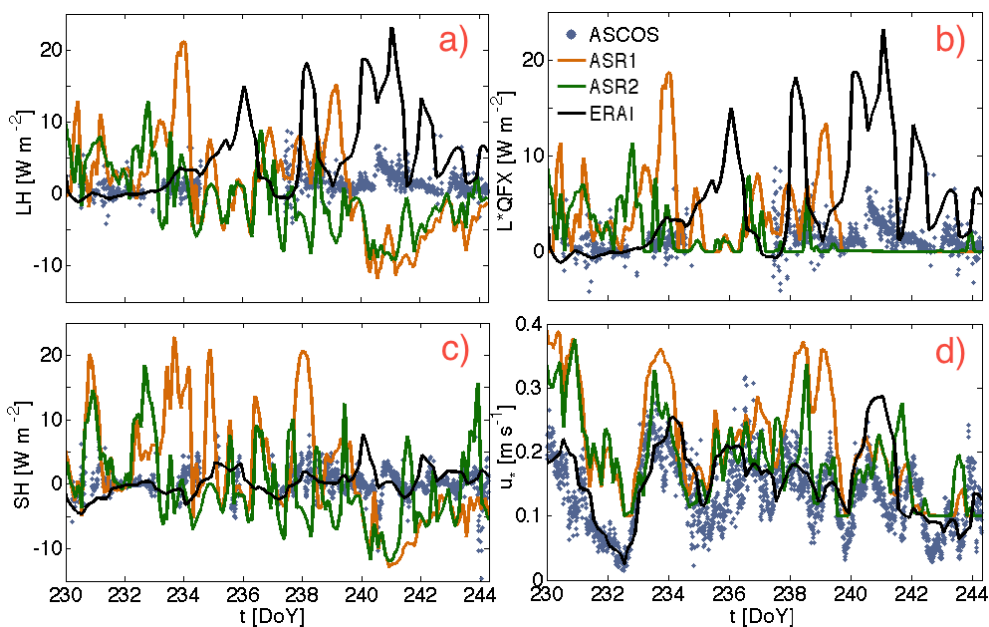


Fig. 15. Time series of the turbulent surface fluxes from the ASCOS ice drift showing (a) latent heat flux at the surface (LH); (b) moisture flux times the latent heat of vaporization at 0 °C (L^*QFX); (c) sensible heat flux (SH) in $W m^{-2}$; and (d) the friction velocity, u_* , in ms^{-1} . The turbulent fluxes are defined positive upwards.

addition, the normalized standard deviation of the shortwave radiation (SWU, SWD) derived from ASR displays a value closer to 1, indicating that the shortwave fluxes are relatively well represented in ASR.

The reanalyses exaggerate the fluctuations in the turbulent fluxes compared to observations (Fig. 15), both negatively and positively; this is consistent with the findings from Tjernström et al. (2005). The two top panels in Fig. 15 show the traditional latent heat flux at the surface (LH) and the upward moisture flux at the surface multiplied with the latent heat of vaporization at 0 °C (QFX), respectively. These two quantities should be exactly the same, but QFX is derived differently in ASR compared to ERA-Interim. ASR excludes negative values, which are set to zero. From the observations it is clear that the latent heat occasionally can be negative, but only rarely and with very small values. While this fix in the ASR solves the problem of the frequently appearing, rather large downward moisture flux in the model by simply canceling it, this also means that the moisture budget in ASR is violated. In both versions of the ASR, there are some indications of a diurnal cycle in the sensible heat flux (SH, Fig. 15c), which is seen neither in the observations nor in the ERA-Interim (Fig. 15c). This feature could be an artifact from the too large values of SWD in ASR. During this time of the year, all available energy (SW) is used for melting the surface; hence the temperature does not change as much over the day. While the SH and LH is negative during the cloudless period in ASR (DoY 239–244), the L^*QFX is set to be positive and will therefore influence the development of the surface processes incorrectly.

During ASCOS, observations of the turbulent momentum flux were made for air flow both over ice and open water and the result did not differ significantly (not shown). The temporal variability of the friction velocity (u_* , Fig. 15d) from ERA-Interim agrees relatively well (correlation coefficient=0.82) with the observed values, although the values are biased high, especially during certain periods, e.g., DoY ~ 241 . The momentum fluxes from the different ASR versions are also overestimated, more than in ERA-Interim, although reduced in ASR2 compared to the ASR1. The calculated correlations coefficients for the entire time period between the observations and the different reanalyses are similar between ASR and ERA-Interim (Table 2).

In the statistical analysis (Fig. 16) the histograms are noisy, most likely due in part to the limited length of the analyzed time period for the turbulent fluxes. For LH, the ASR errors are mostly negative (Fig. 16a); the error is also bimodal with a peak at large negative values, especially in ASR2, corresponding to the too cold period when the flux is erroneously downward (cf. Fig. 15a). The errors are closer to zero when the negative values in ASR are replaced by zeros (Fig. 16b). In other words, the lower bound of QFX in ASR improves the comparison with observations, but for the wrong reason. The ERA-Interim LH flux error has very few negative values but more large values, which gives rise to a long positive tail; the error distributions in all three reanalyses are similar when the negative values in ASR are removed and while peaking at zero is often positive, especially in ERA-Interim.

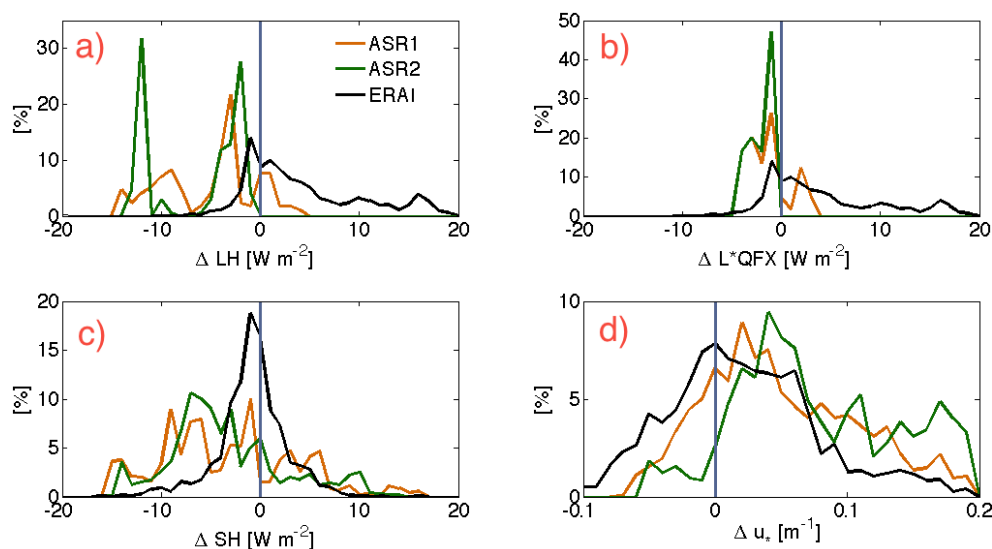


Fig. 16. Histograms (%) of the difference between each reanalysis respectively and the ASCOS observations for the variables in Fig. 15: **(a)** latent heat flux at the surface (LH); **(b)** moisture flux times the latent heat of vaporization at 0 °C (L^*QFX); **(c)** sensible heat flux (SH) in Wm^{-2} ; and **(d)** the friction velocity, u_* , in ms^{-1} .

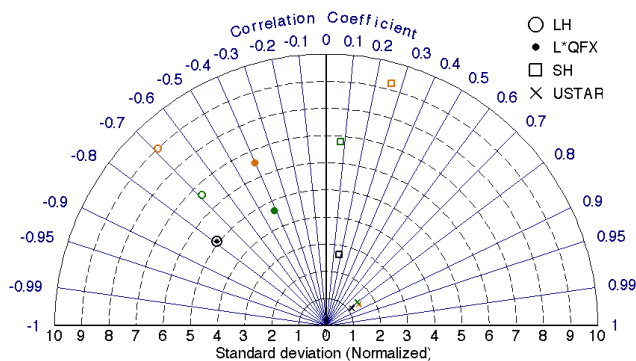


Fig. 17. Taylor diagrams summarizing the normalized standard deviation and the correlation coefficient for each of the variables in Fig. 15.

The ASR SH is too low compared to observations (Fig. 16c). A closer investigation shows that the underestimate is mostly because the ASR atmosphere is too stable, intermittently before DoY 240 and more consistently later. This happens during the erroneously clear period when the temperatures are too low. For the momentum flux (Fig. 16d), the peak for all reanalyses is broad, with the absolute peak centered around zero for ERA-Interim, while both versions of the ASR mostly have a too high friction velocity.

While the statistics for the variables displayed in previous Taylor diagrams (Figs. 6, 10, 14; cf. also Table 2) all showed positive correlation coefficients, although sometimes with a low correlation, the results from the turbulent fluxes give rise to negative values of the correlation coefficient among the majority of the variables (Fig. 17), although the magnitude is

small. Only the momentum flux agrees relatively well with observations, both in terms of correlation coefficient (around 0.8) and normalized standard deviation (close to 1). For the other heat turbulent fluxes, the correlation coefficient is close to zero, or even below zero, indicating that the presentation of the heat and moisture fluxes in all three reanalyses is sometimes completely disconnected from the variability seen in the observations. The agreement is particularly poor for the latent heat flux. The normalized standard deviations are well above unity, which corroborates the conclusion that the turbulent moisture and heat fluxes are poorly represented in all three reanalyses. This is consistent with a similar analysis of modeled turbulent heat fluxes in several regional models using SHEBA data (e.g., Tjernström et al., 2005).

4 Summary and discussion

The overall synoptic meteorological conditions (T_{2m} , P_s and U_s) are generally well represented by all three reanalyses. However, the reanalyses differ from observations in terms of their representation of small-scale temporal variability and the humidity parameters, which in turn affect cloud formation and the surface energy budget. For example, the analysis in Sect. 3 shows that clouds have a too large vertical extent in the ASR, with a too low cloud base and a too high cloud top. This could indicate that multiple cloud layers exist throughout the atmospheric column. ERA-Interim is shown to often have multiple cloud tops, while the vertical extent of the lower cloud layer generally agrees with observations. Analyzing multiple and precipitating cloud layers from cloud radar is difficult, since the radar's sensitivity to hydrometeor

size saturates the signal and only the uppermost cloud top and the precipitation below is indicated.

Analysis of cloud properties showed that in the ASR, (1) the clouds did not persist during DoY 240–245, when they should have, and (2) even when they did at other times, they were generally too optically thin, especially in ASR2. These results are consistent with the results from the LWP where there is a low amount of liquid water present. This clearly illustrates that for a model to get the cloud cover, cloud geometry and surface energy balance right, the microphysics, such as the amount of water in each phase, also need to be correct.

The amount of incoming solar radiation (SWD) was too high in the latest version of the ASR (ASR2). Additionally, the LWP as well as the IWP were underestimated in ASR2 compared to the observations. This result suggests that ASR2 generates too few clouds or at least clouds that are too optically thin. In the older version of the ASR (ASR1), the error in SWD was smaller, except when clouds were erroneously absent. ASR1 also had substantially higher LWP compared to ASR2, and was in better agreement with the observations. ERA-Interim has too low values of the shortwave radiation but still slightly more realistic LWP.

The relatively poor representation of the longwave radiation in ASR also indicates differences in the cloud layer between the reanalyses and the observations. The incoming longwave radiation (LWD) is underestimated in both versions of the ASR, especially in ASR2. The outgoing longwave radiation (LWU) is also well below the observations in both versions of the ASR and at the same time the surface temperature is too cold (especially DoY 239–245). ASR also effectively has an upper cap on LWU, due to the fact that the surface temperature is, realistically, limited to around zero while snow is melting, combined with an assumption that the emissivity is lower than unity.

For ERA-Interim, a warm bias was found for the surface temperature. LWD and LWU were substantially higher than in ASR and generally agreed more closely with observations. In fact, ERA-Interim uses the same surface emissivity as ASR, but maintains a more realistic LWU because of the higher surface temperatures. While ERA-Interim cloud base was often too high, the cloud-top statistics agreed well with the ASCOS observations. Shortwave radiative fluxes (SWU and SWD) were underestimated, but the amount of liquid water (LWP) was reasonable compared to the measured values. Ice water path was not as well simulated, with ERA-Interim featuring excessive values and ASR underestimating mixed-phase cloud occurrence. However, there is a relatively high uncertainty in the observations of cloud ice.

When the turbulent fluxes in ASR were examined, the latent heat flux was calculated in two different ways. Since ASR uses only the upward part of the moisture flux – here displayed as the moisture flux times the latent heat of vaporization at 0 °C (L^*QFX) – in the calculations rather than the full latent heat flux (LH), it made sense to include both

variables in the analysis. Both results are compared with the latent heat (LH) from ERA-Interim and the ASCOS observations. L^*QFX from ASR obviously does not include any negative values, which affects the evaluation results positively but for the wrong reasons; this should also have an influence on the total energy and moisture budget in the system.

It is well known that mesoscale models produce more spatial variability than coarser resolution models and, as a result, often perform poorer in terms of standard metrics such as bias, RMSE and correlation. Slight spatial shifts in otherwise realistic spatial patterns can give rise to a negative outcome where the spatial resolution affects the results. On the other hand, a too coarse resolution may imply that small regional changes disappear in the outcome of the simulation. In the present case, for the central Arctic Ocean, where terrain plays no role, it seems like other factors, apart from only the spatial resolution, related to the handling of sub-grid-scale processes are more important when comparing the performance of the regional versus global reanalysis.

Quite clearly, the atmospheric model used in ERA-Interim (IFS Cy31r1) is a very well-tuned system: ERA-Interim sometimes showed better agreement with observations than ASR, despite having fewer physically based parameterizations, such as the simple temperature-dependent separation of cloud water into liquid and ice. This may be due to the fact that ERA-Interim is an operational weather model and is constantly evaluated against observations. In addition to the difference in the resolution and model physics between ASR and ERA-Interim, there is also a difference in the assimilation system: ASR uses 3DVar and ERA-Interim a 4DVar system. This may, of course, also have an impact on the outcome of this study.

5 Conclusions

This study has focused on evaluating two versions of the Arctic System Reanalysis (ASR) with observations from the ASCOS field campaign, which was conducted in August and September of 2008. Included in the comparison was also the global ERA-Interim reanalysis in order to give a comprehensive view of the performance of a regional reanalysis and to evaluate the advantages and disadvantages with a global reanalysis; recall that the two versions of the Arctic System Reanalysis used here were forced at the lateral boundaries by ERA-Interim.

The results show that the developmental versions of the ASR provide a good representation of the general meteorological situation over the Arctic Ocean, but there is room for improvement in the representation of moisture and clouds. The performance of ASR in this context critically impacts the surface energy balance and thus surface temperature.

The main conclusions are the following:

- Both versions of the ASR (ASR1 and ASR2, cf. Table 1) and ERA-Interim describe the regional climate dynamics and thermodynamics (T_{2m} , P_s and U_s) in a reasonable way. However, the ERA-Interim has a systematic warm bias (1.3 °C), while the ASR has a cold bias of about the same magnitude on average, but this is mostly concentrated to distinct periods when the temperature was clearly underestimated by ASR. ERA-Interim also shows a positive bias in the near-surface specific humidity. To conclude, the temporal development on the synoptic scale is quite satisfactory in all three reanalyses.
- The vertical profiles of wind, temperature and relative humidity from the ASR and ERA-Interim are in good agreement with the observations. The main problems occur near the surface, but some smaller problems are also found in the free troposphere (> 1 km). It should be noticed that the surface layer parameterization is different between the two versions of the ASR, but the outcome is still very similar.
- Both versions of ASR, with reasonably sophisticated microphysics descriptions, fail to reproduce a week-long period of mixed-phase stratocumulus (one of the most common cloud types found in the Arctic) observed during ASCOS, whereas the much simpler microphysics description in ERA-Interim manages to capture the event in a more realistic way. The lack of clouds in the ASR during this period has large consequences for the surface energy balance and thus for errors in the near-surface temperature and the stability of the boundary layer.
- The newer version of the ASR (ASR2) is not necessarily better than the older in reproducing the evolution of the Arctic cloud layer and the moisture content of the lower atmosphere. The different microphysical parameterizations in ASR1 and ASR2 give rise to different distributions of liquid and ice in the clouds, where the parameterization applied in ASR1 (Morrison et al., 2005) seems to perform slightly better in the Arctic region than ASR2 (Hong et al., 2004), at least for these observed conditions.
- All reanalyses deviated from the observations in terms of their representation of the surface properties: both the surface albedo and the surface emissivity were too low compared to observations. To some extent these errors were found to compensate for other errors in the surface energy balance. The observed turbulent fluxes are generally small and all the reanalyses exaggerate the magnitude of turbulent fluxes of momentum, sensible heat and moisture. ASR has long periods with a

substantial downward moisture flux in contrast to the observations.

The comparison between the reanalyses and the observations clearly illustrate how a modeling problem in one aspect of the atmosphere, here the clouds, immediately feeds back to other parameters especially near the surface and in the boundary layer.

Other regional models have also been evaluated using ship-based measurements from the Arctic (e.g., Tjernström et al., 2005; Jakobson et al., 2012). When validating reanalyses, especially examining the vertical profiles over the central Arctic using the drifting ice station Tara, Jakobson et al. (2012) found large errors in all reanalyses, with warm and moist biases in, for example, ERA-Interim and ERA-40, which has also been observed earlier (Curry et al., 2002; Vihma et al., 2002). Also, Lüpkes et al. (2010) found warm and moist biases in ERA-Interim boundary layer. Previous evaluations of the Polar WRF (Hines and Bromwich, 2008; Bromwich et al., 2009; Hines et al., 2011; Wilson et al., 2011; Wilson et al., 2012) show good agreement in surface pressure, but high-frequency fluctuations in surface temperature due to the variability in liquid and ice clouds. However, the simulations captured the synoptic variability in the Arctic and show improvement in the surface energy balance over Greenland compared to previous generations of the model.

It is intriguing that the much more sophisticated descriptions of the cloud microphysics in both ASR versions compared to ERA-Interim did not significantly improve the modeling of cloud properties. It may be that more advanced microphysics parameterizations, with separate conservation relations for different phases and types of hydrometeors, require a much more careful representation of the aerosol population and aerosol–cloud interactions. For example, the prediction of parameters such as the number of available cloud condensation and ice nuclei may have to be more dynamic. A consequence could be that advanced cloud treatment requires aerosol descriptions at a similar level of complexity (cf. Ekman et al., 2011; Seifert et al., 2012). For this type of clouds in summer, when the clouds are only slightly super-cooled, a cloud scheme such as that in ERA-Interim with only one prognostic cloud-water equation and a simple temperature-dependent separation of cloud water into liquid and ice seems to perform better, but not necessarily for the right reasons. As long as the model produces any cloud water, this formulation dictates that clouds must have some liquid and some ice. In a more sophisticated scheme, such as in ASR, liquid and ice may develop independently and clouds may glaciate and disappear (cloud ice grows at the expense of liquid and the ice falls out of the cloud) if the balance between sources and sinks of cloud liquid and ice is not correct (e.g., Prenni et al., 2007).

It should be kept in mind that this study was based on observations from a limited time period, during one summer, August and early September 2008, and for a limited region

of the Arctic Ocean, north of Svalbard. However, the conclusions should still be valid for similar meteorological conditions, no matter when and where in the Arctic Region they occur, assuming the number of available cloud condensation and ice nuclei is not completely different. A similar evaluation on a longer timescale would be highly valuable but requires long-term observations of processes in the Arctic, which is logistically very challenging. Today, this type of data unfortunately does not exist.

Acknowledgements. This work is part of ASCOS (the Arctic Summer Cloud-Ocean Study). ASCOS is an IPY project under the umbrella of AICIA-IPY and is an endorsed SOLAS project. ASCOS was made possible by funding from the Knut and Alice Wallenberg Foundation, the DAMOCLES Integrated Research Project from the European Union 6th Framework Programme and by the Bert Bolin Centre for Climate Research. We are grateful to the ASCOS science crew for making quality-controlled observations available to us, especially Ian Brooks, Cathryn Birch, Thorsten Mauritsen, Joseph Sedlar, Matthew Shupe and Ola Persson. Caroline Leck and Michael Tjernström were co-chief scientists, and the Swedish Polar Research Secretariat (SPRS) provided access to the icebreaker *Oden* and logistical support. We are also grateful to the SPRS logistical staff and the *Oden's* Captain Mattias Peterson and his crew. The Swedish National Research Council funded M. Tjernström and C. Wesslén for this study under grant 2009-3085. The US National Science Foundation funds the Arctic System Reanalysis (ASR) via grants ARC-0733023 and ARC-1144117 to D. H. Bromwich. G. de Boer acknowledges funding from the US National Science Foundation (ARC-1023366 and ARC 1203902) and the US Department of Energy (DE-SC0008794). We also acknowledge the kind assistance by Dick Dee from ECMWF in answering all our questions regarding ERA-Interim.

Edited by: E. Swietlicki

References

- ACIA: Impact of a warming Arctic: Arctic Climate Impact Assessment, Cambridge University Press, 989–1020, 2005.
- Andreas E., Guest, P. S., Persson, P. O. G., Fairall, C. W., Horst, T. W., Moritz, R. E., and Semmer, S. R.: Near-surface water vapor over sea ice is always near ice saturation. *J. Geophys. Res.*, 107, 8032, doi:10.1029/2000JC000411, 2002.
- Barker, D. M., Huang, X.-Y., Liu, Z., Auligne, T., Zhang, X., Rugg, S., Katheri, A. A., Bourgeios, A., Bray, J., Chen, Y., Demirtas, M., Guo, Y., Henderson, T., Huang, W., Lin, H.-C., Michalakes, J., Rizvi, S., and Zhang, X.-Y.: The Weather Research and Forecasting (WRF) Model's Community Variational/Ensemble Data Assimilation System: WRFDA. *B. Am. Meteor. Soc.*, 93, 831–843, 2011.
- Bromwich, D. H., Hines, K. M., and Bai, L.-S.: Development and testing of Polar WRF: 2. Arctic Ocean. *J. Geophys. Res.*, 114, D08122, doi:10.1029/2008JD010300, 2009.
- Bromwich D. H., Kuo, Y.-H., Serreze, M. C., Walsh, J. E., Bai, L.-S., Barlage, M., Hines, K. M., and Slater, A.: Arctic System Reanalysis: Call for Community Involvement, *EOS*, 91, 13–20, 2010.
- Cesana, G., Kay, J. E., Chepfer, H., English, J. M., and de Boer, G.: Ubiquitous low-level liquid-containing Arctic clouds: New observations and climate model constraints from CALIPSO-GOCCP. *Geophys. Res. Lett.*, 39, L20804, doi:10.1029/2012GL053385, 2012.
- Curry, J. A., Schramm, J. L., Alam, A., Reeder, R., Arbetter, T. E., and Guest, P.: Evaluation of data sets used to force sea ice models in the Arctic Ocean, *J. Geophys. Res.*, 107, 8027, doi:10.1029/2000JC000466, 2002.
- Chapman, W. L. and Walsh, J. E.: Simulations of Arctic temperature and pressure by Global Coupled Models, *J. Climate*, 20, 609–632, 2007.
- de Boer, G., Chapman, W., Kay, J. E., Medeiros, B., Shupe, M. D., Vavrus, S., and Walsh, J. E.: A Characterization of the Present-Day Arctic Atmosphere in CCSM4, *J. Clim.*, 25, 2676–2695, doi:10.1175/JCLI-D-11-00228.1, 2012.
- 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., Holm, E. V., Isaksen, I., Källberg, 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., Thepaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.*, 137, 553–597, doi:10.1002/qj.828, 2011.
- Ekman, A. M. L., Engström, A., and Söderberg, A.: Impact of two-way aerosol-cloud interaction and changes in aerosol size distribution on simulated aerosol-induced deep convective cloud sensitivity. *J. Atmos. Sci.*, 68, 685–698, 2011.
- Fouquart, Y. and Bonnel, B.: Computation of solar heating of the Earth's atmosphere: A new parameterization, *Beitr. Phys. Atmos.*, 53, 35–62, 1980.
- Hanna, S. R.: Mesoscale Meteorological Model Evaluation Techniques with Emphasis on Needs of Air Quality Models, edited by: Pielke, R. A. and Pearce, R. P., in: Mesoscale modeling of the Atmosphere, Meteorological Monographs, Am. Meteorol. Soc., 47–62, 1994.
- Hines, K. M. and Bromwich, D. H.: Development and Testing of Polar WRF. Part I. Greenland Ice Sheet Meteorology. *Mon. Weather Rev.*, 136, 1971–1989, doi:10.1175/2007MWR2112.1, 2008.
- Hines, K. M., Bromwich, D. H., Bai, L.-S., Barlage, M., and Slater, A. G.: Development and testing of Polar WRF – Part III: Arctic land, *J. Climate*, 24, 26–48, doi:10.1175/2010JCLI3460.1, 2011.
- Holland, M. M. and Bitz, C. M.: Polar amplification of climate change in the coupled model intercomparison project, *Clim. Dynam.*, 21, 221–232, 2003.
- Hong, S.-Y., Dudhia, J., and Chen, S.-H.: A revised approach to ice-microphysical processes for the bulk parameterization of cloud and precipitation, *Mon. Weather Rev.*, 132, 103–120, 2004.
- IPCC: Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Solomon S., Qin D., Manning M., Chen Z., Marquis M., Averyt K. B., Tignor M., Miller H. L., Cambridge University Press, Cambridge, UK and New York, NY, USA, 653–685, 2007.

- Jakobson, E., Vihma, T., Palo, T., Jakobson, L., Keernik, H., and Jaagus, J.: Validation of atmospheric reanalyses over the central Arctic Ocean, *Geophys. Res. Lett.*, 39, L10802, doi:10.1029/2012GL051591, 2012.
- Karlsson, J. and Svensson, G.: The simulation of Arctic clouds and their influence on the winter present-day climate in the CMIP3 multi-model dataset, *Clim. Dynam.*, 36, 623–635, 2011.
- Karlsson, J. and Svensson, G.: Consequences of poor representation of Arctic sea-ice albedo and cloud-radiation interactions in the CMIP5 model ensemble, *Geophys. Res. Lett.*, 40, 4374–4379, doi:10.1002/grl.50768, 2013.
- Liu, X., Xie, S., Boyle, J., Klein, S. A., Shi, X., Wang, Z., Lin, W., Ghan, S. J., Earle, M., Liu, P. S. K., and Zelenyuk, A.: Testing cloud microphysics parameterizations in NCAR CAM5 with IS-DAC and M-PACE observations, *J. Geophys. Res.*, 116, D00T11, doi:10.1029/2011JD015889, 2011.
- Liu, J., Song, M., Horton, R. M., and Hu, Y.: Reducing spread in climate model projections of a September ice-free Arctic, *Proc. Natl. Acad. Sci. USA*, 110, 12571–12576, doi:10.1073/pnas.1219716110, 2013.
- Lüpkes, C., Vihma, T., Jakobson, E., König-Langlo, G., and Tetzlaff, A.: Meteorological observations from ship cruises during summer to the central Arctic: A comparison with reanalysis data, *Geophys. Res. Lett.*, 37, L09810, doi:10.1029/2010GL042724, 2010.
- Mauritsen, T., Sedlar, J., Tjernström, M., Leck, C., Martin, M., Shupe, M., Sjogren, S., Sierau, B., Persson, P. O. G., Brooks, I. M., and Swietlicki, E.: An Arctic CCN-limited cloud-aerosol regime, *Atmos. Chem. Phys.*, 11, 165–173, doi:10.5194/acp-11-165-2011, 2011.
- Morrison, H., Curry, J. A., and Khvorostyanov, V. I.: A New Double-Moment Microphysics Parameterization for Application in Cloud and Climate Models – Part I: Description, *J. Atmos. Sci.*, 62, 1665–1677, 2005.
- Murray, M. S., Anderson, L., Cherkashov, G., Cuyler, C., Forbes, B., Gascard, J. C., Haas, C., Schlosser, P., Shaver, G., Shimada, K., Tjernström, M., Walsh, J., Wandell, J., and Zhao, Z.: International Study of Arctic Change: Science Plan, ISAC International Program Office, Stockholm, 2010.
- Prenni, A., Harrington, J. Y., Tjernström, M., DeMott, P. J., Avramov, A., Long, C. N., Kreidenweis, S. M., Olsson, P. Q., and Verlinde, J.: Can Ice-Nucleating Aerosols Affect Arctic Seasonal Climate?, *B. Am. Meteorol. Soc.*, 88, 541–550, 2007.
- Richter-Menge J. and Jeffries M.: The Arctic, in: State of the Climate 2010, *B. Am. Meteorol. Soc.*, 92, S143–S160, 2011.
- Sedlar, J., Tjernström, M., Mauritsen, T., Shupe, M. D., Brooks, I. M., Persson, P. O. G., Birch, C. E., Leck, C., A. Sirevaag, and M. Nicolaus, A transitioning Arctic surface energy budget: the impacts of solar zenith angle, surface albedo and cloud radiative forcing, *Clim. Dynam.*, 37, 1643–1660, 2011.
- Seifert, A., Köhler, C., and Beheng, K. D.: Aerosol-cloud-precipitation effects over Germany as simulated by a convective-scale numerical weather prediction model, *Atmos. Chem. Phys.*, 12, 709–725, doi:10.5194/acp-12-709-2012, 2012.
- Serreze, M. C., Barrett, A. P., Stroeve, J. C., Kindig, D. N., and Holland, M. M.: The emergence of surface-based Arctic amplification, *The Cryosphere*, 3, 11–19, doi:10.5194/tc-3-11-2009, 2009.
- Shupe, M. D., Daniel, J. S., de Boer, G., Eloranta, E. W., Kollias, P., Long, C. N., Luke, E. P., Turner, D. D., and Verlinde, J.: A focus on mixed-phase clouds, The status of ground-based observational methods, *B. Am. Meteorol. Soc.*, 87, 1549–1562, 2008.
- Skamarock, W., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D., Duda, M. G., Huang, X.-Y., and Wang, W.: A Description of the Advanced Research WRF Version 3, NCAR Technical Note NCAR/TN-475+STR, doi:10.5065/D68S4MNVH, 2008.
- Svensson G. and Karlsson, J.: On the Arctic wintertime climate in global climate model. *J. Climate*, 24, 5757–5771, 2011.
- Taylor K. E.: Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*, 106, 7183–7192, 2001.
- Tjernström, M., Žagar, M., Svensson, G., Cassano, J., Pfeifer, S., Rinke, A., Wyser, K., Dethloff, K., Jones, C., and Semmler, T.: Modeling the Arctic Boundary Layer: An evaluation of six AR-CMIP regional-scale models with data from the SHEBA project, *Bound.-Layer Meteorol.*, 117, 337–381, 2005.
- Tjernström M., C. Leck, P. O. G. Persson, M. L. Jensen, S. P. Oncley and A. Targino, The summertime Arctic atmosphere: Meteorological measurements during the Arctic Ocean Experiment (AOE-2001). *Bull. Amer. Meteor. Soc.*, 85, 1305–1321, 2004.
- Tjernström M., Birch, C. E., Brooks, I. M., Shupe, M. D., Persson, P. O. G., Sedlar, J., Mauritsen, T., Leck, C., Paatero, J., Szczodrak, M., and Wheeler, C. R.: Meteorological conditions in the central Arctic summer during the Arctic Summer Cloud Ocean Study (ASCOS), *Atmos. Chem. Phys.*, 12, 6863–6889, doi:10.5194/acp-12-6863-2012, 2012.
- Tjernström, M., Leck, C., Birch, C. E., Brooks, B. J., Brooks, I. M., Bäcklin, L., Chang, R. Y.-W., Granath, E., Graus, M., Hansel, A., Heintzenberg, J., Held, A., Hind, A., de la Rosa, S., Johnston, P., Knulst, J., de Leeuw, G., Di Liberto, L., Martin, M., Matrai, P. A., Mauritsen, T., Müller, M., Norris, S. J., Orellana, M. V., Orsini, D. A., Paatero, J., Persson, P. O. G., Gao, Q., Rauschenberg, C., Ristovski, Z., Sedlar, J., Shupe, M. D., Sierau, B., Sirevaag, A., Sjogren, S., Stetzer, O., Swietlicki, E., Szczodrak, M., Vaattovaara, P., Wahlberg, N., Westberg, M., and Wheeler, C. R.: The Arctic Summer Cloud-Ocean Study (ASCOS): overview and experimental design, *Atmos. Chem. Phys. Discuss.*, 13, 13541–13652, doi:10.5194/acpd-13-13541-2013, 2013.
- Vihma, T., Uotila, J., Cheng, B., and Launiainen, J.: Surface heat budget over the Weddell Sea: buoy results and comparisons with large-scale models, *J. Geophys. Res.*, 107, 3013, doi:10.1029/2000JC000372, 2002.
- Walsh J. E., Kattsov, V. M., Chapman, W. L., Govorkova V., and Pavlova, T.: Comparison of Arctic climate simulations by uncoupled and coupled global models, *J. Climate*, 15, 1429–1446, 2002.
- Westwater, E. R., Han, Y., Shupe, M. D., and Matrosov, S. Y.: Analysis of integrated cloud liquid and precipitable water vapor retrievals from microwave radiometers during SHEBA, *J. Geophys. Res.* 106, 32019–32030, 2001.
- Wilson, A. B., Bromwich, D. H., and Hines, K. M.: Evaluation of Polar WRF forecasts on the Arctic System Reanalysis domain: Surface and upper air analysis, *J. Geophys. Res.*, 116, D11112, doi:10.1029/2010JD01501, 2011.
- Wilson, A. B., Bromwich, D. H., and Hines, K. M.: Evaluation of Polar WRF forecasts on the Arctic System Reanalysis domain. 2. Atmospheric hydrologic cycle. *J. Geophys. Res.*, 117, D04107, doi:10.1029/2011JD016765, 2012.