



On the possibilities to use atmospheric reanalyses to evaluate the warming structure in the Arctic

C. E. Chung¹, H. Cha¹, T. Vihma², P. Räisänen², and D. Decremer¹

¹School of Environmental Science and Engineering, Gwangju Institute of Science and Technology, Gwangju 500-712, Korea

²Finnish Meteorological Institute, 00101 Helsinki, Finland

Correspondence to: C. E. Chung (eddy@gist.ac.kr)

Received: 17 July 2013 – Published in Atmos. Chem. Phys. Discuss.: 23 August 2013

Revised: 21 October 2013 – Accepted: 22 October 2013 – Published: 18 November 2013

Abstract. There has been growing interest in the vertical structure of the recent Arctic warming. We investigated temperatures at the surface, 925, 700, 500 and 300 hPa levels in the Arctic (north of 70° N) using observations and four reanalyses: ERA-Interim, CFSR, MERRA and NCEP II. For the period 1979–2011, the layers at 500 hPa and below show a warming trend in all seasons in all the chosen reanalyses and observations. Restricting the analysis to the 1998–2011 period, however, all the reanalyses show a cooling trend in the Arctic-mean 500 hPa temperature in autumn, and this also applies to both observations and the reanalyses when restricting the analysis to the locations with available IGRA radiosoundings. During this period, the surface observations mainly representing land areas surrounding the Arctic Ocean reveal no summertime trend, in contrast with the reanalyses whether restricted to the locations of the available surface observations or not.

In evaluating the reanalyses with observations, we find that the reanalyses agree better with each other at the available IGRA sounding locations than for the Arctic average, perhaps because the sounding observations were assimilated into reanalyses. Conversely, using the reanalysis data only from locations matching available surface (air) temperature observations does not improve the agreement between the reanalyses. At 925 hPa, CFSR deviates from the other three reanalyses, especially in summer after 2000, and it also deviates more from the IGRA radiosoundings than the other reanalyses do. The CFSR error in summer T_{925} is due mainly to underestimations in the Canadian-Atlantic sector between 120° W and 0°. The other reanalyses also have negative biases in this longitude band.

1 Introduction

The surface warming in the Arctic is observed to be at least twice as large as the global average warming in the recent decades (Hassol, 2004; Bekryaev et al., 2010). Causes for the so-called Arctic amplification of the warming have been proposed to be the snow and ice feedbacks (Manabe, 1983; Hall, 2004) and the poleward energy transport from lower latitudes (Alexeev et al., 2005) among many proposed causes. Graversen et al. (2008) and Screen and Simmonds (2010) used the vertical structure of Arctic temperature trends in reanalyses to gain insights into this issue. They argue that if the maximum warming occurs much above the surface, the poleward energy transport would be the primary mechanism for the warming amplification. Graversen et al. (2008) found the maximum warming well above the surface, whereas Screen and Simmonds (2010) showed the warming to be largest in the lowest layers. Graversen et al. (2008) and Screen and Simmonds (2010) used different reanalyses and different periods for their analyses.

Chung and Räisänen (2011), on the other hand, addressed the origin of Arctic warming rather than the Arctic amplification. They hypothesize, based on idealized climate model experiments, that if the *summertime* warming is largest well above the surface, the poleward energy transport would be mainly responsible for the Arctic warming irrespective of the warming structure in winter. Taken together, the vertical profile of Arctic warming has emerged as one of the top climate issues.

Here, we use sounding and surface observations, and reanalyses collectively to investigate the recent warming and its vertical structure in the Arctic. In doing so, we also evaluate

the reanalyses in Arctic warming. Reanalysis evaluation is important because reanalyses are so broadly applied in climate research, including Arctic research. Reanalyses are not observation but commonly treated as observation in the literature. For example, the atmospheric forcing for ocean, sea ice, glacier, and hydrological models are often taken from reanalyses. Reanalyses are also employed in the studies of climate variability and trends as well as occurrence of extreme events. Reanalyses are, however, not free of errors (Lüpkes et al., 2010; Bromwich et al., 2011; Jakobson et al., 2012).

Reanalysis is a system where observations are assimilated into a global model in order to provide the atmospheric state continuous in space and time (Saha et al., 2010; Dee et al., 2011; Rienecker et al., 2011). Reanalyses differ from each other due to several reasons. First, the usage of observations varies between reanalyses. While all the reanalyses assimilate radiosonde sounding data, there are differences in the assimilation of satellite data. Also, the assimilation method varies. ERA-Interim applies a method based on four-dimensional data assimilation (4-D-VAR), where exact time of observations is taken into account by sophisticated means, whereas other reanalyses apply simpler methods. Second, there are large differences in the horizontal and vertical resolutions of the models applied. Third, the physical parameterization schemes for radiative transfer, turbulent mixing, cloud physics, and surface processes vary between the models.

Over the Arctic, where there are very few in situ observations, the quality of reanalyses is particularly questioned. Recent studies (e.g., Screen et al., 2012) tend to utilize multiple reanalyses to establish the robustness of the vertical warming structure in the Arctic. In this study, we examine 2 m air temperature, temperature (T) at 925, 700, 500 and 300 hPa, and the temperature difference between 925 and 500 hPa levels in the latest reanalyses so as to estimate the accuracy of each reanalysis product in Arctic warming. Recently, Alexeev et al. (2012) used sounding observations to evaluate the warming structures in the older NCEP/NCAR and ERA-40 reanalyses. Here, we consider the most recent reanalyses and focus on Arctic-averaged temperatures.

2 Surface (air) temperature

We analyze monthly mean 2 m air temperature fields from three latest-generation atmospheric reanalyses: ECMWF ERA-Interim reanalysis (ERA-I) (Dee et al., 2011), NASA Modern-Era Retrospective analysis for Research and Applications (MERRA) (Rienecker et al., 2011) and the NCEP Climate Forecast System Reanalysis (CFSR) (Saha et al., 2010). The 2 m air temperature will be referred to as the surface air temperature for the remainder of the paper. The CFSR surface air temperature analysis is available at two resolutions; here we use the T62 Gaussian grid version. We also analyze NCEP's earlier-generation reanalysis product,

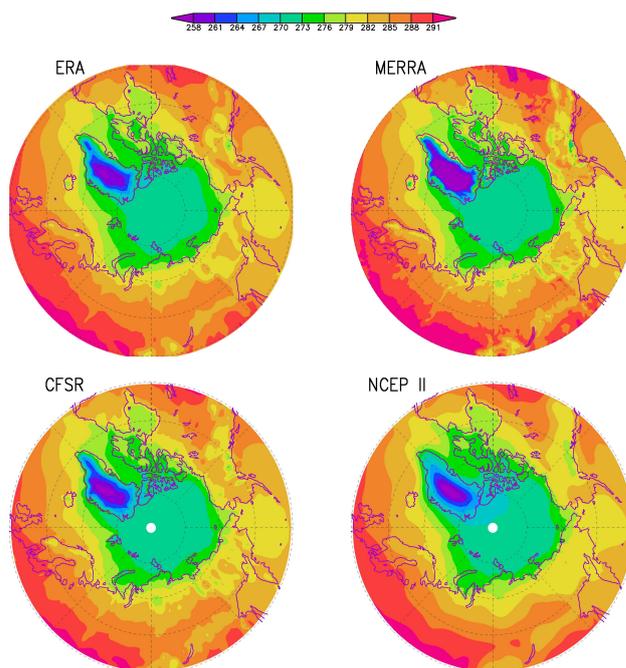


Fig. 1. June–August average surface air temperature climatology for the period 1979–2011 in units of K. In the case of CFSR, the climatology is for 1979–2009.

the so-called NCEP II reanalysis (Kanamitsu et al., 2002). The period of analysis is from 1979 to 2011. Due to data availability, CFSR is only analyzed until the year 2009.

The reanalyses agree on the large-scale features in surface air temperature climatology (see Fig. 1 for the summer season). Furthermore, for all four reanalyses considered, the domain-average air surface temperature for the 70°–90° N region shows a clear warming trend, both for annual and seasonal means (Fig. 2a–e). There are, however, also disagreements between the reanalyses, which will be discussed below.

To evaluate the surface air temperature in the reanalyses, we use the Goddard Institute for Space Studies (GISS) analysis of global surface temperature change (Hansen et al., 2010), referred to here as “GISTEMP”. GISTEMP integrates in situ surface air temperature measurements over land- and ship-based and satellite-derived sea surface temperature (SST) measurements. The SST measurements are, however, only used over year-round ice-free areas. The latter is because GISTEMP data were produced for comparison with the surface air temperature in climate models and only in year-round ice-free areas is SST anomaly a good approximation to surface air temperature anomaly (Hansen et al., 2010). The GISTEMP version we use here is a gridded monthly mean data set with a 250 km smoothing. The use of a smoothing distance of 250 km instead of 1200 km (the default of GISTEMP) avoids the uncertainty related to the extrapolation of temperature measurements made at Arctic

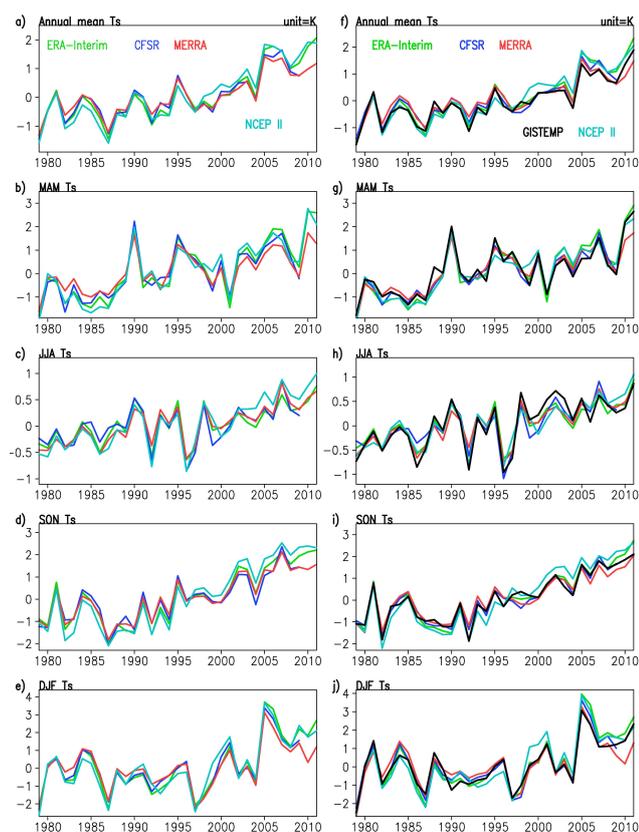


Fig. 2. 70–90° N average T_s anomalies relative to 1979–2009 climatology from four reanalyses and GISTEMP. On the left, the true 70–90° N average is shown. On the right, the average is calculated using only those reanalysis data that correspond to the available GISTEMP data in location and time. Note that T_s refers to surface air temperature (SAT) for the reanalyses, and to a combination of SAT over land and SST over ocean for GISTEMP.

observation sites to large distances over the open ocean or sea ice, thereby providing a more robust point of comparison for the reanalyses. On the other hand, a consequence of the 250 km smoothing distance is that it leaves a large amount of data gaps. The GISTEMP data we use are defined only for a fraction of the Arctic area (on average, 27 % from 1979 to 2011), mainly limited to the vicinity of the observation sites on the land and permanently ice-free parts of the Barents Sea and Greenland Sea (see Fig. 11). The GISTEMP algorithm and its application to the Arctic are discussed in more detail in Hansen et al. (2010).

Figure 2f–j compare the 70–90° N average surface air temperature anomalies in the reanalyses with those of the GISTEMP data. Unlike Fig. 2a–e, which use all reanalysis data north of 70° N, Fig. 2f–j use, after linear interpolation onto the GISTEMP grids, only those months and grid cells corresponding to valid GISTEMP data. In generating Fig. 2f–j, the area averaging is done for each month, and then seasonal or annual means are computed. The temperature anomalies are defined with respect to the climatology of years 1979–2009.

Note that while GISTEMP originally provides temperature anomalies with respect to the 1951–1980 climatology on each grid, the anomalies with respect to the 1979–2009 climatology are formed by subtracting the 1979–2009 mean of Arctic-average GISTEMP data.

The surface air temperature in the reanalyses are evaluated in Fig. 2f–j. In summer, the spread between the reanalyses (Fig. 2h) tends to decrease slightly from that in Fig. 2c. However, in winter, the spread seems to increase. Thus, overall, restricting the analyses to the regions with GISTEMP data does not improve the mutual agreement between the reanalyses. This might actually not be surprising because in the Arctic region GISTEMP is mainly based on in situ surface air temperatures from land, and among the reanalyses considered in this study, these data are only assimilated in ERA-Interim. This is done through a separate surface analysis, which is based on Optimal Interpolation, in contrast to 4dVAR in the main atmospheric analysis (Dee et al., 2011). MERRA assimilates in situ surface air temperatures only over oceans (from ships and buoys; Rienecker et al., 2011), while CFSR and NCEP II do not explicitly assimilate surface air temperature observations (Wang et al., 2011).

3 Upper-air temperatures

As for upper-air temperatures from the reanalyses, we again use ERA-Interim reanalysis, MERRA, CFSR and NCEP II reanalysis. The CFSR upper-air temperature product is available at two resolutions; here we use the $2.5^\circ \times 2.5^\circ$ version. We first discuss the 500 hPa and 925 hPa levels because the temperature trend difference between these two levels is a good measure of the vertical warming structure.

The left panels in Figs. 3–6 show, for each season, the 70–90° N average temperature from the four reanalyses studied. Considering first the 500 hPa level, it is noted that in spring (Fig. 3a) and especially in fall (Fig. 5a) and in winter (Fig. 6a), MERRA is colder than the other reanalyses. In summer the differences between the reanalyses are small, mostly within 0.5 K (Fig. 4a). At the 925 hPa level, NCEP II is systematically colder than the other reanalyses in winter (Fig. 6b). A more striking feature is, however, that in summer, T_{925} for CFSR deviates substantially from the remaining three reanalyses, by -0.5 to -1.7 K, since around the year 2000 (Fig. 4b). This outlier behavior of CFSR is surprising because CFSR is an updated product from NCEP II.

Considering the temperature difference between T_{925} and T_{500} , CFSR again appears as an outlier in summer in the 2000s (Fig. 4c). The low values of T_{925} (Fig. 4b) combined with relatively high T_{500} (Fig. 4a) result in a $T_{925}-T_{500}$ difference that is smaller than in the other reanalyses, by more than 1 K in the mid-2000s (Fig. 4c). Another notable feature in Fig. 4c is that for ERA-Interim, the summertime temperature difference between 925 and 500 hPa increases substantially (by almost 2 K) from late 1990s to early 2000s. This

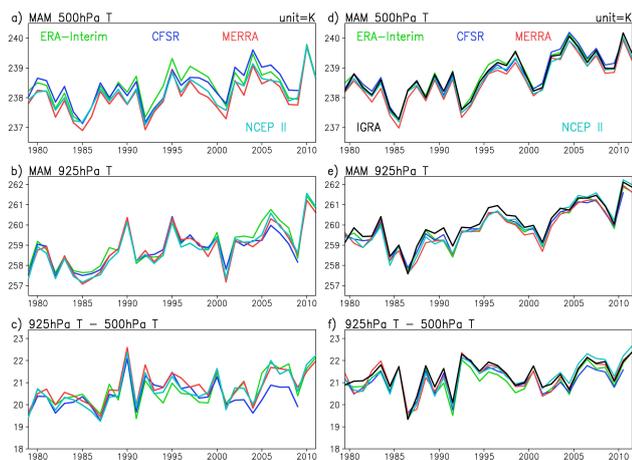


Fig. 3. 70–90° N average T (temperature) from four reanalyses and the IGRA data. On the left, the true 70–90° N average is shown. On the right, the average is calculated using only those reanalysis data that match the available IGRA observations in location and time.

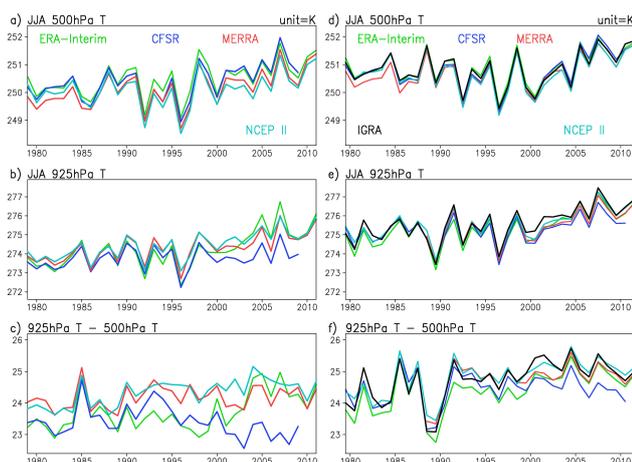


Fig. 4. Same as Fig. 3 except for the JJA (June, July and August) season.

feature is not reproduced by the other reanalyses, and it is mainly linked to a larger increase in T_{925} in ERA-Interim as compared with MERRA, CFSR and NCEP II (Fig. 4b). Furthermore, considering the other seasons, it is noted that in winter the difference between T_{925} and T_{500} is smaller for NCEP II than for the other reanalyses (Fig. 6c) as a result of the relative coldness of NCEP II at 925 hPa (Fig. 6b).

To evaluate the reanalyses, we use the Integrated Global Radiosonde Archive (IGRA) data (Elliott and Gaffen, 1991). IGRA consists of quality-assured soundings over the globe, and has 34 radiosonde stations north of 70° N. IGRA provides monthly and 4 times daily products, but we only use the latter (hereafter referred to as daily IGRA data) for comparing the reanalyses with IGRA data (the IGRA monthly means are not particularly reliable in the Arctic because they are averaged from the available observations, in many cases

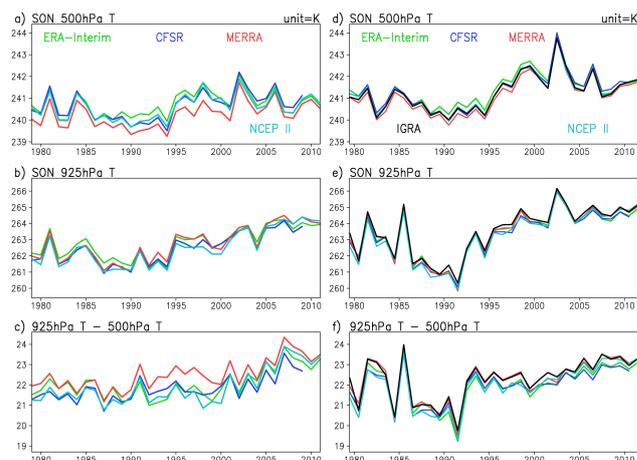


Fig. 5. Same as Fig. 3 except for the SON (September, October and November) season.

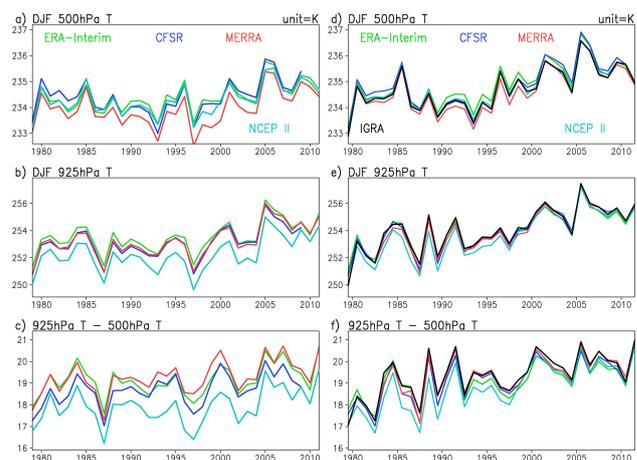


Fig. 6. Same as Fig. 3 except for the DJF (December, January and February) season.

with less than 30 days of data). Moreover, in comparison with the 00Z values of the reanalyses, IGRA soundings within two hours of 00Z are used. Similarly, the 10Z–14Z IGRA soundings are used to evaluate the 12Z reanalysis values. Soundings near 06Z and 18Z are only available from a fraction of the stations and are therefore not included in the analysis below. Naturally, the use of only two observations per day implies that the diurnal cycle is not fully represented. For example, if there are problems specific to a certain part of the diurnal cycle in a reanalysis (such as the representation of stable conditions in the morning, or convective conditions in the afternoon), they could go unnoticed in the analysis. This is mainly a concern when considering data for individual locations (e.g., Fig. 8 below); for the Arctic mean values, this problem is likely alleviated by the fact that different IGRA stations sample different parts of the diurnal cycle.

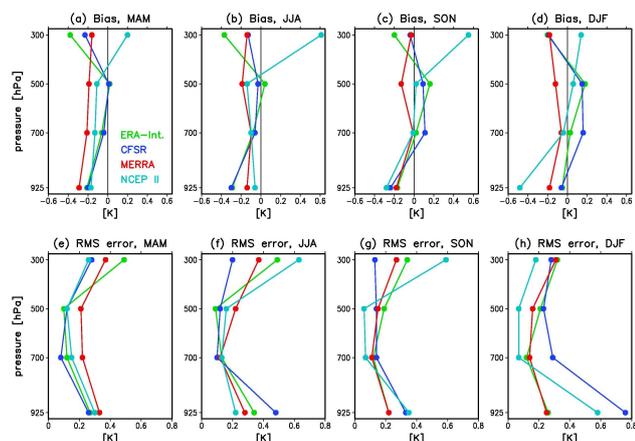


Fig. 7. Mean bias and rms error of seasonal 70–90° N average temperature in each reanalysis, relative to the IGRA data, in units of K. In the computation, we use the reanalysis data that correspond to the available IGRA observations.

In the data processing for reanalysis evaluation, we first linearly interpolated the 00Z and 12Z daily reanalysis data onto the IGRA stations. The interpolated data were then averaged over the station locations north of 70° N over each season for 00Z and for 12Z separately. In this averaging, only the reanalysis data corresponding to the available IGRA data were included. The same averaging procedure was repeated for the IGRA data. Then we averaged the 00Z and 12Z averages, resulting in Tables 1–4 and Figs. 3–7. The averaging of the 00Z and 12Z values was done to avoid a possible bias due to the asymmetry between 00Z and 12Z data volumes; the numbers of 00Z and 12Z soundings for each year and station are not always equal.

The evaluation of the reanalyses, based on the available IGRA observations, is shown in the right panels of Figs. 3–6. The most salient feature of these figures is that the spread between the reanalyses is much smaller than in the left panels. If the spread is quantified in terms of the difference between each reanalysis seasonal Arctic average temperature and the mean over the four reanalyses, the spread decreases by 54 % when averaged over all the seasons and 925, 700, 500 and 300 hPa levels. All the reanalyses incorporated sounding observations, but the reanalyses differ in the extent of other observations incorporated and in the method of doing it. For example, ERA-Interim, CFSR, and MERRA assimilate satellite radiances, whereas NCEP-II assimilates temperature profiles based on satellite data (Saha et al., 2010; Rienecker et al., 2011; Dee et al., 2011). Differences may also exist in the assimilation method of sounding observations (Lüpkes et al., 2010). It is therefore expected that the reanalyses agree better with each other over the available IGRA data than for the true Arctic-mean values (i.e., the average over the entire area north of 70° N). Over the central Arctic regions lacking sounding data, the reanalysis systems have more freedom to

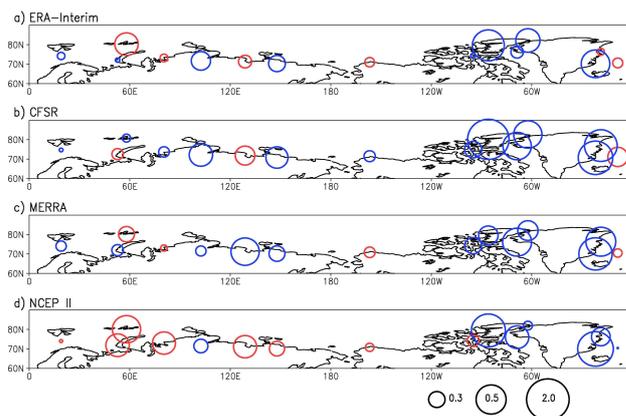


Fig. 8. Evaluation of the reanalyses with the IGRA data in June–August 2000–2009 average 925 hPa T . The reanalysis data used here match the available IGRA observations in location and time. Shown is reanalysis temperature bias at each IGRA station in units of K, with the circle size indicating the magnitude of the bias. When the bias is positive (negative), the circle is colored red (blue).

form their own climate, and thus biases. While spreads between the reanalyses are generally smaller over the available IGRA data, CFSR shows a relatively larger error in summer at 925 hPa since 2000 (Fig. 4e). CFSR shows a cold bias of 0.4–0.8 K at 925 hPa in summer since the year 2000 (Fig. 4e).

The temperature difference between the two levels ($dT = T_{925} - T_{500}$) is shown in the bottom panels of Figs. 3–6. In summer, CFSR shows the largest deviations from the IGRA data, dT for CFSR being less than that for IGRA by almost 1 K for many years in the 2000s (Fig. 4f). In the 1990s, ERA-interim deviates from the IGRA data most, and the deviation is about 0.5 K. ERA's rapid increase of dT for the true Arctic average from the late 1990s to the early 2000s, as shown in Fig. 4c, is significantly reduced in Fig. 4f.

We now extend our investigation to the 300 and 700 hPa levels. Figure 7a–d display, for each reanalysis and season, the bias compared to IGRA data at the 300, 500, 700, and 925 hPa levels, and Fig. 7e–h show the corresponding rms (root mean square) errors. The bias and rms error are defined for the seasonal and Arctic averages of temperature (see Fig. 7). The exact numerical values are provided in Tables 1–2. Different reanalyses exhibit different degrees of bias and rms error. The disagreement between the reanalyses is often particularly large in the lower and upper troposphere (925 and 300 hPa) compared to the middle troposphere (500 and 700 hPa), as shown in Fig. 7.

Looking into seasonal and true Arctic average temperatures at the 300 and 700 hPa levels, the spread between the reanalyses often exceeds 1 K (not shown). When the Arctic average is made over the region with available IGRA data, the spread between the reanalyses is due primarily to an outlier reanalysis. In addition to the aforementioned CFSR behavior at 925 hPa in summer, NCEP II is an outlier in summer

Table 1. The rms error of seasonal and 70–90° N average temperature in the reanalyses in units of K. There are four values separated by “|”, and these four values represent ERA|CFSR|MERRA|NCEP II. The rms error here is between each reanalysis and the IGRA data. To compute the error, we use the reanalysis data that correspond to the available IGRA observations. 00Z and 12Z daily data are used instead of monthly means.

	MAM				JJA				SON				DJF			
	E	C	M	N	E	C	M	N	E	C	M	N	E	C	M	N
300 hPa	0.49	0.28	0.37	0.26	0.49	0.20	0.37	0.63	0.34	0.13	0.27	0.59	0.32	0.28	0.31	0.18
500 hPa	0.10	0.12	0.21	0.12	0.09	0.12	0.22	0.16	0.19	0.14	0.15	0.06	0.21	0.23	0.16	0.07
700 hPa	0.12	0.08	0.22	0.15	0.13	0.10	0.12	0.13	0.12	0.14	0.11	0.07	0.12	0.29	0.14	0.07
925 hPa	0.27	0.26	0.33	0.30	0.34	0.48	0.28	0.22	0.22	0.33	0.22	0.35	0.26	0.76	0.25	0.58

Table 2. Same as Table 1, except for mean bias of seasonal and 70–90° N average temperature of each reanalysis product, relative to the IGRA data, in units of K.

	MAM				JJA				SON				DJF			
	E	C	M	N	E	C	M	N	E	C	M	N	E	C	M	N
300 hPa	-0.38	-0.23	-0.16	+0.20	-0.37	-0.13	-0.14	+0.61	-0.20	-0.04	-0.03	+0.55	-0.20	-0.19	-0.18	+0.14
500 hPa	+0.02	+0.01	-0.19	-0.11	+0.04	-0.03	-0.19	-0.14	+0.16	+0.09	-0.13	+0.02	+0.18	+0.15	-0.12	+0.06
700 hPa	-0.06	-0.04	-0.21	-0.13	-0.08	-0.06	-0.10	-0.10	+0.02	+0.11	-0.01	-0.01	+0.03	+0.16	-0.06	-0.04
925 hPa	-0.21	-0.20	-0.29	-0.17	-0.29	-0.30	-0.14	-0.06	-0.17	-0.24	-0.18	-0.28	-0.05	-0.06	-0.18	-0.48

and fall at 300 hPa. It has a substantial warm bias of about 0.6 K compared to IGRA data, while the other reanalyses show smaller negative biases (−0.37 to −0.03 K). In view of this NCEP II deficiency and the CFSR error in summer at 925 hPa, we note that an outlier reanalysis product tends to correspond to the most erroneous product. Prior to the year 2000 this tendency is, however, less obvious. The tendency we noted here is based on the chosen four reanalyses.

It is further noted from Fig. 7a–d that all the reanalyses have a cold bias at 925 hPa. Also, MERRA has a cold bias at all the 4 levels and 4 seasons. The reasons for this cold bias tendency might be related to free-running climate models having a cold bias in the Arctic. Note that our evaluation of the reanalyses is limited to the use of the IGRA data, which must have been largely incorporated into each reanalysis product. Using sounding observations that are not assimilated over the Arctic, Jakobson et al. (2012) also found the CFSR to perform the worst at 925 hPa (roughly 600 m). However, they found CFSR to be the best reanalysis for temperature in the lowermost 100 m layer, indicating that the performance of each reanalysis depends strongly on altitude.

We further discuss the aforementioned error in CFSR in summer at 925 hPa since 2000 (Fig. 4e). Figure 7b and f confirm that in terms of Arctic-mean values, the old NCEP II reanalysis is superior to the new CFSR in summer at 925 hPa. To better understand this surprising result, we show in Fig. 8 the difference between the reanalysis summertime T_{925} and the IGRA value for each station. The most conspicuous feature of Fig. 8 is that all the reanalyses have negative (cold) biases in the Canadian-Atlantic sector between 120° W and 0°, and less negative or slightly positive biases in the other

longitudes. We found that this bias is not directly related to interpolation of data below the earth’s surface since the surface altitudes of the IGRA stations and those of the nearby reanalysis grid cells are safely below the 925 hPa level. Out of the four reanalyses, the CFSR and NCEP II have particularly negative biases over the 120° W–0° band compared to the bias over the other longitudes (Fig. 8). While the negative bias over the 120° W–0° band is largely canceled out by the positive bias over the other longitudes in the NCEP II, the bias is still negative outside of the 120° W–0° band in the CFSR. In this regard, the better performance of NCEP II for the Arctic mean values is fortuitous.

4 Temperature trends

In this section, we quantify temperature trends at selected levels. Considering first the surface level, the reanalyses and GISTEMP both show clear warming trends from 1979 to 2011 (Fig. 2), and Table 3 quantifies them. For the recent period (i.e., since around 1998), however, the warming trends in GISTEMP agree very well with those in the reanalyses, except in summer (Fig. 2f–j). In summer (Fig. 2h), GISTEMP shows no clear trend since 1998, while all the reanalyses show a warming trend. Hence, we need to have a closer look at the period 1998–2011. The trends of the time series of Figs. 2–6 are computed for the 1998–2011 period in Table 4 and Figs. 9–10. Reverting back to the surface temperature trend, the summertime warming trends in the reanalyses appear for the true Arctic averages as well as when only those data corresponding to the available GISTEMP data are sampled, while GISTEMP does not show a warming trend

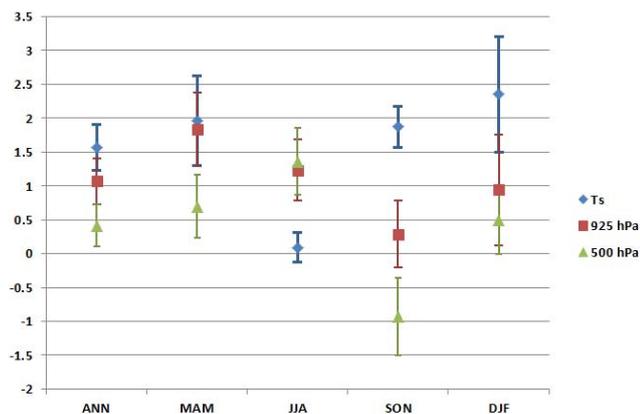


Fig. 9. 70–90° N average temperature trend from 1998 to 2011 in units of change (in K) over the 14 yr, as in Table 4. All the trends here are from pure observations (either GISTEMP or IGRA data) and not from the reanalyses. The error bar represents \pm standard error of the trend.

(Table 4). On the contrary, in spring, autumn and winter, the reanalyses and GISTEMP all show significant warming trends. When annual mean trends are analyzed, all the reanalyses agree with the GISTEMP trend of about $1.6 \text{ K } 14 \text{ yr}^{-1}$, except for ERA-Interim (about $2.1 \text{ K } 14 \text{ yr}^{-1}$). We also note that both GISTEMP and the reanalyses show a large increase ($\sim 1.5 \text{ K}$) in summertime temperature from 1996 to 1998, indicating that the trends are sensitive to the choice of the period considered.

Figure 9 summarizes the observed trends from 1998 to 2011. In summer, the T_s trend is near zero while upper-air temperature trends are positive. Also, the warming at 925 hPa is almost equal to that at 500 hPa. In the other 3 seasons, on the other hand, the warming tendency is larger at lower altitude. Fall is particularly interesting, since the surface shows a very large warming when there has been a statistically significant cooling trend at 500 hPa. The reanalyses agree on the cooling trend at 500 hPa in fall, with the trends of -0.74 K to $-0.03 \text{ K } 14 \text{ yr}^{-1}$ for true Arctic 70–90° N mean values, and with the trends of -1.25 K to $-0.91 \text{ K } 14 \text{ yr}^{-1}$ when sampled according to the availability of IGRA data (Table 4). On the other hand, analyzing the entire 1979–2011 period reveals that the trend is not necessarily negative in fall and the warming tendency does not necessarily become larger at lower altitude (Table 3). In fact, for the period 1979–2011, the layers at 500 hPa and below all show a warming trend in all the chosen reanalyses and observations (Table 3). This raises a hypothesis that the recent warming is associated with different mechanisms than the earlier warming.

The trends shown in Tables 3 and 4 and Fig. 9 can be viewed as originating from one realization sampled from an ensemble of many time series. We assess the significance of the calculated trends using a null hypothesis of an ensemble of no-trend time series. Student's t test (Santer et al., 2000)

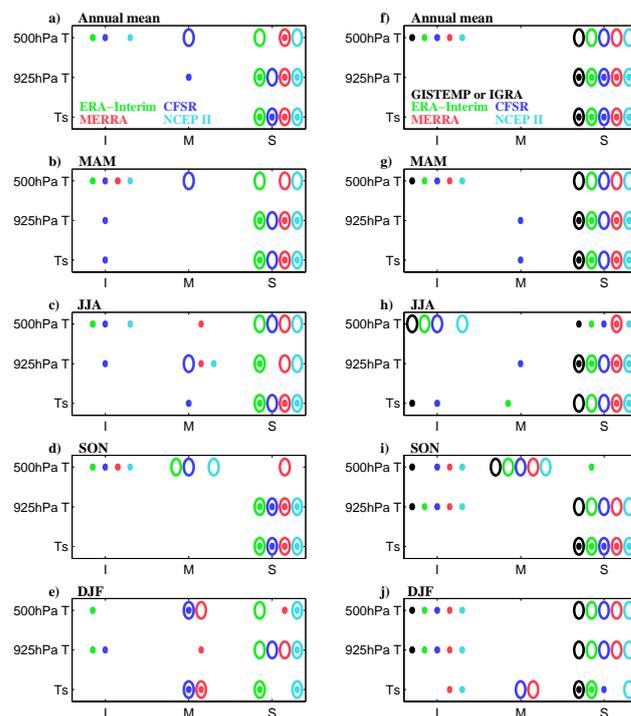


Fig. 10. Significance of the trends in Tables 3 and 4. Significance at the 95 % level is indicated by an oval for the 1979–2011 trend and by a dot for the 1998–2011 trend. When an oval (or dot) is located above I, M or S, it means that the trend is insignificant, may be significant or is significant, respectively. We apply three different significance testing techniques. When all of the techniques establish significance (or insignificance), we determine that the trend is significant (or insignificant). When the techniques give conflicting results, the trend is considered “maybe significant”. All tests are two-tailed tests. The left panels represent the significance of the trends from the true Arctic-mean time series while the right panels address the trends with only the reanalysis data corresponding to the available IGRA or GISTEMP data.

is commonly applied to establish the significance, but this technique does not account for autocorrelation in the time series. Thus, we choose to apply Student's t test after using the prewhitening technique of Wang and Swail (2001) and Student's t test with effective sample size correction (Santer et al., 2000), as well as normal Student's t test. The former two techniques factor in autocorrelation. If all of the three techniques establish the significance, we consider the trend significant, as explained in Fig. 10. Figure 10 shows that the 1979–2011 trends are mostly significant while the 1998–2011 trends are often insignificant. Also, the trends from the true Arctic-mean time series (Fig. 10a–e) tend to be more significant than those from using only the reanalysis data corresponding to the available IGRA or GISTEMP data (Fig. 10f–j). This makes sense because more data would lead to more significant results.

We also assess the significance of the difference in the trends. This can be done in two ways. One way is to construct

Table 3. 70–90° N average temperature trend from 1979 to 2011 in units of temperature change (in K) over the 33 yr. In the case of CFSR, the trend is from 1979 to 2009 in units of temperature change over the 31 yr. There are five values separated by “|”, and these five values represent Observation|ERA|CFSR|MERRA|NCEP II. Observation here refers to either GISTEMP or IGRA data. In the rows with observation (i.e., five values separated by a blank), the 70–90° N average is made with the reanalysis data corresponding to the available observation, whereas in the rows without observation (i.e., four values separated by a blank), true 70–90° N average values are used. The standard error of each trend is shown with \pm in each row.

	Annual mean					JJA					DJF					MAM					SON				
	O	E	C	M	N	O	E	C	M	N	O	E	C	M	N	O	E	C	M	N	O	E	C	M	N
T_s	2.09 2.52 1.75 1.84 2.71	1.05 0.86 0.76 1.01 1.09	2.22 3.03 1.91 1.63 3.16	2.21 3.00 1.94 2.19 2.72	2.87 3.17 2.40 2.52 3.85	\pm .29 \pm .31 \pm .29 \pm .26 \pm .28	\pm .23 \pm .18 \pm .21 \pm .17 \pm .19	\pm .62 \pm .71 \pm .71 \pm .63 \pm .67	\pm .47 \pm .46 \pm .43 \pm .36 \pm .40	\pm .39 \pm .43 \pm .41 \pm .35 \pm .45	2.43 1.78 1.73 2.83	0.71 0.51 0.89 1.18	2.86 2.05 1.55 3.01	2.97 2.12 1.76 3.02	3.19 2.43 2.73 4.11	\pm .34 \pm .32 \pm .27 \pm .31	\pm .17 \pm .20 \pm .15 \pm .18	\pm .73 \pm .72 \pm .64 \pm .71	\pm .50 \pm .50 \pm .39 \pm .47	\pm .45 \pm .44 \pm .37 \pm .46					
	925 hPa T	2.77 2.61 2.41 2.63 3.10	1.78 1.84 0.88 1.56 1.53	3.83 3.30 3.73 3.37 4.43	2.42 2.37 2.12 2.48 2.94	3.07 2.94 2.90 3.11 3.49	\pm .31 \pm .32 \pm .33 \pm .32 \pm .33	\pm .42 \pm .43 \pm .41 \pm .40 \pm .41	\pm .66 \pm .63 \pm .66 \pm .65 \pm .63	\pm .40 \pm .40 \pm .37 \pm .42 \pm .40	\pm .71 \pm .71 \pm .71 \pm .72 \pm .70	2.10 1.50 2.12 2.21	2.06 0.53 1.49 1.60	1.95 1.99 2.26 2.05	2.28 1.30 2.13 2.47	2.12 2.20 2.62 2.74	\pm .32 \pm .27 \pm .28 \pm .31	\pm .44 \pm .34 \pm .34 \pm .31	\pm .57 \pm .60 \pm .58 \pm .64	\pm .46 \pm .43 \pm .49 \pm .47	\pm .41 \pm .40 \pm .42 \pm .43				
500 hPa T	1.36 1.34 1.41 1.42 1.29	0.57 0.61 0.60 0.81 0.43	1.75 1.74 1.87 1.64 1.72	1.65 1.58 1.66 1.68 1.61	1.47 1.42 1.51 1.57 1.39	\pm .21 \pm .20 \pm .25 \pm .22 \pm .21	\pm .36 \pm .35 \pm .37 \pm .35 \pm .38	\pm .37 \pm .36 \pm .42 \pm .39 \pm .37	\pm .30 \pm .30 \pm .31 \pm .30 \pm .29	\pm .42 \pm .42 \pm .45 \pm .43 \pm .41	0.90 0.90 0.98 0.84	1.01 1.04 1.22 0.71	0.97 0.80 0.88 0.94	0.94 0.87 0.97 0.98	0.67 0.87 0.88 0.75	\pm .17 \pm .23 \pm .19 \pm .18	\pm .32 \pm .35 \pm .32 \pm .35	\pm .29 \pm .38 \pm .37 \pm .31	\pm .32 \pm .31 \pm .32 \pm .28	\pm .31 \pm .36 \pm .32 \pm .32					

Table 4. 70–90° N average temperature trend from 1998 to 2011 in units of temperature change (in K) over the 14 yr. In the case of CFSR, the trend is from 1998 to 2009 in units of temperature change over the 12 yr.

	Annual mean					JJA					DJF					MAM					SON				
	O	E	C	M	N	O	E	C	M	N	O	E	C	M	N	O	E	C	M	N	O	E	C	M	N
T_s	1.57 2.06 1.55 1.36 1.59	0.09 0.42 0.45 0.43 0.74	2.35 3.13 2.70 1.72 1.93	1.96 2.33 1.10 1.39 1.50	1.87 2.37 1.94 1.90 2.20	\pm .34 \pm .37 \pm .41 \pm .34 \pm .32	\pm .22 \pm .20 \pm .26 \pm .18 \pm .16	\pm .85 \pm .1.1 \pm .1.1 \pm .1.0 \pm .1.0	\pm .66 \pm .66 \pm .59 \pm .48 \pm .57	\pm .30 \pm .27 \pm .36 \pm .28 \pm .25	2.25 1.60 1.52 1.89	0.49 0.45 0.49 0.80	3.57 3.00 2.32 2.58	2.59 1.14 1.53 1.84	2.35 1.81 1.74 2.33	\pm .37 \pm .44 \pm .34 \pm .32	\pm .18 \pm .26 \pm .17 \pm .19	\pm .1.0 \pm .1.1 \pm .98 \pm .1.0	\pm .70 \pm .70 \pm .52 \pm .66	\pm .37 \pm .51 \pm .40 \pm .37					
	925 hPa T	1.07 1.06 0.83 1.33 1.35	1.23 1.55 0.93 1.39 1.54	0.94 0.72 0.80 1.27 1.01	1.84 1.86 1.45 2.13 2.16	0.29 0.13 0.12 0.51 0.71	\pm .34 \pm .36 \pm .35 \pm .33 \pm .33	\pm .45 \pm .47 \pm .50 \pm .47 \pm .45	\pm .82 \pm .83 \pm .83 \pm .83 \pm .77	\pm .53 \pm .55 \pm .53 \pm .54 \pm .54	\pm .50 \pm .50 \pm .55 \pm .50 \pm .50	1.61 0.79 1.59 1.90	1.70 0.09 0.85 0.81	1.62 1.74 2.00 2.46	1.86 0.19 1.92 2.11	1.28 1.15 1.60 2.20	\pm .41 \pm .35 \pm .34 \pm .32	\pm .56 \pm .47 \pm .48 \pm .38	\pm .83 \pm .85 \pm .81 \pm .83	\pm .72 \pm .56 \pm .70 \pm .72	\pm .36 \pm .39 \pm .44 \pm .41				
500 hPa T	0.41 0.14 0.48 0.40 0.44	1.36 1.17 1.28 1.26 1.34	0.50 0.17 0.77 0.59 0.56	0.70 0.49 0.83 0.71 0.77	−.93 −.125 −.96 −.94 −.91	\pm .31 \pm .30 \pm .33 \pm .30 \pm .30	\pm .49 \pm .49 \pm .51 \pm .48 \pm .49	\pm .51 \pm .52 \pm .52 \pm .54 \pm .49	\pm .47 \pm .45 \pm .46 \pm .44 \pm .44	\pm .57 \pm .55 \pm .60 \pm .58 \pm .55	0.14 0.38 0.74 0.48	0.41 0.50 0.71 0.61	0.63 1.14 1.44 1.13	0.25 0.18 0.83 0.70	−.74 −.29 −.03 −.51	\pm .26 \pm .28 \pm .27 \pm .24	\pm .44 \pm .47 \pm .44 \pm .45	\pm .43 \pm .47 \pm .47 \pm .39	\pm .50 \pm .48 \pm .50 \pm .46	\pm .40 \pm .48 \pm .44 \pm .41					

confidence intervals for individual trends and then determine the overlap between a pair of confidence intervals. A 95 % confidence interval would contain the true trend at a 95 % probability. Thus, when two confidence intervals overlap with each other, the two corresponding trends might have come from the same population. We find confidence interval overlaps in most of the differences in the trends, indicating that the trend differences are too small. Santer et al. (2001) found similar results for the trends they calculated. The other way to assess the significance of a trend difference is to determine the significance of the trend of the time series difference. The advantage of this method is that variability common to the two time series is removed. Applying this method reveals that the trend differences are a good mixture of significant and insignificant differences. Again, the trend difference tends to be more significant for the true Arctic-mean time series.

The spatial distribution of recent summertime temperature trends for GISTEMP and the reanalyses is shown in Fig. 11. The near-zero summertime trend in GISTEMP is related to a negative temperature trend in the Eurasian sector between 20° and 110° E. This cooling trend is not well reproduced in the reanalyses. Also, over the ocean, the reanalyses noticeably differ from GISTEMP (Fig. 11), but we have to bear in mind that, using the 250 km smoothing, there are very few GISTEMP values in the central Arctic. In the sea areas that are open only in summer and autumn (e.g., parts of the Barents, Kara, Laptev, East Siberian, Chukchi, and Beaufort Seas), the GISTEMP data are probably less reliable than the reanalyses for these seasons. This is because (a) reanalyses effectively assimilate satellite SST observations from the seasonally ice-free seas, whereas GISTEMP applies SST observations only from permanently ice-free areas, and (b) the

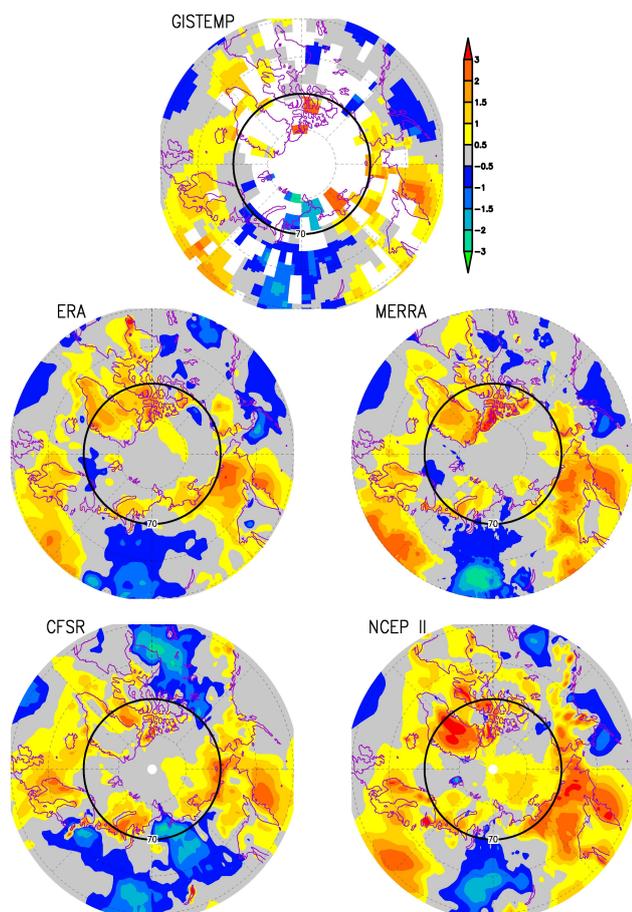


Fig. 11. June–August average T_s trend from 1998 to 2011 in units of change (in K) over the 14 yr for GISTEMP and for the four reanalyses. In the case of CFSR, the trend is from 1998 to 2009 in units of change over the 12 yr. White areas in the GISTEMP plot indicate missing data or insufficient data for the trend analysis. Note that T_s refers to surface air temperature (SAT) for the reanalyses, and to a combination of SAT over land and SST over ocean for GISTEMP. The 70° N circle is thickened in each panel.

GISTEMP values seen in these sea areas are based on the extrapolation of land observations.

With regard to the differences between GISTEMP and reanalyses over the central Arctic (the area where GISTEMP has very few data), one might simply think that reanalyses are more accurate because they assimilate satellite radiances or satellite-based upper-air temperatures, and those should affect the surface air temperatures. However, satellite retrievals have a coarse vertical resolution and cannot well represent the fine structure of temperature in the shallow atmospheric boundary layer in the Arctic. Due to the close coupling of the open ocean and near-surface air, we believe that the surface air temperature in reanalyses is more affected by the assimilation of satellite data on SST than upper-air temperatures.

In addition to differences in satellite data usage, another potential reason for the detected discrepancies (Figs. 2h and

10) is that in GISTEMP T_s refers to a combination of SST and surface air temperature, whereas in reanalyses T_s refers to surface air temperature (SAT). Even though SST observations are assimilated in reanalyses, SAT and SST are not perfectly correlated, and so their trends are not necessarily equal.

5 Discussion

As noted in the introduction, various factors can cause differences between temperature fields in different reanalyses. These include differences in the usage of observations, differences in the atmospheric model used in producing the reanalysis, and the methodology used for data assimilation. In general, it is difficult to pinpoint the exact causes for the differences between different reanalysis temperatures and their trends. However, some general comments can be made. First, as noted in Sect. 3, the spread between reanalysis upper-air temperatures is generally smaller over those regions where IGRA radiosounding data are available. This is expected because most or all of the radiosounding data must have been assimilated in the reanalyses, thereby constraining them. Over the central Arctic regions lacking sounding data, the reanalysis systems have more freedom to form their own climate, and thus biases.

In contrast, for surface air temperature, restricting the analysis to the locations with available GISTEMP data did not reduce the spread between the different reanalysis. As pointed out in Sect. 2, this might be explained by the fact that among the reanalyses considered, only ERA-Interim assimilates surface air temperatures from land stations. At the surface level, there are additional observations that can be used to validate reanalyses. Makshtas et al. (2007) used the observations from Russian drifting stations and Alexeev et al. (2012) used drifting Argo buoys to validate reanalyses over the Arctic. Testing the accuracy of each reanalysis product can perhaps be better carried at the surface level.

While all the reanalyses agree that the Arctic is, in general, warming, they show substantial differences in the details, such as the vertical structure of summertime warming (as characterized by the time series of the temperature difference between 925 and 500 hPa in Fig. 4c). There are two general causes that can explain time-varying differences between the reanalyses. First, the reanalysis systems may respond differently to real changes in the Arctic environment, such as diminishing Arctic sea ice. For instance, the performance of the atmospheric models used in producing the reanalyses may depend on the surface status. Surface status can also vary among the reanalyses despite the assimilation of satellite-based sea ice extent data. The differences may originate from the remote sensing algorithm applied (Valkonen et al., 2008) and because out of the reanalyses included in this study only in CFSR is the atmospheric model coupled with an ocean model and a dynamic-thermodynamic sea ice

model (Saha et al., 2010). Second, while the models used for reanalyses are “frozen”, the availability of observations has changed during the reanalysis period, in particular, due to the increasing amounts of satellite data. Changes in the observation systems may have caused artificial trends and shifts in the reanalyses, such as those demonstrated for ERA-40 by Screen and Simmonds (2011). Overall, while it is beyond the scope of the present work to unravel the exact causes for the differences between the reanalyses, there is clearly a need for such studies.

Despite the aforementioned differences between the reanalyses, there is an optimistic aspect to our results. Alexeev et al. (2012) compared earlier reanalysis versions. The disagreement between ERA-40 and NCEP-I at 925 hPa in winter (as in Fig. 3 of their study) is much greater than the disagreement between the reanalyses we have analyzed (as shown in Fig. 6b). This indicates that as reanalyses are updated, they tend to become in greater agreement with each other.

6 Concluding remark

In this study, we have examined the Arctic warming structure from 1979 to 2011 in observations and reanalyses. Our analysis shows that the warming structure in the recent period (roughly from 1998) differs greatly from that in the earlier period. In the recent period, the surface was not clearly warming in summer, and 500 hPa air became colder in autumn. Before 1998, however, all the layers at 500 hPa or below were warming. These findings are supported by multiple reanalyses together with observations. The cooling trend in both observations and the reanalyses at 500 hPa in fall since 1998 has not been sufficiently recognized in earlier studies.

While examining the warming structure, we have also examined the validity of the reanalyses in surface-air and upper-level temperatures. We have shown for the 1998–2011 period that all the reanalyses reveal warming trends at the surface in all the seasons, while GISTEMP shows no trend in summer. At the 925 hPa level, CFSR shows larger Arctic-mean bias than NCEP II in summer (especially after the year 2000), although CFSR is supposedly an improvement over the NCEP II. This CFSR error was shown to arise primarily from the 120° W–0° longitude band. It is our hope that this study will stimulate further investigations into the root cause of the reanalysis errors and biases. Our results also suggest that studies of the Arctic climate based on reanalyses should be undertaken with extreme caution.

Acknowledgements. The authors are thankful to the NCDC for providing the IGRA data, to the NOAA/OAR/ESRL PSD for providing the GISTEMP data, and to the ECMWF, NCEP, and NASA for providing the reanalysis data. This research was supported by the Korea Meteorological Administration Research and Development Program (CATER 2012-7100). The work of TV

was supported by the Academy of Finland through the CACSI project (contract: 259537).

Edited by: K. Dethloff

References

- Alexeev, V. A., Langen, P. L., and Bates, J. R.: Polar amplification of surface warming on an aquaplanet in “ghost forcing” experiments without sea ice feedbacks, *Clim. Dynam.*, 24, 655–666, doi:10.1007/s00382-005-0018-3, 2005.
- Alexeev, V. A., Esau, I., Polyakov, I. V., Byam, S. J., and Sorokina, S.: Vertical structure of recent arctic warming from observed data and reanalysis products, *Climatic Change*, 111, 215–239, 2012.
- Bekryaev, R. V., Polyakov, I. V., and Alexeev, V. A.: Role of Polar Amplification in Long-Term Surface Air Temperature Variations and Modern Arctic Warming, *J. Climate*, 23, 3888–3906, doi:10.1175/2010JCLI3297.1, 2010.
- Bromwich, D. H., Nicolas, J. P., and Monaghan, A. J.: An Assessment of precipitation changes over antarctica and the southern ocean since 1989 in contemporary global reanalyses, *J. Climate*, 24, 4189–4209, 2011.
- Chung, C. E. and Räisänen, P.: Origin of the Arctic warming in climate models, *Geophys. Res. Lett.*, 38, L21704, doi:10.1029/2011GL049816, 2011.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavalato, C., Thépaut, J. N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. Roy. Meteor. Soc.*, 137, 553–597, 2011.
- Elliott, W. P. and Gaffen, D. J.: On the utility of radiosonde humidity archives for climate studies, *B. Am. Meteorol. Soc.*, 72, 1507–1520, 1991.
- Graversen, R. G., Mauritsen, T., Tjernstrom, M., Kallen, E., and Svensson, G.: Vertical structure of recent Arctic warming, *Nature*, 451, 53–56, doi:10.1038/nature06502, 2008.
- Hall, A.: The Role of Surface Albedo Feedback in Climate, *J. Climate*, 17, 1550–1568, doi:10.1175/1520-0442(2004)017<1550:trosaf>2.0.co;2, 2004.
- Hansen, J., Ruedy, R., Sato, M., and Lo, K.: Global surface temperature change, *Rev. Geophys.*, 48, RG4004, doi:10.1029/2010RG000345, 2010.
- Hassel, S. J.: Impacts of a Warming Climate – Arctic Climate Impact Assessment Cambridge University Press, New York, 144 pp., 2004.
- Jakobson, E., Vihma, T., Palo, T., Jakobson, L., Keernik, H., and Jaagus, J.: Validation of atmospheric reanalyses over the central Arctic Ocean, *Geophys. Res. Lett.*, 39, L10802, doi:10.1029/2012GL051591, 2012.
- Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S. K., Hnilo, J. J., Fiorino, M., and Potter, G. L.: NCEP-DOE AMIP-II reanalysis (R-2), *B. Am. Meteorol. Soc.*, 83, 1631–1643+1559, 2002.

- Lüpkes, C., Vihma, T., Jakobson, E., König-Langlo, G., and Tetzlaff, A.: Meteorological observations from ship cruises during summer to the central Arctic: A comparison with reanalysis data, *Geophys. Res. Lett.*, 37, L09810, doi:10.1029/2010gl042724, 2010.
- Makshtas A., D. Atkinson, M. Kulakov, S. Shutilin, R. Krishfield, and A. Proshutinsky, 2007: Atmospheric forcing validation for modeling the central Arctic, *Geophys. Res. Lett.*, 37, L20706, doi:10.1029/2007GL031378, 2007.
- Manabe, S.: Carbon Dioxide and Climatic Change, in: *Advances in Geophysics*, edited by: Barry, S., Elsevier, Cambridge, MA, 39–82, 1983.
- Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G. K., Bloom, S., Chen, J., Collins, D., Conaty, A., Da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and Woollen, J.: MERRA: NASA's modern-era retrospective analysis for research and applications, *J. Climate*, 24, 3624–3648, 2011.
- 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., 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., Schemm, J. K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C. Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R. W., Rutledge, G., and Goldberg, M.: The NCEP climate forecast system reanalysis, *B. Am. Meteorol. Soc.*, 91, 1015–1057, 2010.
- Santer, B. D., Wigley, T. M. L., Boyle, J. S., Gaffen, D. J., Hnilo, J. J., Nychka, D., Parker, D. E., and Taylor, K. E.: Statistical significance of trends and trend differences in layer-average atmospheric temperature time series, *J. Geophys. Res.-Atmos.*, 105, 7337–7356, doi:10.1029/1999jd901105, 2000.
- Screen, J. A. and Simmonds, I.: The central role of diminishing sea ice in recent Arctic temperature amplification, *Nature*, 464, 1334–1337, doi:10.1038/nature09051, 2010.
- Screen, J. A. and Simmonds, I.: Erroneous Arctic Temperature Trends in the ERA-40 Reanalysis: A Closer Look, *J. Climate*, 24, 2620–2627, doi:10.1175/2010jcli4054.1, 2011.
- Screen, J. A., Deser, C., and Simmonds, I.: Local and remote controls on observed Arctic warming, *Geophys. Res. Lett.*, 39, L10709, doi:10.1029/2012GL051598, 2012.
- Valkonen, T., Vihma, T., and Doble, M.: Mesoscale modeling of the atmosphere over Antarctic sea ice: A late-autumn case study, *Mon. Weather Rev.*, 136, 1457–1474, 2008.
- Wang, X. L. and Swail, V. R.: Changes of extreme wave heights in Northern Hemisphere oceans and related atmospheric circulation regimes, *J. Climate*, 14, 2204–2221, 2001.
- Wang, W., Xie, P., Yoo, S. H., Xue, Y., Kumar, A., and Wu, X.: An assessment of the surface climate in the NCEP climate forecast system reanalysis, *Clim. Dynam.*, 37, 1601–1620, 2011.