

This discussion paper is/has been under review for the journal Atmospheric Chemistry and Physics (ACP). Please refer to the corresponding final paper in ACP if available.

# Near-surface meteorology during the Arctic Summer Cloud Ocean Study (ASCOS): evaluation of reanalyses and global climate models

G. de Boer<sup>1,2</sup>, M. D. Shupe<sup>1,2</sup>, P. M. Caldwell<sup>3</sup>, S. E. Bauer<sup>4,5</sup>, P. O. G. Persson<sup>1,2</sup>, J. S. Boyle<sup>3</sup>, M. Kelley<sup>5</sup>, S. A. Klein<sup>3</sup>, and M. Tjernström<sup>6</sup>

<sup>1</sup>The University of Colorado, Cooperative Institute for Research in Environmental Sciences, Boulder, CO, USA

<sup>2</sup>NOAA Earth System Research Laboratory, Physical Sciences Division, Boulder, CO, USA

<sup>3</sup>Lawrence Livermore National Laboratory, Livermore, CA, USA

<sup>4</sup>Columbia University, The Earth Institute, New York, NY, USA

<sup>5</sup>NASA Goddard Institute for Space Studies, New York, NY, USA

<sup>6</sup>Stockholm University, Department of Meteorology, Stockholm, Sweden

Received: 27 June 2013 – Accepted: 3 July 2013 – Published: 23 July 2013

Correspondence to: G. de Boer (gijs.deboer@colorado.edu)

Published by Copernicus Publications on behalf of the European Geosciences Union.

## Abstract

Atmospheric measurements from the Arctic Summer Cloud Ocean Study (ASCOS) are used to evaluate the performance of three reanalyses (ERA-Interim, NCEP/NCAR and NCEP/DOE) and two global climate models (CAM5 and NASA GISS ModelE2) in simulation of the high Arctic environment. Quantities analyzed include near surface meteorological variables such as temperature, pressure, humidity and winds, surface-based estimates of cloud and precipitation properties, the surface energy budget, and lower atmospheric temperature structure. In general, the models perform well in simulating large scale dynamical quantities such as pressure and winds. Near-surface temperature and lower atmospheric stability, along with surface energy budget terms are not as well represented due largely to errors in simulation of cloud occurrence, phase and altitude. Additionally, a development version of CAM5, which features improved handling of cloud macro physics, is demonstrated to improve simulation of cloud properties and liquid water amount. The ASCOS period additionally provides an excellent example of the need to evaluate individual budget terms, rather than simply evaluating the net end product, with large compensating errors between individual surface energy budget terms resulting in the best net energy budget.

## 1 Introduction

Both modeling and observational studies demonstrate that the Arctic is warming at a rate faster than the rest of the globe (e.g. IPCC, 2007; Serreze et al., 2009; Serreze and Francis, 2006; Rigor et al., 2000). While planetary warming trends are thought to be largely the result of elevated greenhouse gas concentrations, the “Arctic Amplification” described above is yet to be fully understood. Various ideas have been presented on possible drivers for this Arctic amplification, including feedbacks resulting from changes to snow and ice areal coverage (e.g. Curry et al., 1995), changes to clouds and atmospheric composition (e.g. Kay and Gettelman, 2009; Screen and

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

◀

▶

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



Simmonds, 2010; Boé et al., 2009), large scale circulation patterns in both the atmosphere and ocean (e.g. Graversen et al., 2008) and natural, low-frequency variability (e.g. Chylek et al., 2009). To date, there is little consensus on which of these processes is most important to understanding Arctic amplification, or even whether Arctic Amplification as described constitutes a robust signal (Polyakov et al., 2002).

Part of the reason for this lack of understanding is a general dearth of observations at Arctic latitudes. Because of this limitation, many of the studies mentioned in the previous paragraph utilize modeling tools such as reanalysis products and global climate models to arrive at their conclusions. Ironically, the reasoning behind the need for using these modeling tools (limited observational records) also creates a challenging environment for model validation and limits the ability of data assimilation techniques to constrain models. In particular, the representation of cloud and radiation processes has been demonstrated to be problematic at high latitudes. Walsh et al. (2008) used measurements from the Atmospheric Radiation Measurement (ARM) program's North Slope of Alaska site at Barrow ( $71.3^{\circ}$  N,  $156.6^{\circ}$  W) to evaluate cloud and radiation fields in four different atmospheric reanalyses. These included the National Center for Environmental Prediction (NCEP) and National Center for Atmospheric Research (NCAR) reanalysis (hereafter R-1, Kalnay et al., 1996), the European Centre for Medium-Range Weather Forecasting 40 yr reanalysis (ERA-40, Uppala et al., 2005), the NCEP-NCAR North American Reanalysis (NARR, Mesinger et al., 2006) and the Japan Meteorological Agency and Central Research Institute of Electric Power Industry 25 yr reanalysis (JRA-25, Kazutoshi et al., 2007). This work illustrated that the reanalyses generally under-predicted area-weighted cloud fraction, resulting in a corresponding over-prediction of downwelling shortwave and an under-prediction of downwelling longwave radiative flux densities at the earth's surface. These differences are alarming in that the measurements were obtained at a location that includes routine radiosonde launches, which are subsequently assimilated into the reanalysis products. This direct integration of local measurements should result in simulated large-scale atmospheric conditions similar to those observed, a luxury not shared by more remote portions of the Arctic.

**ASCOS model evaluation**

G. de Boer et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

◀

▶

◀

▶

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



More recently, Zib et al. (2012) used measurements from the Baseline Surface Radiation Network (BSRN) stations at Barrow and Ny-Alesund ( $78.9^{\circ}$  N,  $11.9^{\circ}$  E) to evaluate cloud and radiative properties in the National Aeronautics and Space Administration (NASA) Modern-Era Retrospective Analysis for Research and Applications (MERRA, 5 Rienecker et al., 2011), the NCEP Climate Forecast System Reanalysis (CFSR, Saha et al., 2010), the National Oceanographic and Atmospheric Administration (NOAA) Twentieth Century Reanalysis Project (20CR, Compo et al., 2011), the ECMWF-Interim reanalysis (hereafter ERA-I, Dee et al., 2011), and the NCEP-Department of Energy (DOE) reanalysis (hereafter R-2, Kanamitsu et al., 2002). This next generation of re- 10 analyses demonstrates large differences in cloud occurrence from one product to the next, including inconsistencies between relative cloud amounts from one site to the other. Radiative flux densities, while marginally better at Barrow, are still demonstrated to be problematic, with R-2 demonstrating the largest biases. Most recently, Jakobson et al. (2012) evaluated the performance of several reanalysis products over the central 15 Arctic Ocean. Using measurements from the Tara drifting ice station (Gascard et al., 2008; Vihma et al., 2008) the authors demonstrate that the ERA-Interim reanalysis outperforms several others, including R-2 and MERRA, using a ranking system. These rankings were calculated through comparison of the analysis results from these products for fields including air temperature, specific humidity, relative humidity and wind 20 speed.

Global climate models have similarly struggled with simulation of Arctic surface meteorology, clouds, and surface radiation. Walsh et al. (2002) demonstrated that while early climate models produced seasonal cloud cycles that were similar to observational estimates, the amount of cloud cover varied dramatically from one model to the 25 next. As may be expected, these discrepancies led to major differences in radiation, with summertime differences of nearly a factor of three in the surface radiative budget. A broader look at models involved with the 3rd Coupled Model Intercomparison Project (CMIP3) by Svensson and Karlsson (2011) illustrated large variability among the models in their representation of net surface energy flux on an Arctic-wide scale ( $> 66.6^{\circ}$  N)

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[|◀](#)[▶|](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

for present-day wintertime conditions. These discrepancies make deriving concrete results for the causes behind Arctic amplification challenging, particularly given the limited model validation efforts that have taken place. More recently, de Boer et al. (2012) performed an evaluation of 20th century simulations completed using the Community Climate System Model version 4 (CCSM4). They demonstrated that Arctic clouds, including their phase partitioning and liquid and ice water paths were grossly misrepresented in CCSM4, with notable impacts on the surface energy budget. Interestingly, these errors appeared to have only small impacts on the simulated surface air temperature, with model cold biases on the order of 1–2 K when compared to ERA-40, which itself has been shown to have 1.5 K warm biases compared to International Arctic Buoy Programme/Polar Exchange at the Sea Surface (IABP/POLES) measurements (Liu et al., 2007). Similar cloud and radiation discrepancies were brought out in work by Inoue et al. (2006) and Tjernström et al. (2008) while evaluating regional climate model performance over the Surface Heat Budget of the Arctic Ocean (SHEBA) Experiment site.

In the current work, we evaluate some of the tools described above using measurements obtained during the Arctic Summer Cloud Ocean Study (ASCOS, Tjernström et al., 2013). The ASCOS dataset provides us with a unique opportunity to assess the performance of models at high latitudes. Evaluations of basic surface meteorology and the surface energy budget as represented in atmospheric reanalyses and global climate models are completed using this dataset. Because this is a multi-global-model evaluation, with only limited variables and output formats available, averaging and interpolation of model results is necessary at times to ensure the most fair comparison possible. Not all variables are handled in the same way, and therefore these actions are outlined in the individual result sections as necessary. Section 2 provides a brief overview of ASCOS, Sect. 3 provides some background on the models, Sect. 4 outlines results of the evaluation, and finally Sect. 5 provides discussion of these results and a summary.

[Title Page](#)  
[Abstract](#) [Introduction](#)  
[Conclusions](#) [References](#)  
[Tables](#) [Figures](#)

[◀](#) [▶](#)  
[◀](#) [▶](#)  
[Back](#) [Close](#)  
[Full Screen / Esc](#)

[Printer-friendly Version](#)  
[Interactive Discussion](#)



## 2 ASCOS

During August of 2008, the Swedish ice breaker *Oden* served as a drifting base camp for the Arctic Summer Cloud Ocean Study (ASCOS, Tjernström et al., 2013). The vessel headed north from Svalbard with a goal to spend as much time as possible in the central Arctic ice pack. Between 12 August and 2 September, *Oden* was moored to a  $3 \times 6$  km ice floe just north of  $87^{\circ}$  N latitude (see Fig. 1). Much of the analysis performed in the current work is completed for this drifting time period. During this time, the expedition first experienced a ten day period where surface air temperatures ( $T_{\text{air, sfc}}$ ) were near, or slightly above the freezing point and low cloud cover was abundant. This was followed by a slightly colder period with  $T_{\text{a, sfc}}$  between 266 and 272 K. Finally, starting around 29 August, temperatures fell further (below 263 K) and open water began to close with the initiation of the fall freeze-up of sea ice.

A wide variety of instrumentation was deployed during ASCOS in order to comprehensively sample the atmosphere and surface. Included were a series of in-situ instruments measuring properties such as air temperature, air pressure, relative humidity, and wind speed and direction. This includes sensors at the surface as well as sensors used for profiling, with profiles obtained four times daily via radiosonde. In addition to in-situ instrumentation, surface based remote sensors were deployed to measure clouds and radiation. These included a 35 GHz millimeter cloud radar (MMCR, Moran et al., 1998), ceilometers, a scanning 60 GHz radiometer, a dual wavelength (24/31 GHz) microwave radiometer (MWR, Westwater et al., 2001), an S-Band cloud and precipitation radar, a 449-MHz wind profiler, and a suite of broadband radiometers measuring surface short- and longwave radiation flux densities (Sedlar et al., 2011).

When combined, this suite of instruments provide an in-depth view of atmospheric processes. In addition to derivation of cloud macrophysical properties (height, thickness, fractional coverage, etc.), information on cloud phase is obtainable via methods described in Shupe (2007). Additionally, estimates of cloud liquid water path (LWP) are

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



obtained from the MWR, with an estimated error of approximately  $25\text{ g m}^{-2}$  (Westwater et al., 2001).

Unlike several other large measurement campaigns over the Arctic Ocean, radiosonde measurements from ASCOS were not submitted to the Global Telecommunications System (GTS, Birch et al., 2012), meaning that they were not included in the datasets used for model initialization and data assimilation. Conversely, 6 hourly surface-based wind and pressure observations from the *Oden* were submitted to the GTS. While the one month observation period does not provide us with an long-term record against which to evaluate our models, this data set does represent one of the longest-lasting, comprehensive sets of high-Arctic measurements not included in the GTS, to our knowledge. Therefore, ASCOS measurements represent an independent data set that can be used to directly evaluate reanalysis model performance.

### 3 Models evaluated

#### 3.1 ERA-interim

The European Center for Medium Range Weather Forecasting (ECMWF) has released multiple reanalysis products. Here, we evaluate the performance of the most recent version, the ERA-Interim (hereafter ERA-I, Dee et al., 2011). ERA-I provides global analyses of atmospheric and surface state variables from 1989 to the present. It builds upon the successful ERA-40 (Uppala et al., 2005) product which covered the years between 1957–2002, with notable differences. ERA-I extends the number of atmospheric pressure levels archived from 23 to 37, is run at a higher resolution (T255) and includes a number of additional cloud and state variables in its output. Advances in data assimilation techniques include the introduction of a 12 hourly 4-D-varitional (4-D-VAR) scheme, improved formulation of background error constraint, improved humidity analysis, improved model physics, and improved quality control, to name a few. Observationally, ERA-I utilizes all of the observational datasets from the ERA-40 project,

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

and adds altimeter wave height information, winds and clear sky radiances from EU-METSAT, ozone profiles and radio occultation measurements. Boundary forcing fields come from a combination of ERA-40 reanalysis output (before 2001) and the ECMWF operational analysis (after 2001).

### 5 3.2 NCEP-NCAR (R-1)

The National Centers for Environmental Prediction (NCEP) teamed up with the National Center for Atmospheric Research (NCAR) to produce a reanalysis product to support research and climate monitoring communities (NCEP-NCAR, hereafter R-1, Kalnay et al., 1996). Originally planned to cover the 40 yr period from 1957–1996, 10 the project was later expanded to include years up to present day. The underpinning model's horizontal resolution (T62) is significantly lower than that of ERA-I, as is the number of vertical levels (28 for R-1). This results in a lower-atmospheric vertical resolution ranging from 80 m (lowest grid box) to 384 m (7th grid box at 850 mb). Examples of assimilated datasets include a variety of satellite-based measurements, radiosondes 15 from the GTS, sea ice characteristics from the ECMWF, and surface ocean data. R-1 utilizes a 3-D-varitional (3-D-VAR) analysis scheme known as a spectral statistical interpolation. Unlike for ERA-I, forecast products are not archived for R-1 beyond the 6 h forecast time. These 6 h forecast fields are in addition to analysis time values provided for limited variables.

### 20 3.3 NCEP-DOE (R-2)

To improve upon R-1, NCEP teamed up with the Program of Climate Model Diagnosis and Intercomparison (PCMDI) at Lawrence Livermore National Laboratory (LLNL) to create the NCEP-DOE AMIP-II reanalysis (hereafter R-2, Kanamitsu et al., 2002). R-2 utilizes the same spatial and temporal resolution as R-1 (T62, 28 levels, 6 h), and 25 makes use of similar raw observational datasets. Differences in the datasets include the removal of data from the Special Sensor Microwave Imager (SSM/I), and the addition

of limited additional data after 1993. Additionally, errors pertaining to bogus data in the Southern Hemisphere, snow cover analysis, humidity diffusion, oceanic albedo, relative humidity discontinuities and snowmelt were fixed for the R-2 product. New system components included in R-2 include rainfall assimilation over land to improve surface soil moisture, a smoothed orography and an updated treatment of snow. Other differences result from improvements to model physics made between the creation of R-1 and R-2. These include a new planetary boundary layer scheme, new shortwave radiation scheme, a retuned convective parameterization, improved cloud-top radiative cooling, updated cloud-tuning coefficients and further improvements to the radiation scheme. As with R-1, only 6 h forecasts are distributed, along with limited analysis fields.

### 3.4 CAM5

Another set of simulations evaluated in this study were completed using a recent version of the Community Atmosphere Model (CAM5.1, Neale et al., 2010). In order to recreate conditions from the ASCOS period, these simulations were completed using the Cloud-Associated Parameterizations Testbed (CAPT, Phillips et al., 2004). CAPT utilizes operational analyses from numerical weather prediction centers to initialize CAM5 and produce short term forecasts. In this instance, the European Center for Medium-Range Weather Forecasts (ECMWF) Year of Tropical Convection (YOTC) re-analysis was used to initialize forecasts within CAPT. The analysis data are interpolated from the finer-resolution analysis grid of  $0.150^{\circ}$  and 91 levels to the CAM5 grids using procedures outlined in Boyle et al. (2005). These procedures use a slightly different interpolation approach for each of the dynamic state variables (i.e. horizontal winds, temperature, specific humidity and surface pressure), along with careful adjustments to account for the difference in representation of the earths topography between models. A series of 6 day hindcasts are initialized every day at 00:00 UTC from the ECMWF analysis for the entire YOTC period from 1 May 2008 to 30 April 2010. Only the atmospheric winds, temperature and moisture are initialized, while the rest of the

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[|◀](#)[▶|](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

initial variables (land and atmosphere) come from an additional ECMWF-nudged run of the same model. Skin surface temperature and sea ice are prescribed using the NOAA Optimum Interpolation (OI) Sea Surface Temperature (SST) V2. These data are weekly means on a  $1^\circ \times 1^\circ$  grid and are interpolated in time from weekly to the model time step. Since the model has a spin-up period to adjust to ECMWF conditions, the ASCOS time series are created by concatenating hours 24–48 from each hind cast.

CAM5 was run with the finite volume dynamic core at resolution of  $0.90^\circ \times 1.250^\circ$  in the horizontal and utilizes 30 vertical levels. This version of CAM contains a range of significant enhancements and improvements in the representation of physical processes. Except for the deep convection scheme, most other physical parameterizations have been updated from CAM4 to CAM5 (Neale et al., 2010). In addition, a three-mode modal aerosol scheme (MAM3) has been implemented in CAM5 to provide internally mixed representations of number concentrations and mass for Aitken, accumulation and coarse aerosol modes (Liu et al., 2011). These major physics enhancements permit new research capability for assessing the impact of aerosol on cloud properties. In particular, they provide a physically based estimate of the impact of anthropogenic aerosol emissions on the radiative forcing of climate by clouds.

We also analyze simulations from a development version of CAM5 which includes several changes expected to improve model clouds (CAM5-PF). In particular, this new simulation includes a new parameterization for cloud macrophysics (which combines stratiform liquid cloud fraction, condensation, and evaporation) based on a truncated Gaussian PDF for sub-grid variability in saturation excess (defined as total water mixing ratio minus ice mixing ratio minus liquid saturation mixing ratio). This truncated Gaussian also replaces the Gamma sub-grid scale distribution previously used for microphysics. This improves inter-process consistency and slightly reduces microphysical depletion rates. Further improvements include sub-stepping macro- and microphysics to improve coupling between condensational growth and microphysical erosion of cloud and fixing an inconsistency between the liquid water content and droplet number used by microphysics. The latter change distributes liquid across more droplets (further re-

## ASCOS model evaluation

G. de Boer et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[|◀](#)

[▶|](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



ducing microphysical depletion), while the former keeps parameterizations from bouncing liquid between physically unrealistic states. These collective changes are expected to increase high-latitude cloud fraction and liquid water path, while handling microphysical depletion more appropriately.

### 5 3.5 NASA GISS-ModelE2

GISS-ModelE2 simulations were completed using an updated version of the NASA GISS GCM ModelE2 specifically developed for the 5th IPCC assessments (CMIP5). The CMIP5 version of the GISS-ModelE2 is improved over that used for CMIP3 (and described in Schmidt et al. (2006) and Hansen et al. (2007)) in a number of respects (Schmidt et al., in preparation). Firstly, the model has a higher horizontal and vertical resolution (2 lat  $\times$  2.5 longitude, 40 layers). The vertical layers are distributed on a non-uniform grid, with spacing of roughly 25 mb (250 m) from the surface to 850 mb, and roughly 40–50 mb (400–700 m) from 850 to 415 mb. Secondly, various physics components have been upgraded from the CMIP3 version, namely the convection scheme, stratiform cloud scheme and gravity wave drag. The simulations discussed here further include the aerosol microphysics scheme MATRIX (Bauer et al., 2008).

For this work the model is run continuously for 2008, covering the ASCOS campaigns described above. In order to force representative meteorology in the GCM, the model uses prescribed sea surface temperatures and sea ice, and the horizontal wind components of the model are nudged towards the R-2 reanalysis. R-2 winds are available on a 6 hourly time step and are linearly interpolated to the model 30 min timestep. The aerosol scheme uses the CMIP5 emissions by Lamarque et al. (2010). This setup has previously been used by Bauer and Menon (2012) and de Boer et al. (2013) to evaluate forecast-mode simulations using ModelE.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

◀

▶

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



## 4 Notes on sampling

A major consideration in evaluations such as in this study is how to best analyze available measurements to appropriately represent the scales inherent to model grid boxes (McComiskey and Feingold, 2012). While in-situ and surface-based observations such as those obtained during ASCOS have the potential to capture process-level relationships, they do not necessarily capture the spatial variability included within a model grid box without alteration of obtained measurements. A simple approach is aggregation (averaging) of data over time scales that begin to capture the spatial variability in the model grid. Naively, it may be assumed that it would be appropriate to average over a time period that covers the full scale of the grid box assuming some advective velocity (e.g.,  $10\text{ ms}^{-1}$ ). At  $2^\circ$ , this requires averaging of periods on the order of 6–7 h. Using this technique is not practical, however, as it blurs the evolution of quantities occurring within the diurnal cycle. An alternative approach entails averaging over shorter periods (e.g., one hour) in order to capture some of the sub-grid scale variability in the measurements, while maintaining signals inherent in an evolving atmosphere. This short aggregation timescale is very appropriate for time periods featuring consistent large scale meteorological conditions and a relatively homogeneous surface, but may fail during frontal passages or at coastal sites. In this work, comparison is additionally complicated by the way in which parameters are presented within the reanalysis products. For example, while most variables are provided on a 6 hourly timescale, some of the values presented (e.g. liquid water path, cloud fraction) represent instantaneous values for that time (but still averaged across the model grid box), other variables (e.g. precipitation, radiation) are provided as average values over the 6 h period. Complicating things further, the observational datasets used for this study represent different time scales as well, with most quantities aggregated into hourly averages, and others represented as 3 hourly averages. In this evaluation, we have done our best to be consistent in our comparisons. For variables represented in the reanalyses as averages, we have taken the time period over which this average is taken and computed averages

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[|◀](#)[▶|](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[◀](#)

[▶](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



from the measurements and GCM output as well. Likewise, for instantaneous values, we have interpolated the measurements and GCM results to most closely match the time of the reanalysis dataset. The (at least) hourly averaging used on the measurements ensures that some spatial variability is accounted for in those datasets, even when “instantaneous” evaluations are done since the highest temporal resolution used is a one-hour average. Further details are provided in the individual sections in Sect. 5.

## 5 Results

### 5.1 Surface meteorology

Figure 2 gives an overview of the performance of different modeling products in simulating basic surface meteorological quantities for the time that *Oden* was drifting with the pack ice north of 87° N latitude. Included are (from top to bottom) 2 m air temperature ( $T_{\text{air, sfc}}$ , K), surface air pressure ( $P_{\text{air, sfc}}$ , mb), 2 m air relative humidity ( $\text{RH}_{\text{air, sfc}}$ , %), the 10 m zonal wind component ( $U$ ,  $\text{ms}^{-1}$ ) and the 10 m meridional wind component ( $V$ ,  $\text{ms}^{-1}$ ). For all variables, the model results have been linearly interpolated in space to the exact location of the *Oden* at a given time, reducing (but not eliminating) the influence of resolution that would result from a nearest grid cell comparison, and eliminating the influence of jumping between grid boxes with the movement of *Oden* during the campaign. The time series represented in Fig. 2 provide a comparison between the observations, reanalyses and GCMs. Biases are calculated and presented as distributions in the right hand column of Fig. 2, with black circles represent the median difference, the box representing the interquartile range (IQR) and the whiskers representing the extent of the 5th and 95th percentiles of the biases. Additionally, values for median biases and correlations for all evaluated variables are presented in Table 1. For the variables compared here, the reanalyses provide these variables as instantaneous values at the 6 hourly intervals. In order to most closely compare these quantities between the GCMs (which provide 1 h averages and 3 hourly instantaneous

output for CAM5 and GISS, respectively), the observations (which are provided as one-minute averages) and the reanalyses, observations are averaged on a one-hour scale to account for sub grid variability within the models, and then interpolated, along with the GCM output, to the 6 hourly reanalysis times.

5 Looking first at  $T_{\text{air, sfc}}$  (Fig. 2a), it is apparent that there is quite a bit of variability from one product to the next, and that there does not appear to be a clear “best” model in terms of agreement with observations. Statistically, R-2 outperforms other models in terms of the median bias, though it is clear that this bias is derived from a distribution with significant variability. ERA-I is correlated much more closely with the observations  
10 (correlation of 0.82) than the other models are. As noted above, the first portion of this period (roughly 12–17 August) features  $T_{\text{air, sfc}}$  near the freezing point. This is captured by the models, with most within 1–2° of the observations. There is a short period of cooler temperatures on 16 August which only the CAM5 models appear to simulate well. Unlike the original CAM5 version, CAM5-PF  $T_{\text{air, sfc}}$  does not recover back to the  
15 freezing point as quickly as in the observations, taking nearly two days to recover. Beyond 17 August, the observations, ERA-I and R-1 remain in this near-freezing point state for another four days. The ModelE2 has a cooler period from 17–19 August, while CAM5, CAM5-PF, and R-2 all cool off quickly on 18 August. Both versions of CAM5 cool to temperatures around 265 K by 20 August, which is cooler than the observed  $T_{\text{air, sfc}}$   
20 decrease to 267 K on 21 August. R-2 captures this cool period most successfully, with ERA-I, R-1 and ModelE2 all remaining too warm. The time period from 24 August until 31 August is an interesting one in that the observations along with R-1, ModelE2 and ERA-I reanalyses all feature temperatures that are only slightly below the freezing point, while CAM5 and R-2 generally have much colder (around 265 K) temperatures  
25 and CAM5-PF oscillates between these states. Physical mechanisms driving these differences will be explored later in the paper.

All of the model products generally perform well in simulating synoptic scale weather phenomena. Figure 2b illustrates a comparison of  $P_{\text{air, sfc}}$  at *Oden* between the different products. With the exception of R-1 and R-2, differences from the observations are

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

## ASCOS model evaluation

G. de Boer et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[|](#)[|](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

Ultimately, this works to reduce the median bias for these models (Fig. 2c, right-hand column), and ultimately result in some of the lowest biases of all of the models with poor correlation to the observations (Table 1).

## 5.2 Clouds and precipitation

5 ASCOS was in general very cloudy, a fact reflected in the observed cloud fractions (CF) in Fig. 3a. Three-hourly averaged values were at 100 % for most of the campaign, with only a few time periods in early September where values dropped significantly. As with surface meteorology, the models and observations provide different quantities for reporting cloud fraction. ERA-I, R-1 and R-2 all provide an instantaneous value every

10 six hours, while GISS ModelE2 provides a 3 hourly instantaneous value. The observations are provided as a 3 hourly average, while CAM5 provides 1 hourly averages. In order to attempt to make a more consistent comparison, CAM5, GISS and observational values are interpolated to the reanalysis times without averaging further. While CF is only a rough means for evaluating model performance in cloud simulation, the

15 model results do not compare favorably with the observations. ERA-I featured CFs that were closest to observations, with only a couple of periods where the CF fell below 100 %. The fact that ERA-I only reports instantaneous values at the 6 hourly times, while the other models (and observations) report 6 hourly averages, does make it likely that some differences will occur. During times where ERA-I CF did stray from 100 %, it

20 was not by much, falling only to values around 70–80 %. Towards the end of the measurement period, ERA-I retains 100 % CF, while the observations feature less clouds. In comparison, the other models generally feature less clouds than both the observations and ERA-I. Both R-1 and R-2 consistently produce far too few clouds. This is particularly true for the period between 26–29 August, where low, stratiform clouds persisted in the observations but both NCEP products have cloud fractions near 10 %. ModelE2 features significant variability in its representation of cloud fraction, spanning values between 0–100 %. The two CAM5 versions both under-predict cloud cover, and while the original CAM5 version has fewer very low values, CAM5-PF performs better during

25

13, 19421–19470, 2013

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

◀

▶

◀

▶

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



## ASCOS model evaluation

G. de Boer et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[|◀](#)[▶|](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

## ASCOS model evaluation

G. de Boer et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[|](#)[|](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

## 5.3 Surface energy budget

Figures 4 and 6 provide an overview of the surface energy budget terms as governed by:

$$Q_{\text{SFC}} = F_{\text{LW}} + F_{\text{SW}} - F_{\text{SH}} - F_{\text{LH}} \quad (1)$$

5 where  $F_{\text{LW}}$  and  $F_{\text{SW}}$  represent the net longwave and shortwave radiative fluxes, respectively,  $F_{\text{SH}}$  represents the surface sensible heat flux,  $F_{\text{LH}}$  represents the surface latent heat flux and  $Q_{\text{SFC}}$  represents the residual flux, including subtracted conduction terms. The influence of the conduction terms on the overall energy budget can be significant, but because we are evaluating the atmosphere these terms are not discussed here.

10 The signs of radiative terms are in line with their impact on the surface, with positive values acting to heat the surface and negative values acting to cool the surface, while the turbulent latent and sensible heat flux terms follow their traditionally applied sign convention, with negative values warming the surface and positive values cooling the surface. In all panels of Figs. 4 and 6 the ASCOS observations are presented in black, with various model results indicated in colored lines. For the surface energy budget terms, ERA-I, R-1 and R-2 all provide 6 hourly averaged values, while CAM5 and the observations provide 1 hourly averages and GISS ModelE2 provides 3 hourly instantaneous values. To best compare these different values, the 1 hourly averages from CAM5 and the observations, as well as the 3 hourly instantaneous GISS ModelE2 values are averaged over the 6 h period represented by the reanalyses.

25 Looking first at surface shortwave radiation (Fig. 4a–c), there are substantial differences between the different models. ERA-I appears to be the only model that comes close to resembling observed values, with CAM5, R-1 and R-2 featuring excessive downwelling shortwave radiation and ModelE2 featuring too little downwelling shortwave radiation, particularly in the first of the observational period. While some of this may be the result of differences in atmospheric chemical composition and the radiative transfer codes applied, this result should generally not be considered surprising if

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



## ASCOS model evaluation

G. de Boer et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[|](#)[|](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

## ASCOS model evaluation

G. de Boer et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[|◀](#)[▶|](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[◀](#)

[▶](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



## ASCOS model evaluation

G. de Boer et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[|](#)[|](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

## ASCOS model evaluation

G. de Boer et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[|◀](#)[▶|](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[|◀](#)[▶|](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

Somewhat incredibly, R-1, which was illustrated to be among the worst performing models for many of the individual terms discussed above, seems to most closely portray the net  $Q_{SFC}$  term with a bias of only  $2.97 \text{ W m}^{-2}$ . ERA-I and GISS ModelE2 are the only other products that demonstrate a positive bias in  $Q_{SFC}$  ( $7.24$  and  $22.38 \text{ W m}^{-2}$ , respectively). The other models all excessively lose heat at the surface, with R-2 featuring a median bias of  $-11.68 \text{ W m}^{-2}$ , CAM5 a bias of  $-18.62 \text{ W m}^{-2}$  and CAM5-PF a bias of  $-7.57 \text{ W m}^{-2}$ . A large portion of these biases is demonstrated to come from long wave radiative biases, with most of those being the result of the downwelling long wave. This illustrates the important role that cloud physics play in regulating this vital component of the climate system. Additionally, this result demonstrates clearly the influence of compensating errors on overall evaluations, and that in general, care must be taken to incorporate all of the budget terms in model evaluations, rather than simply looking at the net product term. If we instead look at the distribution of the absolute values of biases of the budget terms, only including the individual components ( $SW_{up}$ ,  $SW_{down}$ ,  $LW_{up}$ ,  $LW_{down}$ ,  $F_{SH}$ ,  $F_{LH}$ ), (Fig. 7) we see a very different representation of model performance. Here, ERA-I outperforms the other models with significantly smaller biases in the budget terms, with GISS ModelE2 a close second. CAM5PF demonstrates significant improvement over CAM5, with R-1 and R-2 falling in between these two. This figure can be thought of as a step towards diagnosing the models' abilities to get the right answer for the right reasons, rather than simply getting the right answer.

## 5.4 Lower tropospheric temperature structure

In order to evaluate the impact of the surface radiative balance terms on atmospheric state, here we assess the models' ability to simulate lower atmospheric temperature structure. Figure 9 illustrates temperature biases relative to radiosonde observations for the lowest 3000 m of the atmosphere for all five model time series. On the right of each figure is a profile demonstrating the mean (bold line) temperature bias profile, along with the interquartile range of these biases for each model grid box level. Biases detected for the lowest model levels follow biases in the surface energy budget, with

CAM5, CAM5-PF and R-2 all featuring cold biases close to the Earth's surface. CAM5-PF has a reduced cold bias when compared to CAM5, but also features a 1–2° warm bias higher in the atmosphere. R-1 and ERA-I have slight warm biases low in the atmosphere, which is not surprising given the neutral or positive biases they demonstrate in  $Q_{SFC}$ . Both R-1 and R-2 feature cold biases in the upper portion of the evaluated domain, while ERA-I and CAM5 have very small net biases at higher altitudes and CAM5-PF has a warm bias. All models except for ERA-I and GISS ModelE2 have warm biases between 500–1000 m, resulting from insufficient low-level cloud cover in those simulations.

10 Looking at this in a different way, we can evaluate the models' ability to simulate the potential temperature difference between the surface and 850 mb (Hereafter  $LTS_{850}$ , Fig. 10). For the ASCOS period, 850 mb fell between 1200–1600 m in altitude.  $LTS_{850}$  can be thought of as representative of the mean stability of the lower atmosphere, with stability increasing with the difference between these temperatures. All of the models accurately represent  $LTS_{850}$  during the first days of the observational period (12 August–15 August). After this, the models begin to diverge somewhat, with R-1 and R-2 generally being less stable than observations until 24 August, when R-2 suddenly shifts to generally being more stable than the observations. Looking back at Fig. 9, we can see that the instability (relative to observations) between 15–24 August in R-1 and 15 R-2 is the result of a combination of near-surface warm bias and slight 850 mb cold bias. After 24 August, R-2 features a cold near-surface temperature, which results in its shift to a more stable lower atmosphere. When looking at the distributions of  $LTS_{850}$  (Fig. 10b), the general picture is confirmed, with R-1 having an atmosphere that is generally less stable than observations, and R-2 more closely matching observations. 20 ERA-I generally has lower-atmospheric stabilities that are similar to those observed, with two exceptions. The largest differences between ERA-I and observations occur from 21–26 August and 1–2 September, where ERA-I is less stable than observations. During the 21–26 August time period, ERA-I exhibits a near-surface warm bias, resulting mainly from a high bias in the net surface long wave radiation (Fig. 4f). It ap-

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[|◀](#)[▶|](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

pears as though a similar bias in surface radiation may have occurred between 1–2 September, though we do not have surface radiation measurements for much of that time period. In addition to the surface radiation-induced near-surface temperature bias, ERA-I also has a cold bias around 850 mb for the 1–2 September time period, adding to the reduced stability. Overall, ERA-I is shown to be less stable than observations, and in particular does not capture more stable environments well (right hand tail of distributions in Fig. 10b). CAM5 and CAM5-PF simulations are generally too stable after 18 August. The exceptions to this are the periods from 23–24 August where both agree more closely with observed stability, and then the 24–27 August period where CAM5-PF agrees closely with observed stability. The excessive stability is largely the result of near-surface temperatures that are too cold in both versions of this model (Fig. 9). The biases in lower atmospheric stability again closely track the net long wave radiation at the surface (Fig. 4f), which for both versions of CAM5 is generally under predicted. Time periods mentioned above with improved performance match the time periods in which net long wave radiation is in better agreement with observations. Distributions of lower atmospheric stability for the entire ASCOS period demonstrate the slightly improved simulation of this quantity with CAM5-PF, with more of the values moving towards the less stable (left) side of Fig. 10b. Finally, evaluation of ModelE2 provides us with a somewhat puzzling result. Despite large biases in both the net surface energy flux, an evaluation of ModelE2 lower atmospheric stability results in a very favorable comparison with observations. The main exceptions to this occur between 27–28 August (too unstable) and 28–31 August (too stable). The GISS ModelE2 distribution of  $LTS_{850}$  matches observations more closely than any of the other models. Figure 10c and d demonstrate the relative contributions of near surface air temperature and 850 mb temperature on  $LTS_{850}$ . These values were obtained by using one of the observed values (either  $\theta_{SFC}$  or  $\theta_{850}$ ) together with the model value for the other in order to calculate  $LTS_{850}$ , and then calculating the bias in  $LTS_{850}$  relative to values calculated directly from the observations. As discussed above, most of the discrepancies result from differences in the near surface potential temperature, where variability

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[|◀](#)[▶|](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

is somewhat larger. The two main exceptions to this are the event on 27 August that all of the models appeared to misrepresent at 850 mb, and the general contribution of CAM5-PF's warm bias at 850 mb. CAM5's excessive stability is clearly illustrated to result mainly from the near-surface potential temperature, with  $LTS_{850}$  biases much larger when using the model  $\theta_{SFC}$  than when using the model  $\theta_{850}$ .

## 6 Summary and discussion

The different models had variable success in simulating the evaluated properties. A brief summary of the results is included here:

- *Surface meteorology*: with some minor exceptions, all of the models had small biases in wind speed and direction. Surface pressure was generally well simulated, with the exception of R-1 (−3.01 mb bias) and R-2 (4.93 mb bias). Variables more closely tied to clouds and radiation such as near surface temperature and humidity were shown to be more poorly simulated. Median near-surface temperature biases ranged between −3.39 K (CAM5) to 1.15 K (R-1), while median relative humidity biases ranged between −7.80 % (GISS ModelE2) to 2.39 % (ERA-I) with all models demonstrating substantial variability. ERA-I generally featured superior correlation with observations over the other models for these variables.
- *Clouds and precipitation*: cloud-related processes continue to represent a major stumbling block for accurate simulation of the Arctic environment. With the exception of ERA-I and GISS ModelE2, models struggled to produce the amount of cloud cover observed. Cloud liquid water path was underestimated by three of the four models that provided this quantity, with GISS ModelE2 producing the closest amount of liquid water (median bias of  $2.87 \text{ gm}^{-2}$ ), though care needs to be taken since the observations represent a lower limit, and likely an underestimate of the true liquid water present. CAM5-PF did demonstrate improvement over the standard version in simulating liquid water amount. While overall precip-

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[|◀](#)[▶|](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

5  
itation amounts were low during ASCOS, for what was recorded, median model errors were generally between 1–5 %, which is likely well within the errors associated with measuring precipitation. However, variability in those numbers was large, with models not reproducing three of the four more substantial precipitation events observed during ASCOS.

10  
– *Surface energy budget*: model biases in cloud-related processes discussed above result in large biases in the radiative components of the surface energy budget. Shortwave radiation biases are generally as may be expected, with models that have too few clouds also featuring too much downwelling shortwave radiation at the surface, and GISS ModelE2 agreeing most directly with measurements (median bias of  $3.97 \text{ W m}^{-2}$ ). Biases in upwelling shortwave radiation generally mirror those in downwelling shortwave, although the relative magnitude is modulated by the model surface albedos, which were demonstrated to vary widely with ModelE2 having the lowest (0.29–0.55) and R-2 having the highest (0.85–0.88). Net shortwave median biases were generally less than  $10 \text{ W m}^{-2}$  with R-1 and GISS ModelE2 being exceptions (median biases of 48.08 and  $19.57 \text{ W m}^{-2}$ , respectively). Longwave radiation biases are more variable. Except for ERA-I all of the models have negative median biases in downwelling long wave radiation resulting from a lack of optically thick clouds in the lower atmosphere. All of the models demonstrate a reduction in upwelling long wave radiation with a reduction in surface temperature, as is present in the observations, but the magnitude of upwelling long wave values varies from model to model and includes a feedback from the downwelling long wave caused by clouds (or a lack thereof). Net long wave radiation biases generally follow those in the downwelling long wave, with ERA-I comparing most favorably (median bias of  $-1.43 \text{ W m}^{-2}$ ) and other models featuring excessive radiation to the atmosphere. Sensible and latent turbulent heat fluxes are observed and modeled to be small compared to radiative fluxes, with R-2 featuring the largest biases in sensible heat fluxes and R-1 featuring the largest latent heat flux biases. Overall, ERA-I outperforms other models in repre-

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[|◀](#)[▶|](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

senting surface energy budget terms, as indicated by the absolute bias illustrated in Fig. 7. Somewhat incredibly, despite relatively large absolute biases, compensating errors result in R-1 having the smallest net surface flux bias and best correlation. The ModelE2 performs the worst, with a net surface energy median bias of  $22.38 \text{ W m}^{-2}$ .

– *Lower atmospheric temperature structure*: the models vary widely on their representation of lower atmospheric temperature. R-2 and both versions of CAM5 both feature near-surface cold biases for much of the ASCOS period, and all models except ERA-I feature a slight warm bias between 500–1200 m. R-1 and R-2 both have significant (around 2 K) cold biases near 3000 m. From a bulk stability perspective, R-1 and ERA-I tend to have lower atmospheres that is slightly less stable than observations, while CAM5 (both versions) tend to be too stable. ModelE2 and R-1 most closely match the observed lower tropospheric stability, as defined by the temperature difference between the surface and 850 mb.

The demonstrated model biases can have wide ranging impacts. For one, reanalysis output is sometimes used to force large-scale or local sea ice and ocean models to evaluate their performance against available observations (e.g., Brodeau et al., 2010; Miller et al., 2007). Curry et al. (2002) evaluated several datasets used to force sea ice models during the SHEBA time period and noted that substantial differences can be found between the different datasets. Our evaluation results in a similar conclusion, with notable differences between products in quantities relevant to ocean processes and sea ice growth and decay, such as winds, precipitation and radiation. Additionally, reanalyses are often used to evaluate the performance of climate models in simulation of the present-day atmosphere (e.g., de Boer et al., 2012; Schmidt et al., 2006). For variables with well-characterized, small biases, this may be acceptable, but for other variables (generally those related to clouds, precipitation and radiation), this sort of evaluation is generally inappropriate given the relatively comparable magnitudes of errors in the reanalyses and differences between reanalyses and models.

Biases in GCMs can cause serious issues with determining equilibrated climatic states. Since small biases in a fully-coupled climate model can run away into new climatic regimes, biases in the model themselves can, over several year periods, result in large errors in predicted climate. This is particularly true in a sensitive area such as 5 the Arctic where small shifts in climatic equilibrium can result in stark transitions, such as those from ice-covered to melted surfaces. One challenging aspect of a model evaluation such as this one is distinguishing between the errors resulting from the model itself and those resulting from the dataset from which forecasts were initialized. In the 10 present study, there are some GCM biases that appear to result from the model itself, such as the lack of low clouds in CAM5 and the associated surface energy balance errors and biases in lower atmospheric stability.

The ASCOS campaign helps to illustrate some of the issues faced by models in the Arctic environment. For the reanalyses, there may be significant advances in product accuracy with the integration of additional observations into the GTS. This would help 15 to improve the analyses used for forecast initialization, giving the model less room to stray from the measured atmospheric state. In terms of model performance, current parameterizations continue to be challenged in correctly simulating Arctic processes. While previous campaigns (e.g., SHEBA, ASCOS) have greatly improved our understanding of some of these processes, others could use additional improved observations to enhance our understanding. In particular, cloud processes are demonstrated 20 here to cause problems for all of the models involved, with errors in the representation of clouds translating to surface energy and near-surface temperature errors. This can have significant impacts on climate simulation as these imbalances will likely drive the climate into an altered state through various feedback processes. The short duration of ASCOS ultimately results in an incomplete evaluation of these models and similar 25 datasets need to be obtained for other times of year and other parts of the Arctic to complete such evaluations. Ultimately, increased observations are necessary to better understand and consequently simulate this environment. This is particularly true over

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[|◀](#)[▶|](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

minimally-studied areas such as the marginal ice zone, which continues to expand with the shrinking sea ice pack.

*Acknowledgements.* The authors wish to thank ECMWF for making YOTC analysis data available for research purposes. ASCOS was made possible by funding from the Knut and Alice Wallenberg Foundation, the DAMOCLES European Union 6th Framework Program, the Swedish National Research Council (VR), the US National Science Foundation (NSF), the National Atmospheric and Oceanic Administration (NOAA) and the UK Natural Environment Research Council (NERC). CAM5 and the CESM project are supported by the National Science Foundation and the Office of Science (BER) of the US Department of Energy. NCEP Reanalysis data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at <http://www.esrl.noaa.gov/psd/>. This work was prepared in part at the Cooperative Institute for Research in Environmental Sciences (CIRES) with support in part from the National Oceanic and Atmospheric Administration, US Department of Commerce, under cooperative agreement NA17RJ1229 and other grants. The statements, findings, conclusions, and recommendations are those of the author and do not necessarily reflect the views of the National Oceanic and Atmospheric Administration or the Department of Commerce. This research was supported in part by the Director, Office of Science, Office of Biological and Environmental Research of the US Department of Energy under Contract No. DE-AC02-05CH11231. Additionally, this work was supported by the National Science Foundation under grant numbers ARC-1023366 and ARC1203902 as well as the US Department of Energy under grant DE-SC0008794. Resources supporting this work were additionally provided by the NASA High-End Computing (HEC) Program through the NASA Center for Climate Simulation (NCCS) at Goddard Space Flight Center. The efforts of PMC, JSB, and SAK were supported by the Earth System Modeling program of the United States Department of Energy's Office of Science and were performed under the auspices of the United States Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344.

## References

Bauer, S. and Menon, S.: Aerosol direct, indirect, semi-direct and surface albedo effects from sector contributions based on the IPCC AR5 emissions for pre-industrial and present day conditions, *J. Geophys. Res.*, 117, D01206, doi:10.1029/2011JD016816, 2012. 19431

30

## ASCOS model evaluation

G. de Boer et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

Bauer, S. E., Wright, D. L., Koch, D., Lewis, E. R., McGraw, R., Chang, L.-S., Schwartz, S. E., and Ruedy, R.: MATRIX (Multiconfiguration Aerosol TRacker of mIXing state): an aerosol microphysical module for global atmospheric models, *Atmos. Chem. Phys.*, 8, 6003–6035, doi:10.5194/acp-8-6003-2008, 2008. 19431

5 Birch, C. E., Brooks, I. M., Tjernström, M., Shupe, M. D., Mauritsen, T., Sedlar, J., Lock, A. P., Earnshaw, P., Persson, P. O. G., Milton, S. F., and Leck, C.: Modelling atmospheric structure, cloud and their response to CCN in the central Arctic: ASCOS case studies, *Atmos. Chem. Phys.*, 12, 3419–3435, doi:10.5194/acp-12-3419-2012, 2012. 19427

10 Boé, J., Hall, A., and Qu, X.: Current GCMs' unrealistic negative feedback in the Arctic, *J. Climate*, 22, 4682–4695, 2009. 19423

15 Boyle, J., Williamson, D., Cederwall, R., Fiorino, M., Hnilo, J., Olson, J., Phillips, T., Potter, G., and Xie, S.: Diagnosis of Community Atmosphere Model 2 (CAM2) in numerical weather forecast configuration at atmospheric radiation measurement sites, *J. Geophys. Res.*, 110, D15S15, doi:10.1029/2004JD005042, 2005. 19429

20 Bretherton, C. and Park, S.: A new moist turbulence parameterization in the Community Atmosphere Model, *J. Climate*, 22, 3422–3448, 2009.

25 Brodeau, L., Barnier, B., Treguier, A.-M., Penduff, T., and Gulev, S.: An ERA40-based atmospheric forcing for global ocean circulation models, *Ocean Model.*, 31, 88–104, doi:10.1016/j.ocemod.2009.10.005, 2010. 19450

30 Chylek, P., Folland, C., Lesins, G., Dubey, M., and Wang, M.: Arctic air temperature change amplification and the Atlantic multidecadal oscillation, *Geophys. Res. Lett.*, 36, L14801, doi:10.1029/2009GL038777, 2009. 19423

35 Compo, G., Whitaker, J., Sardeshmukh, P., Matsui, N., Allan, R., Yin, X., Gleason, B., Vose, R., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R., Grant, A., Groisman, P., Jones, P., Kruk, M., Kruger, A., Marshall, G., Maugeri, M., Mok, H., Nordli, O., Ross, T., Trigo, R., Wang, X., Woodruff, S., and Worley, S.: The twentieth century reanalysis project, *Quart. J. Roy. Meteor. Soc.*, 137, 1–28, 2011. 19424

40 Curry, J., Schramm, J., and Ebert, E.: Sea ice-albedo climate feedback mechanism, *J. Climate*, 8, 240–247, 1995. 19422

45 Curry, J., Schramm, J., Alam, A., Reeder, R., Arbetter, T., and Guest, P.: Evaluation of data sets used to force sea ice models in the Arctic Ocean, *J. Geophys. Res.*, 107, C10, doi:10.1029/2000JC000466, 2002. 19450

## ASCOS model evaluation

G. de Boer et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

de Boer, G., Chapman, W., Kay, J., Medeiros, B., Shupe, M., Vavrus, S., and Walsh, J.: A characterization of the present-day Arctic atmosphere in CCSM4, *J. Climate*, 25, 2676–2695, 2012. 19425, 19450

de Boer, G., Bauer, S., Toto, T., Menon, S., and Vogelman, A.: Evaluation of aerosol–cloud interaction in the GISS ModelE using ARM observations, *J. Geophys. Res.*, doi:10.1002/jgrd.50460, 2013. 19431

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A., Van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hersbach, H., Hólm, E., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A., Monge-Sanz, B., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Quart. J. Roy. Meteor. Soc.*, 137, 553–597, 2011. 19424, 19427

Ebert, E. and Curry, J.: An intermediate one-dimensional thermodynamic sea ice model for investigating ice–atmosphere interactions, *J. Geophys. Res.*, 98, 10085–10109, 1993. 19440

Gascard, J.-C., Festy, J., LeGoff, H., Weber, M., Bruemmer, B., Offermann, M., Doble, M., Wadhams, P., Forsberg, R., Hanson, S., Skourup, H., Gerland, S., Nicolaus, M., Metaxian, J.-P., Grangeon, J., Haapala, J., Rinne, E., Haas, C., Heygster, G., Jakobson, E., Palo, T., Wilkinson, J., Kaleschke, L., Claffey, K., Elder, B., and Bottenheim, J.: Exploring Arctic transpolar drift during dramatic sea ice retreat, *Trans. Amer. Geophys. Union*, 89, 21–28, 2008. 19424

Graversen, R., Mauritsen, T., Tjernström, M., Källén, E., and Svensson, G.: Vertical structure of recent Arctic warming, *Nature*, 541, 53–56, 2008. 19423

Hansen, J., Sato, M., Ruedy, R., Kharecha, P., Lacis, A., Miller, R., Nazarenko, L., Lo, K., Schmidt, G., Russel, G., Aleinov, I., Bauer, S., Baum, E., Cairns, B., Canuto, V., Chandler, M., Cheng, Y., Cohen, A., Del Genio, A., Faluvegi, G., Fleming, E., Friend, A., Hall, T., Jackman, C., Jonas, J., Kelley, M., Kiang, N., Koch, D., Labow, G., Lerner, J., Menon, S., Novakov, T., Oinas, V., Perlitz, J., Perlitz, J., Rind, D., Romanou, A., Schmunk, R., Shindell, D., Stone, P., Sun, S., Streets, D., Tausnev, N., Thresher, D., Unger, N., Yao, M., and Zhang, S.: Climate simulations for 1880–2003 with GISS modelE, *Clim. Dynam.*, 29, 661–696, 2007. 19431

Iacono, M., Delamere, J., Mlawer, E., Shephard, M., Clough, S., and Collins, W.: Radiative forcing by long-lived greenhouse gases: calculations with the AER radiative transfer models, *J. Geophys. Res.*, 113, doi:10.1029/2008JD009944, D13103, 2008.

Inoue, J., Liu, J., Pinto, J., and Curry, J.: Intercomparison of Arctic regional climate models: modeling clouds and radiation for SHEBA in May 1998, *J. Climate*, 19, 4167–4178, 2006. 19425

IPCC: Climate Change 2007: The Physical Science Basis, Cambridge University Press, 2007. 5 19422

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, doi:10.1029/2012GL051591, L10802, 2012. 19424

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., 10 White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The NCEP/NCAR 40 yr reanalysis project, *B. Am. Meteorol. Soc.*, 77, 437–471, 1996. 19423, 19428

Kanamitsu, M., Ebisuzaki, W., Woollen, J., S.-K., Y., Hnilo, J., Fiorino, M., and Potter, G.: NCEP- 15 DOE AMIP-II reanalysis (R-2), *B. Am. Meteorol. Soc.*, 83, 1631–1643, 2002. 19424, 19428, 19440

Kay, J. and Gettelman, A.: Cloud influence on and response to seasonal Arctic sea ice loss, *J. Geophys. Res.*, 114, D18204, doi:10.1029/2009JD011773, 2009. 19422

Kazutoshi, K., Tsutsui, J., Koide, H., Sakamoto, M., Kobayashi, S., Hatsushika, H., 20 Matsumoto, T., Yamazaki, N., Kamahori, H., Takahashi, K., Kadokura, S., Wada, K., Kato, K., Oyama, R., Ose, T., Mannoji, N., and Ryusuke, T.: The JRA-25 reanalysis, *J. Meteor. Soc. Japan*, 85, 369–432, 2007. 19423

Lamarque, J.-F., Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., Lee, D., Liousse, C., 25 Mieville, A., Owen, B., Schultz, M. G., Shindell, D., Smith, S. J., Stehfest, E., Van Aardenne, J., Cooper, O. R., Kainuma, M., Mahowald, N., McConnell, J. R., Naik, V., Riahi, K., and van Vuuren, D. P.: Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application, *Atmos. Chem. Phys.*, 10, 7017–7039, doi:10.5194/acp-10-7017-2010, 2010. 19431

Liu, J., Zhang, Z., Hu, Y., Chen, L., Dai, Y., and Ren, X.: Assessment of surface air temperature 30 over the Arctic Ocean in reanalysis and IPCC AR4 model simulations with IABP/POLES observations, *J. Geophys. Res.*, 113, doi:10.1029/2007JD009380, D10105, 2007. 19425

Liu, X., Easter, R. C., Ghan, S. J., Zaveri, R., Rasch, P., Shi, X., Lamarque, J.-F., Gettelman, A., Morrison, H., Vitt, F., Conley, A., Park, S., Neale, R., Hannay, C., Ekman, A. M. L., 35

## ASCOS model evaluation

G. de Boer et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## ASCOS model evaluation

G. de Boer et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

## ASCOS model evaluation

G. de Boer et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[|◀](#)[▶|](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

Gu, W., Joiner, J., Koster, R., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C., Reichle, R., Robertson, F., Ruddick, A., Sienkiewicz, M., and Woollen, J.: MERRA: NASA's modern-era retrospective analysis for research and applications, *J. Climate*, 24, 3624–3648, 2011. 19424

5 Rigor, I., Colony, R., and Martin, S.: Variations in surface air temperature observations in the Arctic, 1979–97, *J. Climate*, 13, 896–914, 2000. 19422

10 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.-M. H., 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., Chemm, 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, R., Rutledge, G., and Goldberg, M.: The NCEP climate forecast system reanalysis, *B. Am. Meteorol. Soc.*, 91, 1015–1057, 2010. 19424

15 Schmidt, G., Ruedy, R., Hansen, J., Aleinov, I., Bell, N., Bauer, M., Bauer, S., Cairns, B., Canuto, V., Cheng, Y., Del Genio, A., Faluvegi, G., Friend, A., Hall, T., Hu, Y., Kelley, M., Kiang, N., Koch, D., Lacis, A., Lerner, J., Lo, K.-W., Miller, R., Nazarenko, L., Oinas, V., Perlitz, J., Perlitz, J., Rind, D., Romanou, A., Russell, G., Sato, M., Shindell, D., Stone, P., Sun, S., Tausnev, N., Thresher, D., and Yao, M.-S.: Present day atmospheric simulations using GISS ModelE: comparison to in-situ, satellite and reanalysis data, *J. Climate*, 19, 153–192, 2006. 19431, 19450

20 Schmidt, G. A., Kelley, M., Nazarenko, L., Ruedy, R., Russell, G. L., Aleinov, I., Bauer, M., Bauer, S., Bhat, M. K., Bleck, R., Canuto, V., Chen, Y., Cheng, Y., Clune, T. L., DelGenio, A., de Fainchtein, R., Faluvegi, G., Hansen, J. E., Healy, R. J., Kiang, N. Y., Koch, D., Lacis, A. A., LeGrande, A. N., Lerner, J., Lo, K. K., Marshall, J., Mathews, E. E., Menon, S., Miller, R. L., Oinas, V., Olsos, A., Perlitz, J., Puma, M. J., Putman, W. M., Rind, D., Romanou, A., Sato, M., Shindell, D. T., Sun, S., Syed, R., Tausnev, N., Tsigaridis, K., Unger, N., Voulgarakis, A., Yao, M.-S., and Zhang, J.: Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive, *J. Climate*, in preparation, 2013. 19431

25 Screen, J. and Simmonds, I.: The central role of diminishing sea ice in recent Arctic temperature amplification, *Nature*, 464, 1334–1337, 2010. 19422

30 Sedlar, J., Tjernström, M., Mauritzen, T., Shupe, M., Brook, I., Persson, P., Birch, C., Leck, C., Sirevaag, A., and Nicolaus, M.: A transitioning Arctic surface energy budget: the impacts of

solar zenith angle, surface albedo and cloud radiative forcing, *Clim. Dynam.*, 37, 1643–1660, doi:10.1007/s00382-010-0937-5, 2011. 19426

Serreze, M. and Francis, J.: The Arctic amplification debate, *Climatic Change*, 76, 241–264, 2006. 19422

5 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. 19422

Shupe, M.: A ground-based multisensor cloud phase classifier, *Geophys. Res. Lett.*, 34, L22809, doi:10.1029/2007GL031008, 2007. 19426

10 Shupe, M. and Intrieri, J.: Cloud radiative forcing of the Arctic surface: the influence of cloud properties, surface albedo, and solar zenith angle, *J. Climate*, 17, 616–628, 2004. 19443

Shupe, M., Uttal, T., and Matrosov, S.: Arctic Cloud Microphysics Retrievals from Surface-Based Remote Sensors at SHEBAs, *J. Appl. Meteor.*, 44, 1544–1562, 2005.

15 Shupe, M., Matrosov, S., and Uttal, T.: Arctic Mixed-Phase Cloud Properties Derived from Surface-Based Sensors at SHEBA, *J. Atmos. Sci.*, 63, 697–711, 2006.

Svensson, G. and Karlsson, J.: On the Arctic wintertime climate in global climate models, *J. Climate*, 24, 5757–5771, 2011. 19424

Tjernström, M., Sedlar, J., and Shupe, M.: How well do regional climate models reproduce radiation and clouds in the Arctic?, *J. Appl. Meteor. Climatol.*, 47, 2405–2422, 2008. 19425

20 Tjernström, M., Leck, C., Birch, C., Bottenheim, J., Brooks, B., Brook, I., 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., Mauritzen, T., Müller, M., Norris, S., Orellana, M., Orsini, D., Paatero, J., Persson, P., Gao, Q., Rauschenberg, C., Ristovski, Z., Sedlar, J., Shupe, M., Sierau, B., Sirevaag, A., Sjogren, S., Stetzer, O., Swietlicki, E., Szczodrak, M., Vaattovaara, P., Wahlberg, N., Westberg, M., and Wheeler, C.: 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. 19425, 19426

25 Uppala, S., Källberg, P., Simmons, A., Andreae, U., Da Costa Betchold, V., Fiorino, M., Gibson, J., Haseler, J., Hernandez, A., Kelley, G., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R., Andersson, E., Arpe, K., Balmaseda, M., Beljaars, A., Van de Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B., Isaksen, L., Janssen, P., Jenne, R., McNally, A., Mah-

## ASCOS model evaluation

G. de Boer et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

fouf, J.-F., Morcrette, J.-J., Rayner, N., Saunders, R., Simon, P., Sterl, A., Trenberth, K., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 re-analysis, *Quart. J. Roy. Meteor. Soc.*, 131, 2961–3012, 2005. 19423, 19427

5 Vihma, T., Jaagus, J., Jakobson, E., and Palo, T.: Meteorological conditions in the Arctic Ocean in spring and summer 2007 as recorded on the drifting ice station Tara, *Geophys. Res. Lett.*, 35, L18706, doi:10.1029/2008GL034681, 2008. 19424

Walsh, J., Kattsov, V., Chapman, W., Govokova, V., and Pavlova, T.: Comparison of Arctic climate simulations by uncoupled and coupled global models, *J. Climate*, 15, 1429–1446, 2002. 19424

10 Walsh, J., Chapman, W., and Portis, D.: Arctic cloud fraction and radiative fluxes in atmospheric reanalyses, *J. Climate*, 22, 2316–2334, 2008. 19423

Westwater, E., Han, Y., Shupe, M., and Matrosov, S.: Analysis of integrated cloud liquid and precipitable water vapor retrievals from microwave radiometers during the surface heat budget of the Arctic Ocean project, *J. Geophys. Res.*, 106, 32019–23030, 2001. 19426, 19427

1075 Zib, B., Dong, X., Xi, B., and Kennedy, A.: Evaluation and intercomparison of cloud fraction and radiative fluxes in recent reanalyses over the Arctic using BSRN surface observations, *J. Climate*, 25, 2291–2305, 2012. 19424

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[|◀](#)

[▶|](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



**Table 1.** Median biases and correlation (in parentheses) to observations for each of the evaluated variables. *Italicized* correlations are not statistically significant at the 95 % level. The best performing statistics for each variable are included in **bold**.

Variable	ERA-I	R-1	R-2	CAM5	CAM5PF	GISS ModelE2
$T_{\text{sfc}}$	0.65 ( <b>0.82</b> )	1.15 (0.62)	<b>-0.02</b> (0.54)	-3.39 (0.59)	-0.90 (0.37)	0.82 (0.58)
$P_{\text{sfc}}$	-0.71 (0.99)	-3.01 ( <b>1.00</b> )	4.93 (0.99)	<b>-0.04</b> ( <b>1.00</b> )	-0.90 (0.96)	-0.55 (0.99)
$RH_{\text{sfc}}$	2.39 ( <b>0.65</b> )	-4.39 (0.41)	-2.99 ( <i>0.06</i> )	<b>-1.15</b> (0.27)	-1.17 ( <i>-0.15</i> )	-7.80 (0.10)
$U_{\text{sfc}}$	-0.36 ( <b>0.92</b> )	-0.65 (0.91)	-0.48 (0.83)	-0.55 (0.84)	0.22 (0.41)	<b>-0.07</b> (0.83)
$V_{\text{sfc}}$	0.20 ( <b>0.91</b> )	<b>0.09</b> (0.88)	1.12 (0.77)	1.20 (0.82)	2.13 (0.50)	1.78 (0.76)
CF	<b>0.00</b> ( <i>-0.04</i> )	-58.30 ( <b>0.28</b> )	-47.00 ( <b>0.28</b> )	-12.88 ( <i>-0.04</i> )	-13.38 (0.01)	<b>0.00</b> ( <i>-0.08</i> )
LWP	-7.50 (0.55)	N/A	N/A	-63.56 ( <b>0.69</b> )	-37.18 (0.21)	<b>2.87</b> (0.42)
Precip	0.005 ( <b>0.52</b> )	-0.006 ( <i>-0.01</i> )	-0.005 (0.05)	<b>0.003</b> (0.35)	0.005 ( <i>-0.05</i> )	0.015 (0.28)
$F_{\text{sw, down}}$	-22.73 (0.39)	89.10 (0.77)	55.14 ( <b>0.78</b> )	69.20 (0.71)	27.75 (0.70)	<b>3.97</b> (0.37)
$F_{\text{sw, up}}$	28.53 (0.41)	-40.66 (0.74)	-52.24 ( <b>0.77</b> )	-59.81 (0.69)	-25.05 (0.67)	<b>19.57</b> (0.27))
$F_{\text{sw, net}}$	7.11 (0.31)	48.08 (0.62)	<b>5.33</b> ( <b>0.70</b> )	9.21 (0.74)	9.47 (0.62)	30.26 (0.52)
$F_{\text{lw, down}}$	<b>1.80</b> (0.35)	-59.32 (0.48)	-45.79 ( <b>0.52</b> )	-62.60 (0.51)	-19.76 (0.25)	-19.41 (0.15)
$F_{\text{lw, up}}$	-2.69 (0.35)	8.01 (0.27)	9.17 (0.51)	30.06 ( <b>0.69</b> )	6.52 (0.24)	<b>0.70</b> (0.49)
$F_{\text{lw, net}}$	<b>-1.43</b> (0.25)	-54.58 (0.38)	-31.07 ( <b>0.39</b> )	-34.12 (0.26)	-15.50 (0.23)	-19.30 (0.00)
$F_{\text{SH}}$	<b>-0.85</b> ( <i>-0.01</i> )	6.38 ( <i>-0.08</i> )	17.84 ( <b>0.11</b> )	5.78 ( <i>-0.06</i> )	6.37 ( <i>-0.10</i> )	2.60 ( <i>-0.03</i> )
$F_{\text{LH}}$	2.21 (0.25)	-7.47 (0.35)	-0.47 (0.28)	<b>-0.37</b> ( <b>0.37</b> )	-1.61 (0.09)	6.28 ( <i>-0.13</i> )
$Q_{\text{sfc}}$	7.24 (0.22)	<b>2.97</b> ( <b>0.48</b> )	-11.68 (0.42)	-18.62 ( <b>0.48</b> )	-7.57 (0.36)	22.38 (0.42)

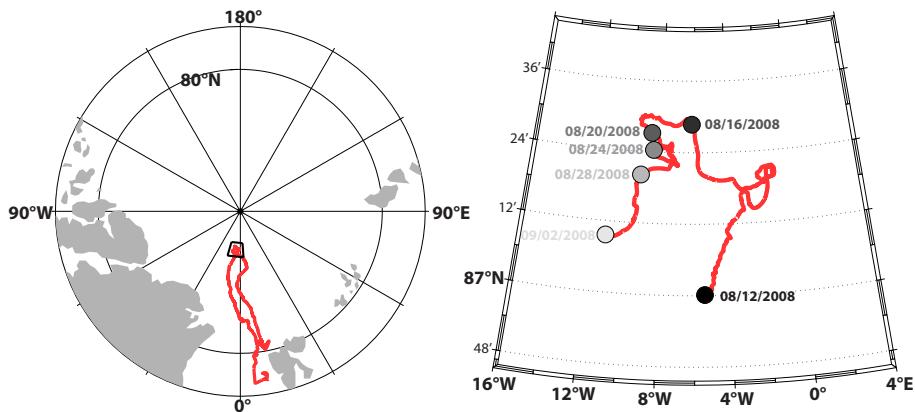
Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures
   	   
Back	Close
Full Screen / Esc	

Printer-friendly Version
Interactive Discussion



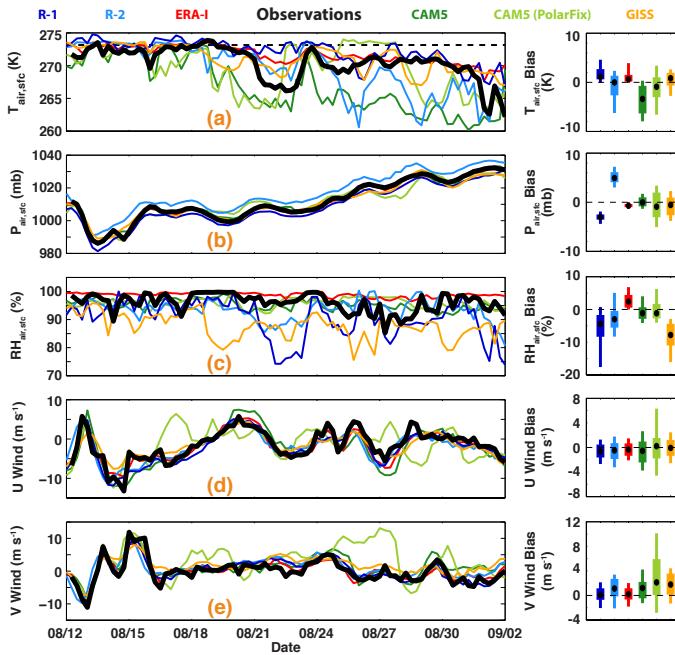
## ASCOS model evaluation

G. de Boer et al.

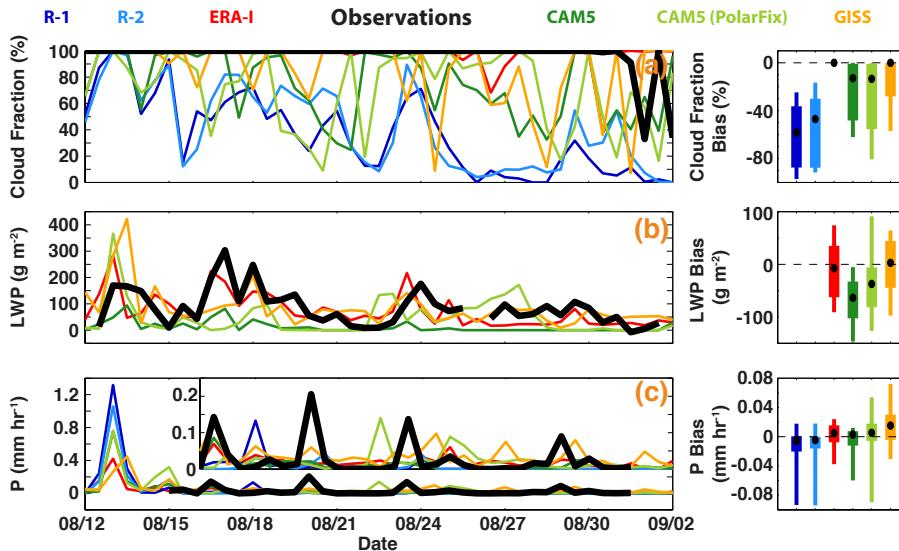


**Fig. 1.** A map illustrating the path of *Oden* (red line) during ASCOS. The region of ice drift (black box) is enlarged, and greyscale points provide the location of the ship on given dates.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[|◀](#)[▶|](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)



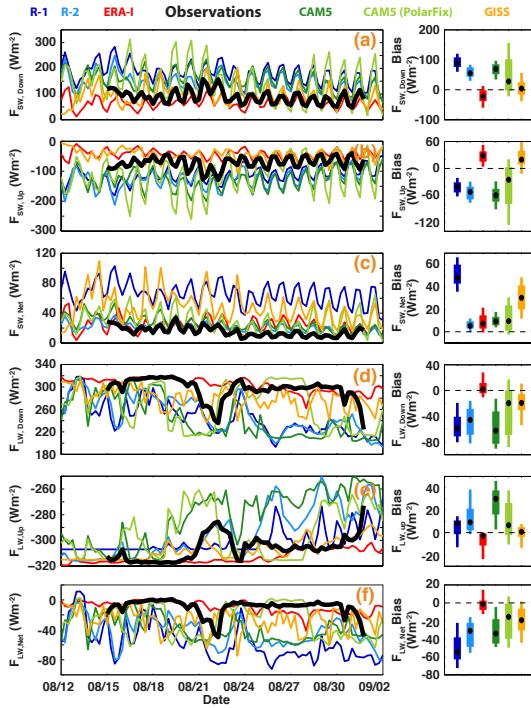
**Fig. 2.** Time series of basic meteorological quantities (from top to bottom: surface air temperature, surface air pressure, surface air relative humidity, 10 m  $U$  wind component, and 10 m  $V$  wind component. Included are lines for the observations from *Oden* (bold black line), CAM5 (standard version, dark green; CAM5-PF, light green), the GISS-ModelE2 (orange), and R-1 (dark blue), R-2 (light blue) and ERA-I (red) reanalyses. ERA-I, R-1 and R-2 lines represent 6 hourly analysis (0 h) time instantaneous values, while the CAM5 lines depicts the an interpolated value for the analysis time from the model 1 hourly averages. The right-hand side of the figure includes distributions depicting the median (black dot), IQR (wide bar) and 10–90th percentiles (thin bar) of the differences between observed and simulated values.



**Fig. 3.** Time series of common cloud and precipitation quantities (from top to bottom: average cloud fraction, average liquid water path, and average precipitation rate). Included are lines for the observations from *Oden* (bold black line), CAM5 (standard version, dark green; CAM5-PF, light green), the GISS-ModelE2 (orange), and R-1 (dark blue), R-2 (light blue) and ERA-I (red) reanalyses. The inserts in (b) and (c) modify the scale of the vertical axis (note log scale for (b)) to better distinguish between models where all have low values. All lines represent six-hour average values. Please note that R-1 and R-2 do not provide liquid water path, and are therefore not included in the center figure. The right-hand side of the figure includes distributions depicting the median (black dot), IQR (wide bar) and 10–90th percentiles (thin bar) of the differences between observed and simulated values.

## Discussion Paper | ASCOS model evaluation

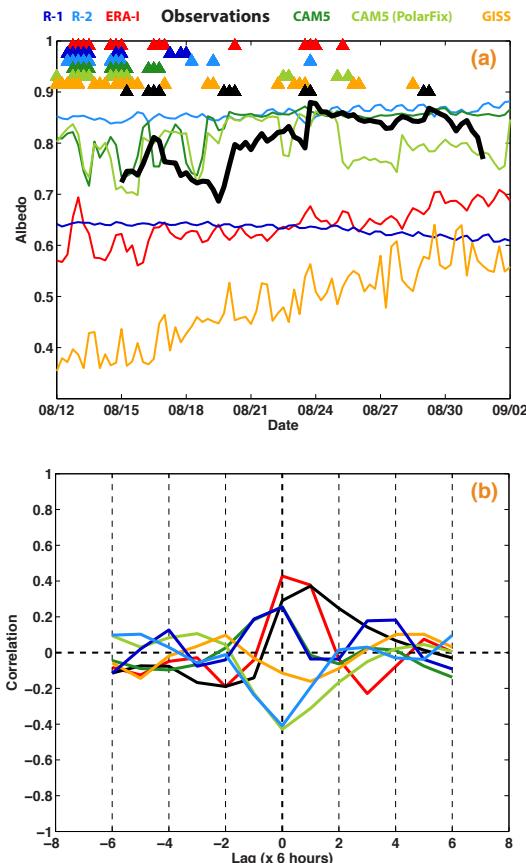
G. de Boer et al.



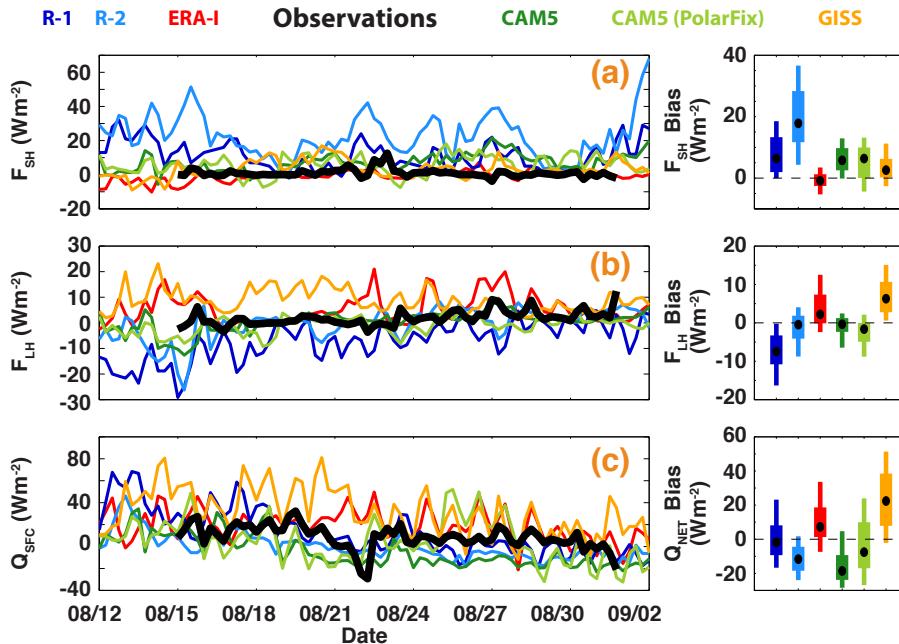
**Fig. 4.** Time series of surface energy budget terms (from top to bottom: average downwelling shortwave flux density, average upwelling shortwave flux density, average net shortwave flux density, average downwelling longwave flux density, average upwelling longwave flux density, average net longwave flux density). Included are lines for the observations from *Oden* (bold black line), CAM5 (standard version, dark green; CAM5-PF, light green), the GISS-ModelE2 (orange), and R-1 (dark blue), R-2 (light blue) and ERA-I (red) reanalyses. All lines represent six-hour average values. The right-hand side of the figure includes distributions depicting the median (black dot), IQR (wide bar) and 10–90th percentiles (thin bar) of the differences between observed and simulated values.

<a href="#">Title Page</a>	<a href="#">Abstract</a>	<a href="#">Introduction</a>
<a href="#">Conclusions</a>	<a href="#">References</a>	
<a href="#">Tables</a>	<a href="#">Figures</a>	
<a href="#">◀</a>	<a href="#">▶</a>	
<a href="#">◀</a>	<a href="#">▶</a>	
<a href="#">Back</a>	<a href="#">Close</a>	
<a href="#">Full Screen / Esc</a>		

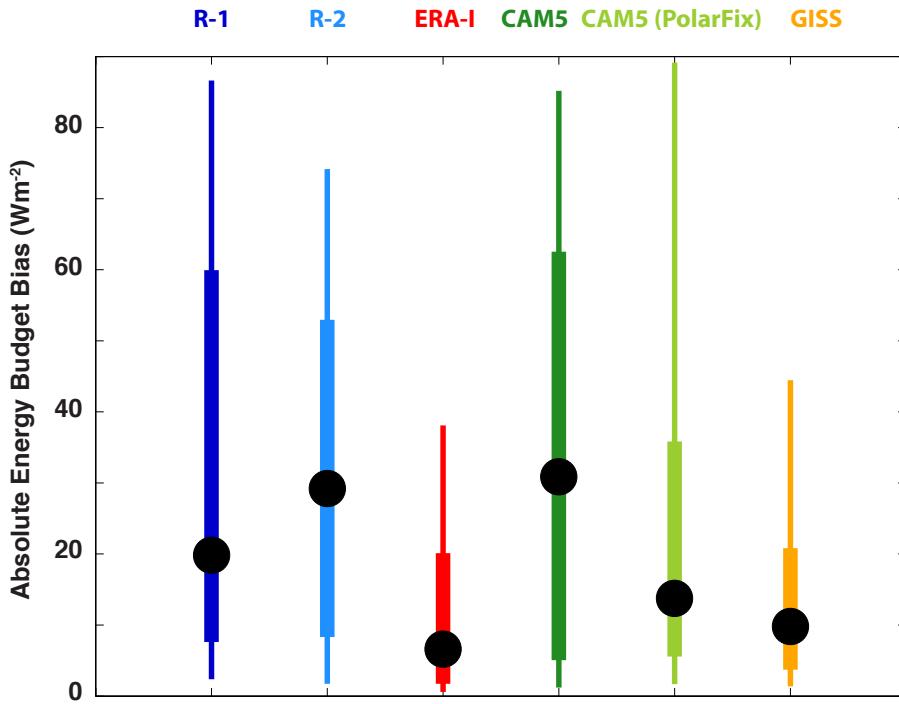
<a href="#">Printer-friendly Version</a>
<a href="#">Interactive Discussion</a>



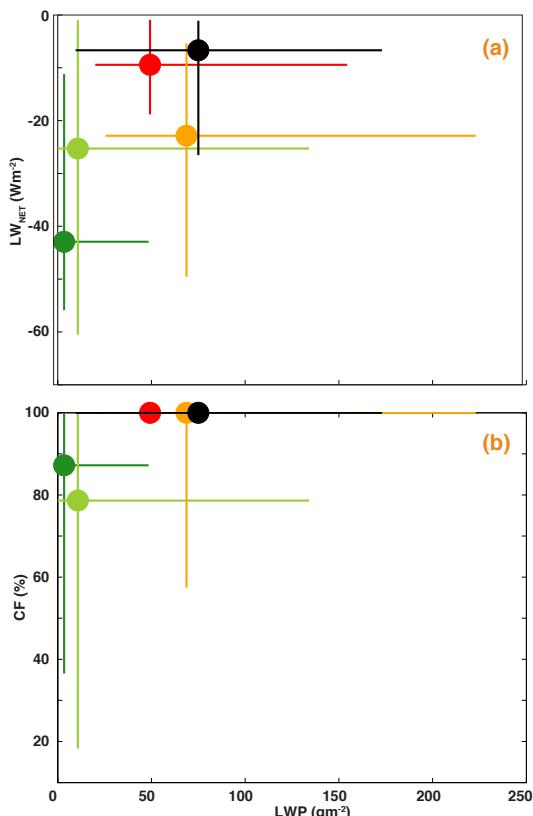
**Fig. 5.** Time series of surface albedo as calculated from the 6 h average shortwave radiation terms. Included are lines for the observations from *Oden* (bold black line), CAM5 (standard version, dark green; CAM5-PF, light green), the GISS-ModelE2 (orange), and R-1 (dark blue), R-2 (light blue) and ERA-I (red) reanalyses. Triangles along the top of the figure represent 6 hourly periods that averaged precipitation of  $0.5 \text{ mm hr}^{-1}$  or greater.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

**Fig. 6.** Time series of surface energy budget terms (from top to bottom: average sensible heat flux density, average latent heat flux density, and net surface energy (radiation, latent heat and sensible heat)). Included are lines for the observations from *Oden* (bold black line), CAM5 (standard version, dark green; CAM5-PF, light green), the GISS-ModelE2 (orange), and R-1 (dark blue), R-2 (light blue) and ERA-I (red) reanalyses. All lines represent six-hour average values. The right-hand side of the figure includes distributions depicting the median (black dot), IQR (wide bar) and 10th-90th percentiles (thin bar) of the differences between observed and simulated values.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

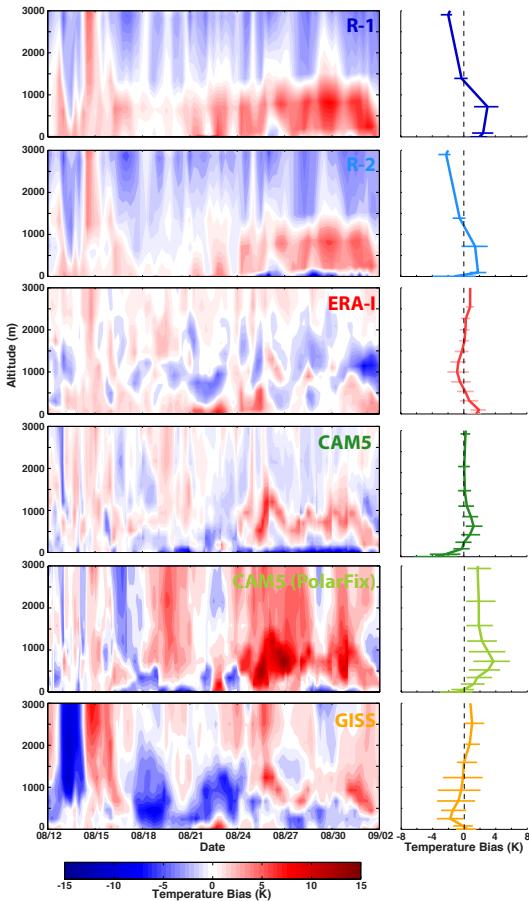
**Fig. 7.** Combined distributions of the absolute values of biases in  $LW_{up}$ ,  $LW_{down}$ ,  $SW_{up}$ ,  $SW_{down}$ ,  $F_{SH}$  and  $F_{LH}$ . Distributions are depicted including the median (black dot), IQR (wide bar) and 10th-90th percentiles (thin bar).



**Fig. 8.** Figures demonstrating the range of values of liquid water path for observations and models that provide it, as well as its relationship to surface net longwave radiation (**a**), and cloud fraction (**b**). The circles represent the mean value over the ASCOS observational period, and the whiskers demonstrate the extent of the 10th and 90th percentiles of the datasets.

## Discussion Paper | ASCOS model evaluation

G. de Boer et al.



**Fig. 9.** Time-height cross sections of air temperature bias (model-observations, left). Profiles on the right side illustrate the mean temperature bias with height (bold line) and the interquartile range of the bias (lighter, horizontal lines).

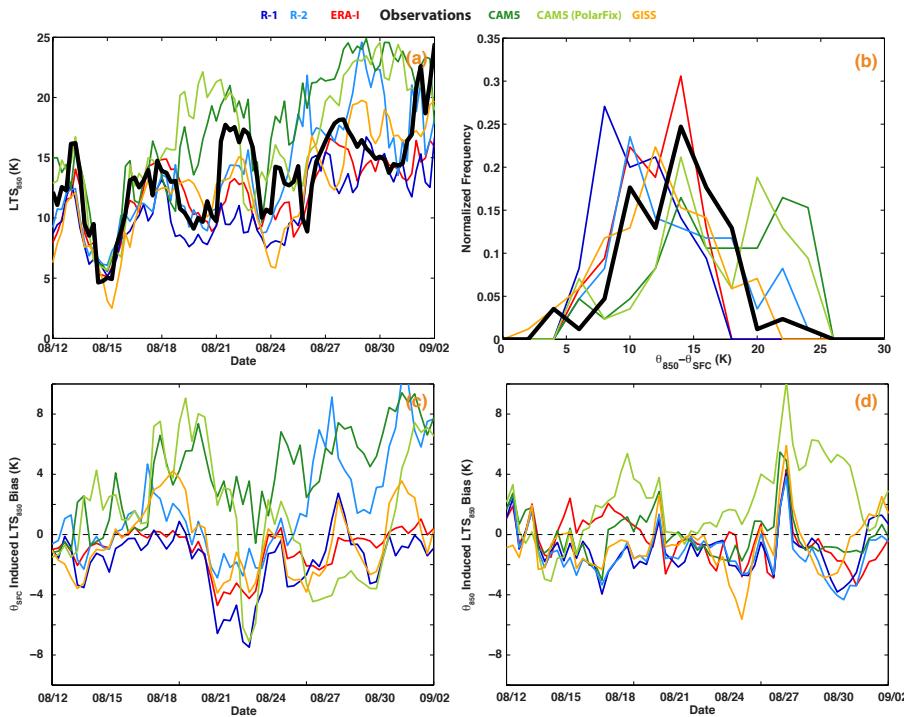
[Title Page](#) [Abstract](#) [Introduction](#)  
[Conclusions](#) [References](#)  
[Tables](#) [Figures](#)

[◀](#) [▶](#)  
[◀](#) [▶](#)  
[Back](#) [Close](#)  
[Full Screen / Esc](#)

[Printer-friendly Version](#)  
[Interactive Discussion](#)

## Discussion Paper | ASCOS model evaluation

G. de Boer et al.



**Fig. 10.** Timeseries (a) and histograms (b) of lower tropospheric stability (LTS<sub>850</sub>), as defined by the potential temperature difference between 850 mb and the surface. The lower panels (c, d) illustrate the relative contributions of near-surface potential temperature and the potential temperature at 850 mb to biases in LTS<sub>850</sub>.

- Discussion Paper | Title Page
- Abstract | Introduction
- Conclusions | References
- Tables | Figures
- | Back | Close
- Full Screen / Esc
- Printer-friendly Version
- Interactive Discussion