2013

Validation of the Diurnal Cycles in Atmospheric Reanalyses Over Antarctic Sea Ice

Esa-Matti Tastula
University of South Florida

Timo Vihma
Finnish Meteorological Institute

Edgar L. Andreas
NorthWest Research Associates, Lebanon, New Hampshire

Boris Galperin
University of South Florida, bgalperin@usf.edu

Validation of the diurnal cycles in atmospheric reanalyses over Antarctic sea ice

Esa-Matti Tastula,[1] Timo Vihma,[2] Edgar L. Andreas,[3] and Boris Galperin[1]

Received 24 October 2012; revised 12 March 2013; accepted 14 March 2013; published 28 May 2013.

[1] The diurnal cycles of near-surface meteorological parameters over Antarctic sea ice in six widely used atmospheric reanalyses are validated against observations from Ice Station Weddell. The station drifted from February through May 1992 and provided the most extensive set of meteorological observations ever collected in the Antarctic sea ice zone. For the radiative and turbulent surface fluxes, both the amplitude and shape of the diurnal cycles vary considerably among different reanalyses. Near-surface temperature, specific humidity, and wind speed in the reanalyses all feature small diurnal ranges, which, in most cases, fall within the uncertainties of the observed cycle. A skill score approach revealed the superiority of the ERA-Interim reanalysis in reproducing the observed diurnal cycles. An explanation for the shortcomings in the reanalyses is their failure to capture the diurnal cycle in cloud cover fraction, which leads to errors in other quantities as well. Apart from the diurnal cycles, NCEP-CFSR gave the best error statistics.

Citation: Tastula, E.-M., T. Vihma, E. L. Andreas, and B. Galperin (2013), Validation of the diurnal cycles in atmospheric reanalyses over Antarctic sea ice, *J. Geophys. Res. Atmos.*, *118*, 4194–4204, doi:10.1002/jgrd.50336.

1. Introduction

[2] Atmospheric reanalyses are broadly applied in Earth sciences. For example, the atmospheric forcing for ocean, sea ice, lake, and discharge hydrology models are often taken from reanalyses. Other applications include studies of the climate variability and trends as well as occurrence of extreme values. Reanalyses do, however, include errors [*Walsh and Chapman*, 1998; *Bromwich et al*., 2007; *Lüpkes et al*., 2010; *Bromwich et al*., 2011; *Screen and Simmonds*, 2011; *Bracegirdle and Marshall*, 2012; *Jakobson et al*., 2012]. To improve the situation, major numerical weather prediction (NWP) centers have been active in producing new reanalyses. These include the European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis (ERA-Interim)[*Dee et al*., 2011], the Japanese Meteorological Agency Reanalysis (JRA-25) [*Onogi et al*., 2007], the National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR) [*Saha et al*., 2010], and the National Aeronautics and Space and Administration (NASA) Modern Era Retrospective-Analysis for Research and Applications (MERRA) [*Cullather and Bosilovich*, 2011].

[3] The diurnal cycle in surface and boundary-layer variables is important for the global climate system [*Dai and Trenberth*, 2004], weather forecasting [*Atlaskin and Vihma*, 2012], dispersion and transport of pollutants [*Panday et al*., 2009], and production of wind energy [*He et al*., 2012]. The importance of accurately modeling the diurnal cycle is also well recognized by the NWP and climate modeling communities, which have carried out extensive model intercomparisons, in particular, under the Global Energy and Water Cycle Experiment of the World Climate Research Programme [*Svensson et al*., 2011]. In the previous validation studies of reanalyses, however, the diurnal cycle has not received much attention; the validation results have been mostly presented as annual, seasonal, monthly, or daily error statistics.

[4] Modeling the diurnal cycle of the Earth's surface temperature and the near-surface air temperature, humidity, and wind speed is a major challenge. Stratification in the atmospheric boundary layer (ABL) typically varies between daytime convective and nighttime stably stratified conditions. Hence, the model should be able to reproduce the morning and evening transitions in the state of the ABL. In addition to the transitions, modeling the stable boundary layer is problematic because of its shallowness and small heat capacity; the complex interaction between turbulence and gravity waves; and, in very stable conditions, the intermittency of the turbulence [*Mahrt*, 1999]. All these reasons reduce the validity of the Monin-Obukhov similarity theory [*Monin and Obukhov*, 1954] that NWP and climate models use as the basis for parameterizing the turbulent surface fluxes. Hence, it is not surprising that NWP and climate models are not particularly successful in reproducing the diurnal cycle of surface and near-surface variables [*Steeneveld et al*., 2008; *Svensson et al*., 2011].

[1]College of Marine Science, University of South Florida, St. Petersburg, Florida, USA.
[2]Finnish Meteorological Institute, Helsinki, Finland.
[3]NorthWest Research Associates, Lebanon, New Hampshire, USA.

Corresponding author: E.-M. Tastula, College of Marine Science, University of South Florida, 830 1st St. Southeast, St. Petersburg, FL 33701, USA. (tastulae@mail.usf.edu)

2169-897X/13/10.1002/jgrd.50336

[5] A particular challenge is to simulate the diurnal cycle at high latitudes, where the diurnal range of the solar zenith angle is small. Furthermore, the high surface albedo of snow and sea ice reduces the net shortwave radiation, the principle forcing behind the diurnal cycle. The variables directly affected are outgoing shortwave radiation and surface temperature. These further affect outgoing longwave radiation, surface albedo, and the turbulent fluxes of sensible and latent heat. The diurnal cycles of other meteorological variables in the ABL are mostly generated via the surface fluxes. Over sea ice, the heat flux from the ocean, via open leads and conduction through the ice, tends to keep the diurnal cycle of such variables small compared to that over snow-covered land [*Niros et al*., 2002].

[6] Detailed studies on the diurnal cycle over sea ice have been rare. At Ice Station Polarstern (ISPOL) over the Antarctic sea ice at 68°S in early summer, *Vihma et al*. [2009] observed diurnal cycles in the incoming and outgoing shortwave and longwave radiation, net radiation, surface temperature and albedo, turbulent fluxes of sensible and latent heat, air temperature, air specific and relative humidity, wind speed, and base height of low clouds. Other observations over Antarctic sea ice include those from *Wendler et al*. [2005] who mentioned a smaller amplitude for incoming shortwave radiation (due to a more southern latitude of 78°S) but a larger one in the incoming longwave radiation. Even very small variations in the solar zenith angle may generate diurnal cycles: over the Arctic sea ice in summer, as far north as 88–89°N, *Tjernström* [2005; 2007] observed diurnal cycles in incoming shortwave and net radiation, wind speed, cloud base height, and visibility.

[7] There is a crucial need for validating reanalyses over the Antarctic sea ice zone because the sea ice extent has increased during the latest decades, and the reasons for this are complex and not fully understood [*Stammerjohn et al*., 2012]. We are aware of only two papers on validating reanalyses over Antarctic sea ice, *Vihma et al*. [2002] and *Vancoppenolle et al*. [2011], and these studies addressed only the old NCEP-NCAR reanalysis [*Kalnay et al*., 1996]. In addition, *King* [2003] and *Vihma et al*. [2002] have validated the ECMWF operational analyses over Antarctic sea ice, and *Pavelsky et al*. [2011] have compared satellite-based temperature inversion climatology against ERA-40 and NCEP-NCAR reanalyses. Pavelsky *et al*. [2011] was, however, not a validation study, as they did not consider it clear whether reanalyses or satellite data are more reliable, and the years with satellite data did not match with the ERA-40 period.

[8] In the Antarctic sea ice zone, the most extensive data set on ABL structure and processes is that collected on the U.S.-Russian Ice Station Weddell (ISW), which drifted in the western Weddell Sea from February through May 1992 (Figure 1). ISW has provided, by far, the longest data record in the Antarctic sea ice zone, and much of the knowledge of the ABL over the Antarctic sea ice zone originates from ISW [*Andreas et al*., 2000, 2002, 2004, 2005; *Tastula et al*., 2012].

[9] In this paper, we present the first seasonal-scale validation of the second generation of reanalyses with respect to the surface fluxes over Antarctic sea ice. Our objectives are (a) to quantify the accuracy of the most widely used reanalyses in representing the diurnal cycle in the Antarctic sea ice zone, (b) to identify reasons for the errors, and (c) to find out which of the reanalyses performs best. We validate the NCEP-DOE, NCEP-CFSR, ERA-40, ERA-Interim, JRA-25, and NASA-MERRA reanalyses against observations on ISW. The somewhat older NCEP-DOE [*Kanamitsu et al*., 2002] and ERA-40 [*Uppala et al*., 2005] reanalyses are included in the study to quantify the progress made in recent years by NCEP-CFSR and ERA-Interim. We demonstrate large discrepancies among the six reanalyses, all of which deviate from observations, and discuss the potential reasons for the errors detected.

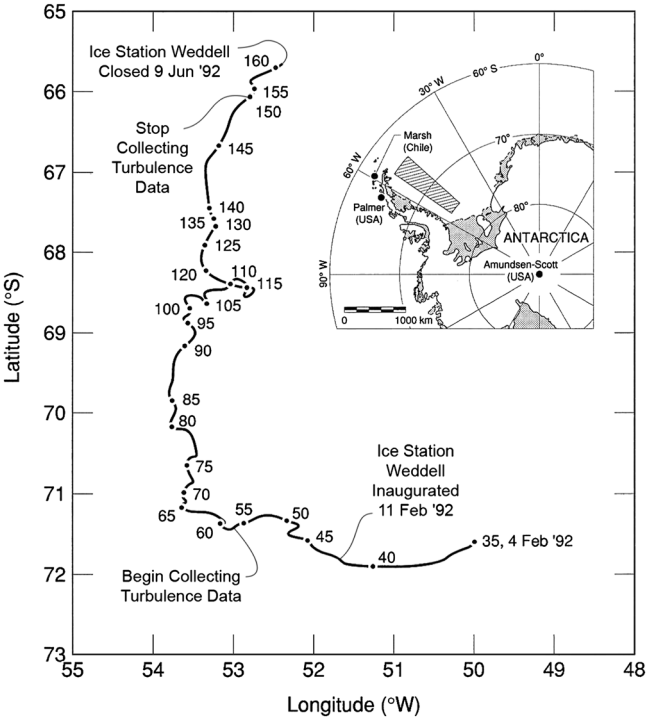

Figure 1. The drift track of Ice Station Weddell. The numbers indicate the Julian day in 1992 [*Andreas et al*., 2000].

2. Observations, Data Processing, and Uncertainties

[10] The ISW measurement period, which ran from 26 February 00 UTC to 28 May 18 UTC (Figure 1), provides hourly time series for incoming and outgoing longwave and shortwave radiation, sensible and latent heat fluxes, cloud cover fraction, near-surface temperature, specific humidity, and wind speed. In this paper, due to the 6-hour temporal resolution of the reanalyses, only every sixth measurement from ISW was used for the validation of the instantaneous reanalysis values. In ERA-Interim and ERA-40, however, the flux variables are not instantaneous but rather are averages over a 6-hour accumulation period. Hence, to obtain a proper comparison with observations, we built an accumulated observational dataset mimicking the ERA-Interim and ERA-40 flux data.

[11] The gaps in the observational data were filled using linear interpolation. The number of missing values was always less than 7% for all parameters except for sensible and latent heat fluxes. For these two parameters, roughly half of the points were missing and the diurnal cycles could not be determined.

[12] For the reanalyses, the value corresponding to the grid point closest to the Ice Station Weddell site was used. Because of the horizontally homogenous surface, no interpolation was necessary. For all the considered quantities, tests confirmed that the difference between the value at the nearest grid point and the value linearly interpolated to the observation site was insignificant.

[13] We calculated the average diurnal cycles by taking the average over all the 00 UTC, 06 UTC, 12 UTC, and 18 UTC values within the 93-day span. To get an idea how these cycles evolve in time, we also averaged cycles separately for days 1–31, 32–62, and 63–93.

[14] Based on the runs test [*Bendat and Piersol*, 1986], our data points are not independent. We therefore applied a Monte Carlo approach when determining uncertainties for the ranges of average diurnal cycles in the reanalyses and observations. To determine these uncertainties, we fitted a first-order autoregressive (AR1) model to the spectra of the residuals [*Press et al*., 1988, pp. 656–706]. A randomly generated 93-day red noise time series with the same variance as in the residuals and with a spectral slope given by the AR1 model was then combined with a time series of the same length, made of 93 duplicates of the average cycle. Next, the range of the average diurnal cycle for this new time series was calculated. This process was then repeated 10,000 times. The resulting histogram for the range of the (artificial) diurnal cycles is a quasi-Gaussian probability density function from which the uncertainties for the 95% confidence level were calculated directly.

[15] Because the diurnal cycle for incoming and outgoing shortwave radiation exhibits large changes during the period of interest, NASA Surface Meteorology and Solar Energy satellite data [*Chandler et al*. 2004], instead of the average cycles, were used as a model to calculate the residuals for these two quantities. This method allows us to estimate a sinusoid with a decreasing amplitude over the 93-day period. The accuracy of the satellite measurements is not the key point here; they were merely used as an approximation for the way the diurnal cycle changes on a seasonal scale. By subtracting the satellite-based shortwave time series instead of the average diurnal cycle from the observed time series, we eliminate the influence of the decreasing daily range on the spectra of the residuals.

3. Diurnal Cycle by Parameter

3.1. Limitations and Main Scientific Questions

[16] Because in most reanalyses the temporal resolution is 6-hourly (00, 06, 12, and 18 UTC), any implied diurnal cycle will be a crude approximation. Even more importantly, because the local solar time (LST) of the ISW observation site is 4 hours behind UTC, the method described above does not give the true amplitude of the diurnal variation because local noon falls between 12 and 18 UTC. Thus, the amplitude obtained does not represent the true amplitude of the diurnal cycle but rather is the amplitude of the cycle based on the available 6-hourly data. Keeping this in mind is important when interpreting the data. For instance, the incoming shortwave radiation in NASA-MERRA (Figure 2) appears to reach a maximum at 12 UTC (08 LST) even though this is just an indication that the 08 LST value in NASA-MERRA is higher than the values before and after it (02 and 14 LST).

[17] To begin with our analysis, we pose two questions: 1. At the 95% confidence level, does the reanalysis/observational dataset produce a diurnal cycle? 2. If there is a diurnal cycle, in the reanalysis, is it, at the same confidence level, different from the observed cycle? The answers are presented in Table 1 for the entire ISW period. The observations feature a diurnal cycle for all the parameters considered. JRA-25 is the only one of the reanalyses that produces this result. There is no diurnal cycle for incoming longwave radiation in ERA-Interim, ERA-40, and NASA-MERRA. The other problematic quantities are cloud cover fraction and near-surface wind speed. Considering all variables, however, when there is a diurnal cycle in the reanalysis data, the cycles are in a vast majority of cases not different from the observed ones. Only in three cases (ERA-40, cloud cover fraction; NASA-MERRA and NCEP-DOE, outgoing shortwave radiation) is the difference statistically significant at the 95% confidence level.

3.2. Incoming and Outgoing Shortwave Radiation

[18] Over the whole period, the incoming shortwave radiation is most problematic for JRA-25 and NCEP-DOE. The diurnal cycles are larger than those observed. The same problem persists separately for each of the three subperiods. For NCEP-CFSR and NASA-MERRA, days 32–93 is the problematic period; both reanalyses yield too small a cycle (Figure 3). Moreover, the shapes of the cycle given by NASA-MERRA, NCEP-DOE, and JRA-25 are erroneous. The first two yield values at 08 LST that are too high compared to values at 14 LST, and the last one underestimates the 08 LST value (Figure 2). Because the fluxes in ERA-Interim and ERA-40 are made of accumulated values, they cannot represent the observed instantaneous peak values. Therefore, the comparison is carried out against the accumulated observational dataset, which also yields a substantially better fit (Figure 2).

[19] With the exception of ERA-Interim and ERA-40, days 1–31 is a particularly difficult period for modeling outgoing shortwave radiation; JRA-25, NASA-MERRA, and NCEP-CFSR all underestimate the range of the diurnal cycle, whereas NCEP-DOE overestimates it (Figure 3). The agreement with the observed amplitude of the outgoing shortwave radiation in the last third of the ISW period is better, though JRA-25 and NCEP-DOE still overestimate the amplitude, whereas NCEP-CFSR and NASA-MERRA underestimate it. The problems concerning the shape of the cycle (involving NASA-MERRA, NCEP-DOE, and JRA-25) are identical to those observed with incoming shortwave radiation (Figure 2).

3.3. Incoming and Outgoing Longwave Radiation

[20] Even though three reanalyses (JRA-25, NCEP-CFSR, and NCEP-DOE) produce a statistically significant diurnal cycle for incoming longwave radiation, none of the reanalyses manage to capture the average shape of the cycle (Figure 2). The observations reach their minimum value at 14 LST, while the maximum occurs at 08 LST. Moreover, the observations yield a statistically significant cycle for days 1–31 (not shown), which is missing in ERA-Interim and NASA-MERRA. Outgoing longwave radiation turns out to be less problematic: there is an overall agreement

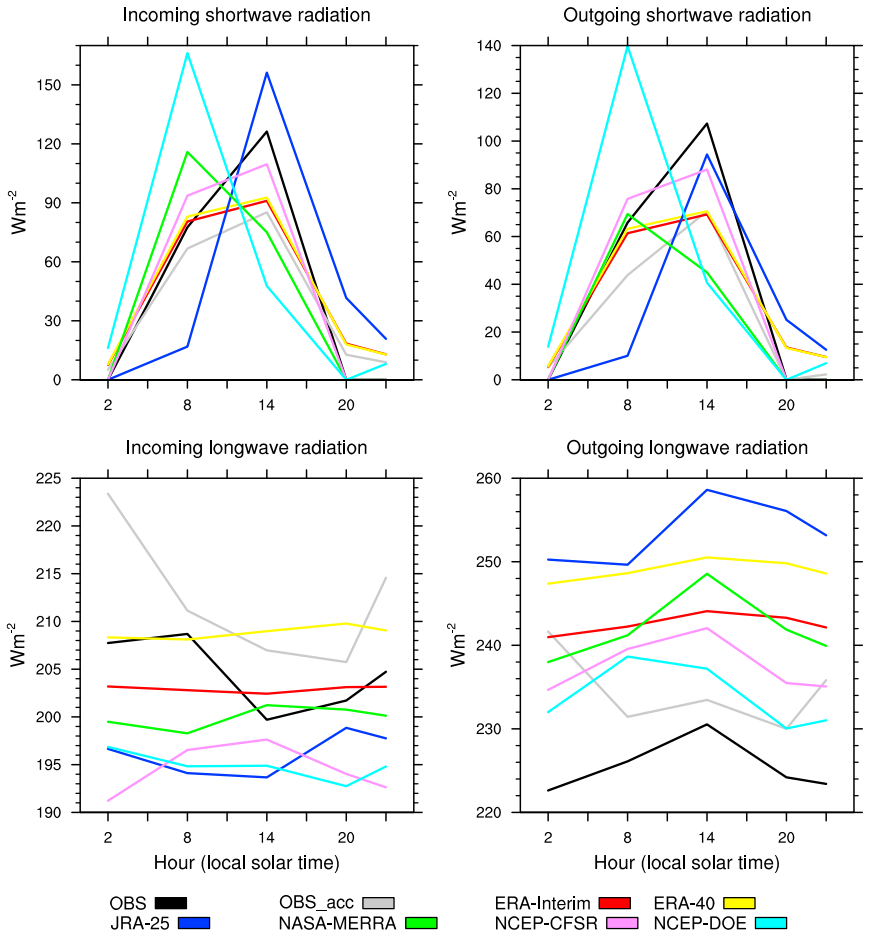

Figure 2. The diurnal cycles for the radiative fluxes averaged for the entire Ice Station Weddell period based on observations and reanalysis values at 02, 08, 14, and 20 LST values.

among the reanalyses on the shape of the diurnal cycle. NCEP-DOE is an outlier. It yields a maximum value for outgoing longwave radiation at 08 LST, whereas the observations and all other reanalyses give it at 14 LST.

3.4. Sensible and Latent Heat Fluxes

[21] The number of missing observations of sensible and latent heat fluxes inhibited the calculation of average observed diurnal cycles for these two quantities. Based on the time series for the respective quantities, it is, however, possible to see fragments of observed diurnal cycles (not shown). The amplitude of these fragmented cycles compares well with the cycles produced by NASA-MERRA and NCEP-CFSR.

[22] The presence of observed cycles in the heat fluxes is also suggested by the average diurnal cycles in the

Table 1. Diurnal Cycles for the 93-Day Period[c]

	Era-Interim		ERA-40		JRA-25		NASA-MERRA		NCEP-CFSR		NCEP-DOE		OBS
Variable	1.	2.	1.	2.	1.	2.	1.	2.	1.	2.	1.	2.	1.
SW in	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
SW out	Yes	No	Yes	No	Yes	No	Yes	Yes (−14%)	Yes	No	Yes	Yes (+0.3%)	Yes
LW in	No	—	No	—	Yes	No	No	—	Yes	No	Yes	No	Yes
LW out	Yes	No	No	—	Yes	No	Yes	No	Yes	No	Yes	No	Yes
CF	No	—	Yes	Yes (−13%)	Yes	No	No	—	No	—	No	—	Yes
q	Yes	No	Yes	No	Yes	No	—[d]	—[d]	Yes	No	Yes	No	Yes
T	Yes	No	Yes	No	Yes	No	—[d]	—[d]	Yes	No	Yes	No	Yes
W	Yes	No	Yes	No	Yes	No	—[d]	—[d]	No	—	No	—	Yes

[a]At the 95% confidence level, is there a diurnal cycle?
[b]If there is a diurnal cycle, is it, at the same confidence level, different from the observed cycle?
[c]If the modeled and observed cycles are different, the percentage in parenthesis gives the difference with respect to the observations. SW in, incoming shortwave; SW out, outgoing shortwave; LW in, incoming longwave; LW out, outgoing longwave; CF, cloud fraction; q, humidity; T, temperature; W, wind speed.
[d]Too many missing values to determine diurnal cycles for near-surface specific humidity, temperature and wind speed.

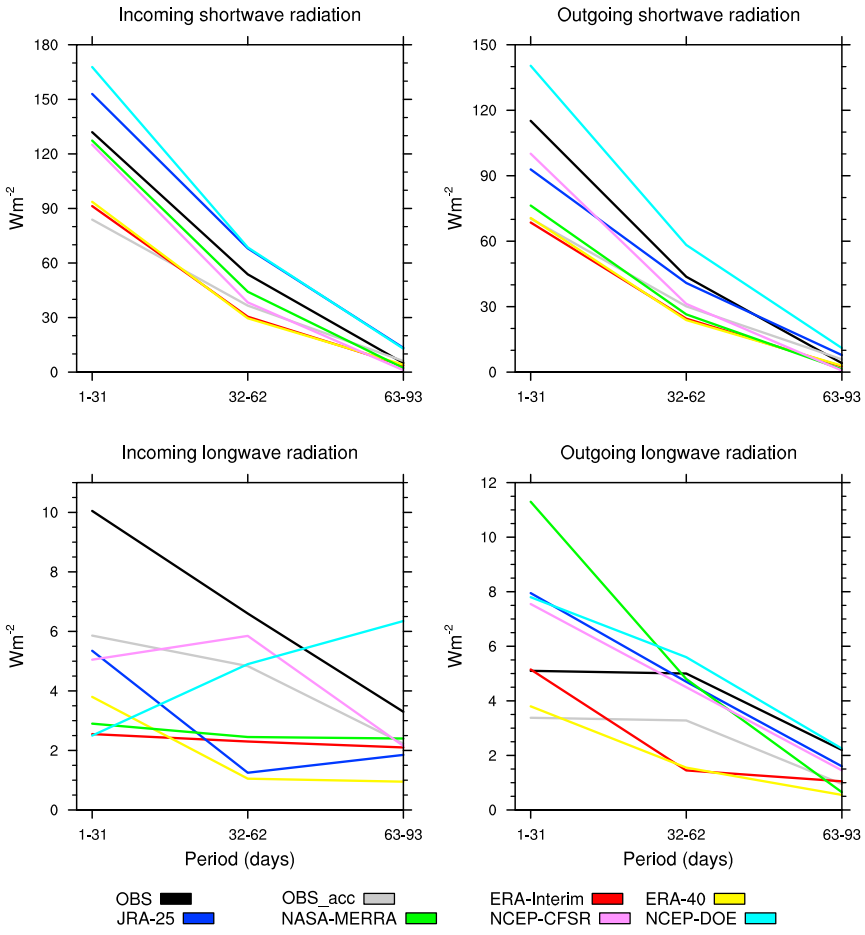

Figure 3. As in Figure 2 but featuring the variability of the amplitude of the diurnal cycle based on the three 31-day subsets.

differences in observed temperature and specific humidity between the snow surface and air (5 m height), 0.4°C and 0.04 g/kg, respectively. The small amplitude (~2 W m^{-2}) for the diurnal cycles of sensible and latent heat fluxes in ERA-Interim and ERA-40 is, at least partly, explained by the averaging of the accumulated fluxes (Figure 4). The variability in the amplitude, as given by the different reanalyses, decreases from summer to fall (Figure 5).

3.5. Other Considered Quantities

[23] We also calculated the diurnal cycles for cloud cover fraction, near-surface temperature and specific humidity, and wind

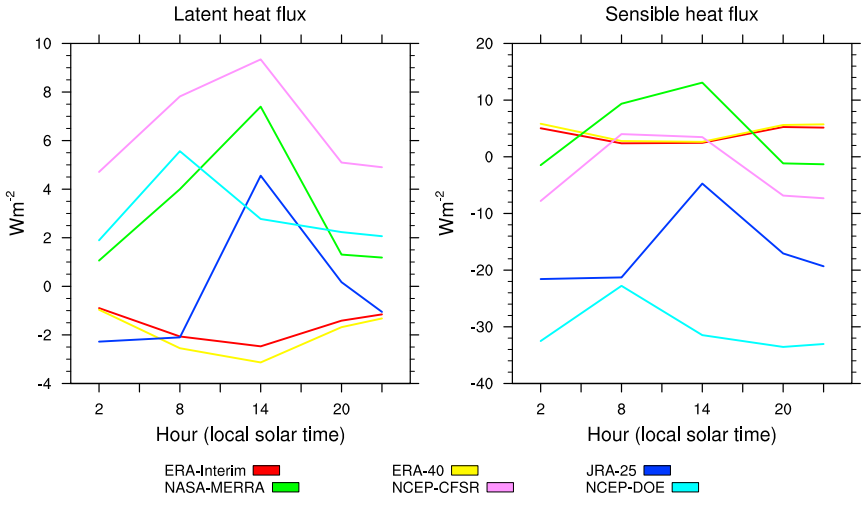

Figure 4. The diurnal cycles for the latent and sensible heat fluxes averaged for the entire Ice Station Weddell period based on observed and reanalysis values at 02, 08, 14, and 20 LST values.

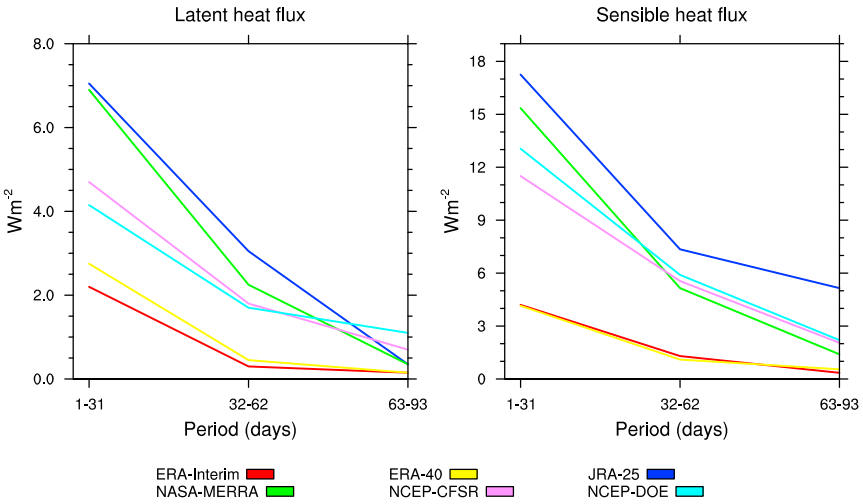

Figure 5. As in Figure 4 but featuring the variability of the amplitude of the diurnal cycle based on the three 31-day subsets.

speed (Figure 6). The diurnal cycles for these quantities are characterized by a small amplitude, and in most cases, the cycles given by the reanalyses fall within the uncertainties of the observed cycle. Cloud cover fraction is an exception. Observations feature a statistically significant, albeit weak, cycle, especially for the first third of the 93-day period of interest. Unsurprisingly, the reanalyses rarely manage to capture such a cycle. The importance of cloud cover fraction for the results is further discussed in Section 7. For near-surface temperature, humidity, and wind speed, the cycles for NASA-MERRA could not be obtained due to a large number of missing values. NASA-MERRA assigns a missing value to a grid

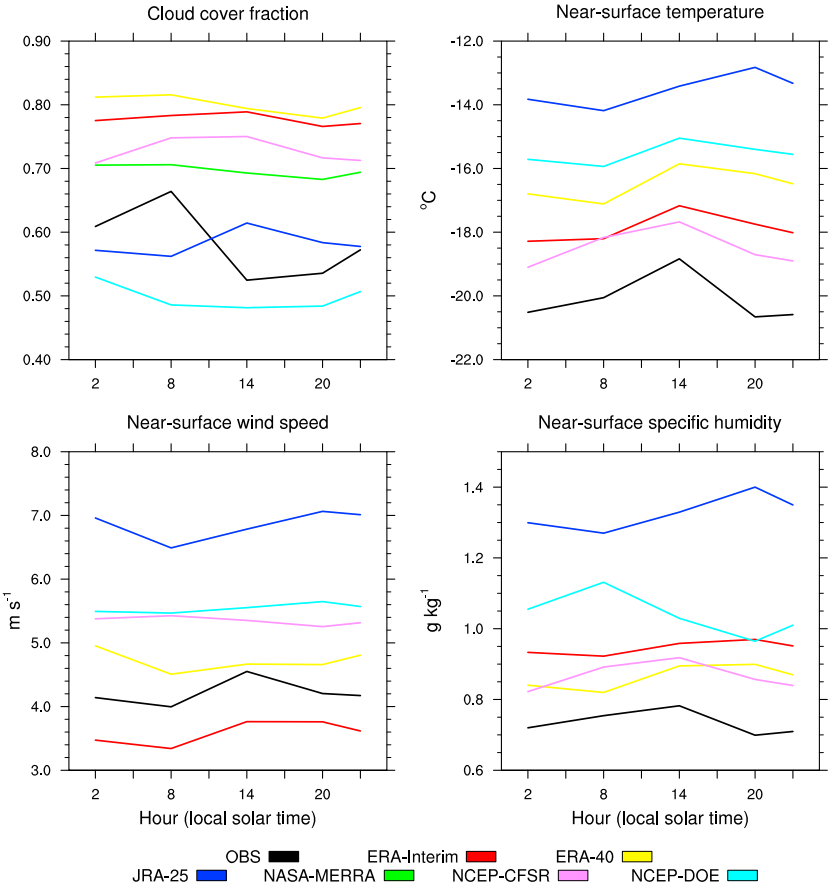

Figure 6. The diurnal cycles for cloud cover fraction, near-surface temperature, wind speed, and specific humidity averaged for the entire Ice Station Weddell period based on observations and reanalysis values 02, 08, 14, and 20 LST values.

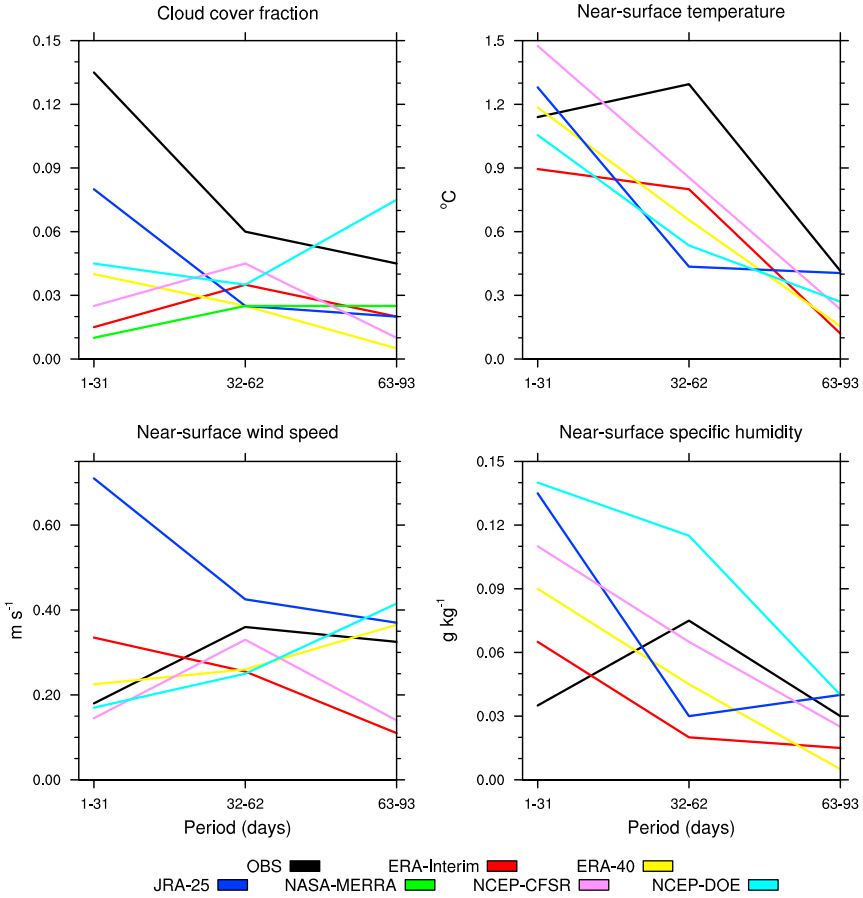

Figure 7. As in Figure 6 but featuring the variability of the amplitude of the diurnal cycle based on the three 31-day subsets.

point when the pressure at a model level is greater than the surface pressure and does not carry out extrapolation beneath the Earth's surface like other reanalyses.

[24] The diurnal cycle of wind speed has a daytime maximum, presumably due to the enhanced downward turbulent transport of momentum in conditions of less stable stratification (Figure 6). ERA-Interim is the only reanalysis that correctly reproduces the observed characteristics of the cycle. Most reanalyses underestimate the amplitude, which is in accordance with the results of *Svensson et al.* [2011] from an extensive model validation study for a clear-sky night in Kansas, USA.

4. Trends in the Diurnal Cycles

[25] As expected, incoming and outgoing shortwave radiation feature a decreasing trend for the diurnal cycle, which is well captured by the reanalyses (Figure 3). All model products also feature a diminishing amplitude for latent and sensible heat fluxes (Figure 5) as well as for the outgoing longwave radiation (Figure 3) though the suggested rates vary. For incoming longwave radiation, however, none of the reanalyses agree with the observed decreasing trend. Other problematic quantities include cloud cover fraction, near-surface wind speed, and near-surface specific humidity (Figure 7). The variations in the amplitude of near-surface temperature are better captured. In the observations, however, the amplitude decreases only from days 32–62 to days 63–93; whereas, in the reanalyses, the amplitude gets smaller throughout the 93-day period.

[26] The diurnal amplitudes of snow surface temperature (~1.5°C) and near-surface air temperature (~1°C) observed at ISW during days 1–62 are of the same order of magnitude as observed at ISPOL in December 2004 to January 2005 [*Vihma et al.*, 2009]. Interestingly, the amplitude of incoming shortwave radiation is much larger at ISPOL (320 W m^{-2}) than at ISW (130 and 50 W m^{-2} on days 1–31 and 32–62, respectively; Figure 5). Although the ISW diurnal cycles are underestimates due to the 6-hourly resolution, such large differences in radiation but approximately similar cycles in temperature at both stations may demonstrate the reduction of the diurnal cycle of surface temperature when the snow surface is melting.

[27] However, the differences between ISPOL and ISW cannot be fully explained by melting; only in a few cases at ISPOL did the snow surface temperature reach the melting point. One should, instead, look at the overall snow characteristics: at ISW, the snow was dry [*Andreas et al.*, 2004], whereas at ISPOL, the snow just a few centimeters below the surface was wet most of the time [*Vihma et al.*, 2009]. The extinction coefficient for shortwave radiation is much larger for dry than wet snow [*Cheng et al.*, 2003]. Hence, at ISPOL, a larger part of the solar radiation penetrated into the snow pack, reducing the diurnal cycle of snow surface temperature, and indirectly, the corresponding cycle for near-surface air temperature.

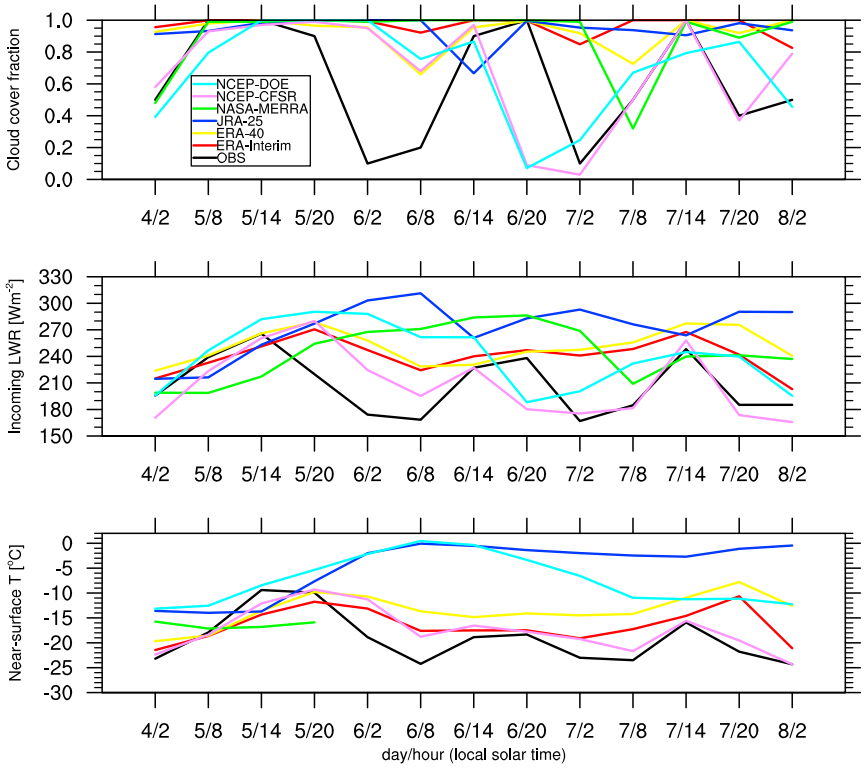

Figure 8. Observed and reanalyses cloud cover fraction, incoming longwave radiation, and near-surface air temperature on Ice Station Weddell on 5–8 April 1992.

[28] The reanalyses validated here do not take into account the penetration of solar radiation into the snow pack. This may have contributed to the excessive diurnal cycle of snow surface temperature (as diagnosed from the outgoing longwave radiation in Figure 3) in all reanalyses during the period of most intense radiation (days 1–31; note that the ERA-Interim and ERA-40 results should be compared against the accumulated observations in Figure 3). This problem practically vanishes when the solar radiation decreases in autumn, but another problem emerges during days 32–62 when all reanalyses strongly underestimate the diurnal cycle of near-surface air temperature (Figure 7). During this period, the observed temperature inversions were strongest [*Tastula et al.*, 2012], but reanalyses yield less stable stratification with more clouds and higher near-surface temperatures (Figure 8 presents a sample of the period). This result suggests an overestimation of the ABL height in the reanalyses. A higher ABL has a larger heat capacity, which reduces the diurnal temperature cycle.

5. General Error Statistics

[29] The root-mean square errors (RMSEs), biases, and correlation coefficients were calculated for all variables over the ISW period (Figure 9). The following general observations can be made. Incoming shortwave and outgoing longwave radiation are, in most cases, overestimated. There is more scatter among the biases from different reanalyses for outgoing shortwave and incoming longwave radiation. NCEP-CFSR stands out by yielding the best or shared-best RMSEs for six quantities out of 10 and the best or shared-best correlation coefficients for five out of the six flux quantities. JRA-25 performs poorly, with near-surface specific humidity, temperature, and wind speed each featuring a significant positive bias.

6. Skill Scores

[30] To rank different reanalyses with respect to their ability to yield diurnal cycles close to the observed ones, we use a skill score approach in which a reanalysis is awarded points based on how well it predicts the range and shape of the cycle. To evaluate the range of the predicted diurnal cycle, we defined the following criterion: if the 95% confidence level showed a diurnal cycle in the observations, a reanalysis was awarded one point if, at the same 95% confidence level, it was possible that the reanalysis gave the same range for the diurnal cycle as was observed. To evaluate the shape of the cycle, we awarded a reanalysis one point if the shape of its diurnal cycle, based on the 02, 08, 14, and 20 LST values, was consistent with the shape of the cycle in the observations. We considered all the quantities studied except the sensible and latent heat fluxes, for which the observed diurnal cycle could not be determined. The results point out the superiority of ERA-Interim and ERA-40 (Table 2). NCEP-CFSR is the third best.

[31] We also applied the skill score approach to general error statistics. RMSEs, biases, and correlation coefficients of all 10 quantities were considered. A reanalysis was awarded one point if it gave the lowest RMSE or bias. In the case of the correlation coefficient, one point went to the reanalysis yielding the highest value. If two analyses were equally good, both were awarded half a point. If more than two shared the best result, no points were given.

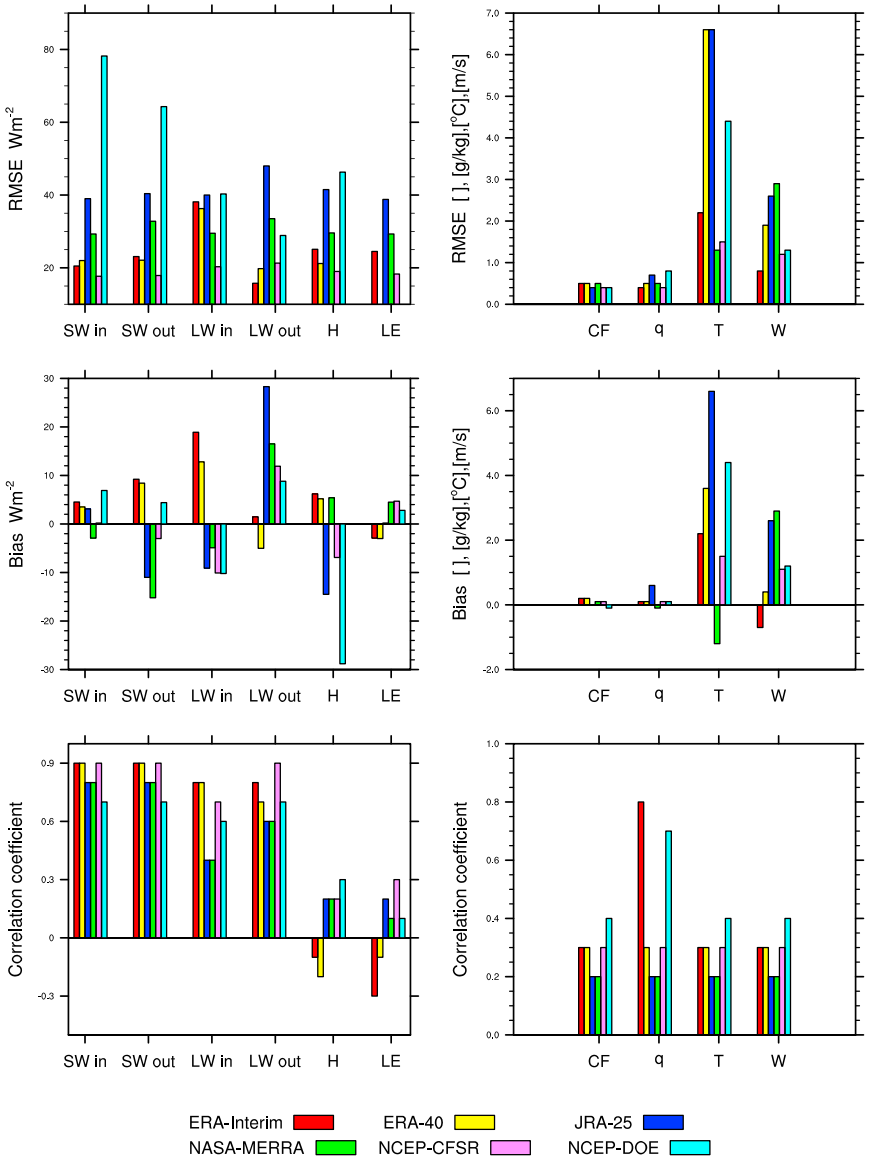

Figure 9. General error statistics for the 93-day period for each analysis. The left panels are root-mean-square errors, biases, and correlation coefficients for incoming (SW in) and outgoing (SW in) shortwave, incoming (LW in) and outgoing (LW in) longwave, and sensible (H) and latent (LE) heat fluxes. The right panels are the same statistics for cloud fraction (CF), near-surface specific humidity (q), temperature (T), and wind speed (W).

NCEP-CFSR ends up getting nine points, almost twice as many as the next best score by NCEP-DOE (Table 2).

7. Challenging Conditions for Reanalyses

[32] Cloud cover fraction is poorly simulated by reanalyses (RMSE typically 0.5 and correlation coefficient 0.2 to 0.4). Therefore, a particularly challenging setup occurs when cloud cover fraction features a strong diurnal signal with variability from overcast to nearly clear skies and back. This bimodal cloud distribution was a hallmark of Ice Station Weddell. The failure to capture such a cycle has a direct effect on the modeled cycles of other quantities. The event we use as an example is from 5 to 8 April (Figure 8). On 5 April, at 14 LST, the sky, according to observations, was overcast. Then a rapid clearing occurred; and 12 h later, the cloud cover fraction was only 0.1. After another 12 h, the cloud cover fraction was back at 0.9.

[33] As demonstrated by Figure 8, none of the reanalyses capture this rapid transition. This failure is immediately reflected in incoming longwave radiation: with decreasing cloud cover, the observed longwave flux decreases drastically. In the reanalyses, however, the flux either starts to decrease too late (NCEP-CFSR, NCEP-DOE, ERA-Interim, and ERA-40) or even features an increase instead (JRA-25 and NASA-MERRA). The failure with cloud cover fraction is also seen in the near-surface temperature time series, where all the reanalyses become positively biased toward the night between 5 and 6 April.

[34] Subsequently, another similar cycle follows in cloud cover fraction. This time, though, NCEP-CFSR succeeds perfectly in simulating cloud cover fraction between 6 April

Table 2. Skill Score Points Awarded[a]

			ERA-Interim	ERA-40	JRA-25	NASA-MERRA	NCEP-CFSR	NCEP-DOE
Diurnal cycles	Range	93 days	8	8	8	3	8	7
		Days 1–31	7	7	6	2	6	7
		Days 32–62	9	9	8	4	8	7
		Days 63–93	9	9	7	5	7	5
	Shape	93 days	5	4	0	0	4	0
		Days 1–31	4	4	0	0	3	1
		Days 32–62	4	3	1	1	2	1
		Days 63–93	1	0	0	0	1	0
	Total		47	44	30	15[b]	39	28
	Rank		1	2	4	6	3	5
General error statistics	Total		5	2	2.5	3	9	5.5
	Rank		3	5	4	3	1	2

[a]See section 6 for explanation.
[b]Too many missing values to determine diurnal cycles for near-surface specific humidity, temperature, and wind speed.

20 LST and 7 April 14 LST. This results in a much better representation for incoming longwave radiation and near-surface temperature for 7 April while the other reanalyses are still clearly erroneous.

[35] The close relationship between cloud cover fraction and incoming longwave radiation is also apparent during other periods. On 22–26 April, overcast conditions are captured by all reanalyses, and the modeled incoming longwave radiation is close to observations. The same situation occurs on 30–31 March, but for clear conditions. Of course, the relationship between cloud cover fraction and incoming longwave radiation is not straightforward. For instance, on 21–25 May, the best cloud cover fraction is given by JRA-25, while the incoming longwave radiation by the same reanalysis has the largest error of all, demonstrating that radiative transfer is not controlled by cloud cover only.

8. Advances Made and Remaining Challenges

[36] One of the objectives of this paper was to address the progress made by ECMWF and NCEP, in the transition from ERA-40 and NCEP-DOE to ERA-Interim and NCEP-CFSR. ERA-40 and ERA-Interim are remarkably similar in reproducing the diurnal cycle. The shape and amplitude of the cycles of incoming and outgoing shortwave radiation compare well with the observed cycles calculated from on the accumulated dataset. However, neither ERA-Interim nor ERA-40 managed to reproduce a statistically significant cycle for incoming longwave radiation. Both reanalyses also feature amplitudes for the diurnal cycles of sensible and latent heat flux that are too small; but these shortcomings are, at least partly, due to the averaging procedure for the accumulated fluxes. When it comes to the shape of the diurnal cycle, however, these two reanalyses give the best performance. In the shape section of the skill score test (Table 2), ERA-Interim gets 14 points; and ERA-40, 11 points. Only NCEP-CFSR was almost equally good with 10 points. The other three analyses received only one to two points. In the overall comparison, ERA-Interim and ERA-40 yielded the two top scores with respect to the diurnal cycle. In the test related to general error statistics ERA-Interim proved to have more skill than its predecessor.

[37] In contrast to ERA-40 and ERA-Interim, NCEP-DOE and NCEP-CFSR are not the least bit similar in their reproduction of the diurnal cycle. The most severe problem with NCEP-DOE is the grossly erroneous cycles of incoming and outgoing shortwave radiation with respect to both the amplitude and the shapes of the cycles over the whole 93-day period. On top of this, the amplitude is too high in both incoming and outgoing shortwave radiation during the last subset (days 63–93) when the amount of incoming solar radiation is small. Though NCEP-CFSR is a much improved version compared to NCEP-DOE, it also fails during the last subset by giving a range that is too small. Other than that, NCEP-CFSR has a substantial advantage over the other renalyses in the case of general error statistics.

[38] JRA-25 is successful in reproducing the ranges for the diurnal cycles (30 points in the range section of the skill score comparison—the shared first place of this section with NCEP-CFSR; Table 2). JRA-25, however, had problems with the shape of the cycle and obtained only one point for skill here. The problem is the 08 LST value which is, in most cases, too low. Even more serious shortcomings lie within the general error statistics: JRA-25 exhibits a strong positive bias for near-surface specific humidity, temperature, and wind speed.

[39] We could not properly evaluate skill scores for NASA-MERRA because of the large number of missing values that precluded our calculating diurnal cycles for the near-surface specific humidity, temperature, and wind speed. For incoming and outgoing shortwave radiation, NASA-MERRA features the same kind of distorted shape of the cycle as was observed with NCEP-DOE, though the 08 LST peak is not as pronounced. Another shortcoming lies in the radiative fluxes during days 1–31: NASA-MERRA significantly underestimates outgoing shortwave radiation and overestimates outgoing longwave radiation. A possible explanation for this is a value for the sea ice albedo that is too low during the summer months.

9. Summary

[40] This study presented the first attempt to evaluate how six reanalyses do in representing the diurnal cycle of surface-level meteorological variables observed over Antarctic sea ice. The reanalyses we evaluated are ERA-Interim, ERA-40, JRA-25, NASA-MERRA, NCEP-CFSR, and NCEP-DOE. As validation data, we used Ice Station Weddell, which provided 93 consecutive days of data from the western Weddell Sea between late February and late May 1992. The variables we considered are surface temperature; surface-level values of air temperature, specific humidity, and wind speed; cloud

fraction; the four radiation components, incoming and outgoing longwave and shortwave radiation; and the turbulent surface fluxes of sensible and latent heat.

[41] All the renalyses exhibit large errors related to the amplitude and shape of the diurnal cycle of these variables in the Antarctic sea ice zone. The most striking errors are related to the cloud cover fraction and incoming longwave radiation. Apart from the diurnal cycles, many reanalyses also have large biases and root-mean-square errors and low correlations when compared with the Ice Station Weddell observations. The largest biases were related to the common overestimation of snow surface temperature as well as near-surface air temperature, specific humidity, and wind speed.

[42] A promising aspect in the results was the progress that both the ECMWF and, especially, NCEP have made in recent years. Their newest reanalyses were ranked among the three best in the validation.

[43] **Acknowledgments.** The work of E.M.T. was partially funded by the ARO grant W911NF-09-1-0018 and the Paul L. Getting Memorial Fellowship. The U.S. National Science Foundation, through Award 10–19322, supported E.L.A.'s participation in this study. NSF also supported the data collection on Ice Station Weddell with previous awards. The Academy of Finland supported the work of T.V. through the AMICO project (contracts 128533 and 263918). We thank Aleksandr Makshtas, Kerry Claffey, and Boris Ivanov for their help in collecting and processing the data from Ice Station Weddell.

References

Andreas, E. L., K. J. Claffey, and A. P. Makshtas (2000), Low-level atmospheric jets and inversions over the western Weddell Sea, *Boundary Layer Meteorol.*, *97*, 459–486.

Andreas, E. L., P. S. Guest, P. O. G. Persson, C. W. Fairall, T. W. Horst, R. E. Moritz, and S. R. Semmer (2002), Near-surface water vapor over polar sea ice is always near ice saturation, *J. Geophys. Res.*, *107*(C10), doi:10.1029/2000JC000411.

Andreas, E. L., R. E. Jordan, and A. P. Makshtas (2004), Simulations of snow, ice, and near-surface atmospheric processes on Ice Station Weddell, *J. Hydrometerol.*, *5*(4), doi:10.1175/1525-7541(2004)005<0611:SOSIAN>2.0.CO;2.

Andreas, E. L., R. E. Jordan, and A. P. Makshtas (2005), Parameterizing turbulent exchange over sea ice: The Ice Station Weddell results, *Boundary Layer Meteorol.*, *114*(2), doi:10.1007/s10546-004-1414-7.

Atlaskin E., and T. Vihma (2012), Evaluation of NWP results for wintertime nocturnal boundary-layer temperatures over Europe and Finland, *Q. J. R. Meteorol. Soc.*, *138*(667), doi:10.1002/qj.1885

Bendat, J. S., and A. G. Piersol (1986), *Random Data: Analysis and Measurement Procedures*, 2nd Ed., pp. 566, John Wiley & Sons, Inc., New York.

Bracegirdle, T. J., and G. J. Marshall (2012), The reliability of Antarctic tropospheric pressure and temperature in the latest global reanalyses, *J. Clim.*, *25*(20), doi:10.1175/JCLI-D-11-00685.1.

Bromwich, D. H., R. L. Fogt, K. I. Hodges, and J. E. Walsh (2007), A tropospheric assessment of the ERA-40, NCEP, and JRA-25 global reanalyses in the polar regions, *J. Geophys. Res.*, *112*, D10111, doi:10.1029/2006JD007859.

Bromwich, D. H., J. P. Nicolas, and A. J. Monaghan (2011), An assessment of precipitation changes of Antarctica and the Southern Ocean since 1989 in contemporary reanalyses, *J. Clim.*, *24*(16), doi:10.1175/2011JCLI4074.1.

Chandler, W. S., C. H. Whitlock, and P. W. Stackhouse Jr. (2004), NASA climatological data for renewable energy assessment, *J. Solar Energy Engineering*, *126*(3), doi:10.1115/1.1748466.

Cheng, B., J. Launiainen, and T. Vihma (2003), Modelling of superimposed ice formation and sub-surface melting in the Baltic Sea, *Geophysica*, *39* (1–2).

Cullather, R. I., and M. G. Bosilovich (2011), The moisture budget of the polar atmosphere in MERRA, *J. Clim.*, *24*(11), doi:10.1175/2010JCLI4090.1.

Dai, A., and K. E. Trenberth (2004), The diurnal cycle and its depiction in the Community Climate System Model, *J. Clim.*, *17*(5), doi:10.1175/1520-0442(2004)017<0930:TDCAID>2.0.CO;2.

Dee, D. P., et al. (2011), The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.*, *137*(656), doi:10.1002/qj.828.

He, Y., N. A. McFarlane, and A. H. Monahan (2012), The influence of boundary layer processes on the diurnal variation of the climatological near-surface wind speed probability distribution over land, *J. Clim.*, *25* (18), doi:10.1175/JCLI-D-11-00321.1.

Jakobson, E., T. Vihma, T. Palo, L. Jakobson, H. Keernik, and J. Jaagus (2012), Validation of atmospheric reanalyses over the central Arctic Ocean, *Geophys. Res. Lett.*, *39*, L10802, doi:10.1029/2012GL051591.

Kalnay, E., et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bull. Amer. Meteorol. Soc.*, *77*, 437–471.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter (2002), NCEP-DOE AMIP-II reanalysis (R-2), *Bull. Amer. Meteorol. Soc.*, *83*(11), doi:10.1175/BAMS-83-11-1631.

King, J. C. (2003), Validation of ECMWF sea level pressure analyses over the Bellingshausen Sea, Antarctica, *Wea. Forecasting*, *18*(3), doi:10.1175/1520-0434(2003)18<536:VOESLP>2.0.CO;2.

Lüpkes, C., T. Vihma, E. Jakobson, G. König-Langlo, and A. Tetzlaff (2010), 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.

Mahrt, L. (1999), Stratified atmospheric boundary layers, *Boundary Layer Meteorol.*, *90*, 375–396.

Monin, A. S., and A. M. Obukhov (1954), Basic laws of turbulent mixing in the atmospheric surface layer, *Trudy Geofiz. Inst. Akad. Nauk* SSSR No. *24* (151), 163–187.

Niros, A., T. Vihma, and J. Launiainen (2002), Marine meteorological conditions and air-sea exchange processes over the Baltic Sea in 1990s, *Geophysica*, *38*(1–2).

Onogi, K., et al. (2007), The JRA-25 reanalysis, *J. Meteorol. Soc. Japan*, *85*.

Panday, A. K., R. G. Prinn, and C. Schär (2009), Diurnal cycle of air pollution in the Kathmandu Valley, Nepal: 2. Modeling results, *J. Geophys. Res.*, *114*, D21308, doi:10.1029/2008JD009808.

Pavelsky, T. M., J. Boé, A. Hall, and E. J. Fetzer (2011), Atmospheric inversion strength over polar oceans in winter regulated by sea ice, *Clim. Dyn.*, *36*(5–6), doi:10.1007/s00382-010-0756-8.

Press, W., B. Flannery, S. Teukolsky, and W. T. Vetterling (1988), *Numerical Recipes in C, 735 pp.*, Cambridge University Press, Cambridge, U.K.

Saha, S., and et al. (2010), The NCEP Climate Forecast System reanalysis, *Bull. Amer. Meteorol. Soc.*, *91*(8), doi:10.1175/2010BAMS3001.1.

Screen, J. A., and I. Simmonds (2011), Erroneous Arctic temperature trends in the ERA-40 reanalysis: A closer look, *J. Clim.*, *24*(10), doi:10.1175/2010JCLI4054.1.

Stammerjohn, S., R. Massom, D. Rind, and D. Martinson (2012), Regions of rapid sea ice change: An inter-hemispheric seasonal comparison, *Geophys. Res. Lett.*, *39*, L06501, doi:10.1029/2012GL050874.

Steeneveld, G. J., T. Mauritsen, E. I. F. de Bruijn, J. Vila-Guerau de Arellano, G. Svensson, and A. A. M. Holtslag (2008), Evaluation of limited area models for the representation of the diurnal cycle and contrasting nights in CASES-99, *J. Appl. Meteorol. Climatol.*, *47*(3), doi:10.1175/2007JAMC1702.1.

Svensson, G., et al. (2011), Evaluation of the diurnal cycle in the atmospheric boundary layer over land as represented by a variety of single column models—The second GABLS experiment, *Boundary Layer Meteorol.*, *140*(2), doi:10.1007/s10546-011-9611-7.

Tastula, E.-M., T. Vihma, and E. L. Andreas (2012), Evaluation of Polar WRF from modeling the atmospheric boundary layer over Antarctic sea ice in autumn and winter, *Mon. Weather Rev.*, *140*(12), doi:10.1175/MWR-D-12-00016.1.

Tjernström, M. (2005), The summer Arctic boundary layer during the Arctic Ocean Experiment 2001 (AOE-2001), *Boundary Layer Meteorol.*, *117*(1), doi:10.1007/s10546-004-5641-8.

Tjernström, M. (2007), Is there a diurnal cycle in the summer cloud-capped Arctic boundary layer? *J. Atmos. Sci.*, *64*(11), doi:10.1175/2007JAS2257.1.

Uppala, S. M., and et al. (2005), The ERA-40 re-analysis, *Q. J. R. Meteorol. Soc.*, *131*(612), doi:10.1256/qj.04.176.

Vancoppenolle, M., et al. (2011), Assessment of radiation forcing data sets for large-scale sea ice models in the Southern Ocean, *Deep Sea Res. Part II*, *58* (9), doi:10.1016/j.dsr2.2010.10.039.

Vihma, T., J. Uotila, B. Cheng, and J. Launiainen (2002), Surface heat budget over the Weddell Sea: Buoy results and model comparisons, *J. Geophys. Res.*, *107*(C2), doi:10.1029/2000JC000372.

Vihma, T., M. M. Johansson, and J. Launiainen (2009), Radiative and turbulent surface heat fluxes over sea ice in the western Weddell Sea in early summer, *J. Geophys. Res.*, *114*, C04019, doi:10.1029/2008JC004995.

Walsh, J. E., and W. L. Chapman (1998), Arctic cloud–radiation–temperature associations in observational data and atmospheric reanalyses, *J. Clim.*, *11* (11), 3030–3045.

Wendler, G., B. Hartmann, C. Wyatt, M. Shulski, and H. Stone (2005), Midsummer energy balance for the southern seas, *Boundary Layer Meteorol.*, *117*(1), doi:10.1007/s10546-004-7090-9.