

Strong Trends in the Skill of the ERA-40 and NCEP–NCAR Reanalyses in the High and Midlatitudes of the Southern Hemisphere, 1958–2001*

DAVID H. BROMWICH AND RYAN L. FOGT

Polar Meteorology Group, Byrd Polar Research Center, and Atmospheric Sciences Program, Department of Geography, The Ohio State University, Columbus, Ohio

(Manuscript received 20 February 2004, in final form 21 June 2004)

ABSTRACT

The European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40) and the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis (NCEP1) data are compared with Antarctic and other mid- to high-latitude station observations for the complete years of overlap, 1958–2001. Overall, it appears that ERA-40 more closely follows the observations; however, a more detailed look at the presatellite era reveals many shortcomings in ERA-40, particularly in the austral winter.

By calculating statistics in 5-yr moving windows for June–July–August (JJA), it is shown that ERA-40 correlations with observed MSLP and surface (2 m) temperatures are low and even negative during the mid-1960s. A significant trend in skill in ERA-40 is observed in conjunction with the assimilation of satellite data during winter, eventually reaching a high level of skill after 1978 that is superior to NCEP1. NCEP1 shows consistency in its correlation with observations throughout time in this season; however, the biases in the NCEP1 MSLP fields decrease significantly with time. Similar problems are also found in the 500-hPa geopotential height fields above the direct influences of the mountainous topography. The height differences between ERA-40 and NCEP1 over the South Pacific are substantial before the modern satellite era throughout the depth of the troposphere. The ability for ERA-40 to be more strongly constrained by the satellite data compared to NCEP1, which is largely constrained by the station observational network, suggests that the differing assimilation schemes between ERA-40 and NCEP1 lead to the large discrepancies seen here. Thus, both reanalyses must be used with caution over high southern latitudes *during the nonsummer months* prior to the assimilation of satellite sounding data.

1. Introduction

The Southern Ocean and the Antarctic continent represent perhaps the largest spatial meteorological data voids on the globe. Climate analysis over the high southern latitudes is limited to the sparse station network, making it challenging to resolve climate signals in the high southern latitudes. Recently, various satellite datasets have become more widely distributed and employed, greatly helping to eliminate the large spatial data gaps in these regions. Satellite-derived surface temperatures (e.g., Comiso 2000), sea ice concentrations derived from passive microwave radiometers (e.g., Zwally et al. 2002), and various cloud and radiation products (including cloud track winds; Pavolonis and Key 2003; Key et al. 2003) derived from the Advanced Very High

Resolution Radiometer (AVHRR) and the Moderate-Resolution Imaging Spectroradiometer (MODIS) data are just a few of the datasets utilized in recent research. However, most satellite data only extend back to 1978, forcing studies involving preceding years to largely depend on the sparse surface station network.

The National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis (hereafter NCEP1; Kistler et al. 2001; Kalnay et al. 1996) project helped to solve the problem of the large data voids, particularly before the availability of the satellite data. NCEP1, unlike many of the available analyses, has the positive benefit of a fixed state-of-the-art assimilation scheme. With more observations included and the better quality control in NCEP1, these reanalysis fields were thought to be the means by which climate studies could finally be conducted over the entire southern high latitudes starting from the International Geophysical Year (IGY: 1957–58).

Since their initial release to the public in 1996 (Kalnay et al. 1996), NCEP1 has been widely used for many climate studies in the high latitudes. Recently, NCEP1

* Byrd Polar Research Center Contribution Number 1300.

Corresponding author address: David H. Bromwich, Polar Meteorology Group, Byrd Polar Research Center, The Ohio State University, 1090 Carmack Road, Columbus, OH 43210.
E-mail: bromwich@polarmet1.mps.ohio-state.edu

has been utilized across Antarctica to study: the katabatic winds (Parish and Cassano 2003); trends in the circumpolar vortex/Antarctic Oscillation (Thompson and Solomon 2002; Marshall 2003; Jones and Widmann 2003); the surface energy budget (Renfrew et al. 2002; Trenberth et al. 2002); the El Niño–Southern Oscillation (ENSO) teleconnection (Bromwich et al. 2000; Genthon et al. 2003); and synoptic-scale cyclone activity (Simmonds 2000), to name only a few. The accessibility and spatial coverage of NCEP1 make it a prime choice to conduct climate studies across these regions of large observational data voids.

However, the shortage of observations still negatively affects the skill and reliability of NCEP1 in the high southern latitudes. Hines et al. (2000) observed artificial trends in the mean sea level pressure (MSLP) fields near Antarctica, due to strong positive biases that decrease with time. They identify this linear trend not only in the surface pressure, but also aloft in the 500-hPa geopotential height fields. Hines et al. (2000) demonstrated that few Antarctic surface observations were assimilated into NCEP1 prior to the availability of the Global Telecommunications System (GTS) data in 1967, despite the fact that many Antarctic stations began collecting data around the IGY. The updated NCEP–Department of Energy (DOE) Atmospheric Model Intercomparison Project-2 (AMIP-2) reanalysis (NCEP2; Kanamitsu et al. 2002), covering 1979–present (versus 1948–present of NCEP1), is very similar to NCEP1 in the period of overlap.

Marshall and Harangozo (2000) also noted the large linear trends in the MSLP fields in the Southern Ocean and West Antarctica. Their study showed that around 70°S the total decrease exceeded 12 hPa across the Southern Ocean, a trend that was statistically significant at the 1% level and was not noticeable in any nearby station observations. This trend remained largely until the 1990s, although it showed some improvement after the assimilation of the satellite sounder data beginning in the late 1970s.

Marshall (2002), in addition to the trends in the geopotential height fields, noted erroneous trends in the stratospheric temperatures in NCEP1. These trends are related to the inability of NCEP1 to capture cooling in the stratosphere associated with the seasonal ozone losses. Marshall found marked improvement in these fields after the assimilation of the satellite sounder data, which helped to further constrain the stratospheric temperatures due to the low observation density. Also noted by Marshall was a rapid drop in the East Antarctic height fields in 1993 that created a significant negative bias between NCEP1 and the observations. This sudden drop was found to be associated with the assimilation of some Australian automatic weather stations (AWS) located over the continent whose specified elevations were erroneously low.

Recently, the European Centre for Medium-Range Weather Forecasts (ECMWF) finished their 40-yr global

Re-Analysis (ERA-40) spanning September 1957–August 2002 (see online at <http://www.ecmwf.int/research/era/>). ERA-40 has the benefit of knowing many of the aforementioned problems encountered in NCEP1, and has taken steps to improve the skill of the reanalysis throughout the entire run. For example, ERA-40 assimilated many Antarctic stations from the start of the run that were not assimilated in NCEP1 until after the GTS data were made available (see online at http://www.ecmwf.int/research/era/Products/Archive/Plan/Archive_plan_2.html#478679 for a detailed list of the observations assimilated in ERA-40). As such, the large trends in the pressure fields (both at the surface and aloft) are expected to have been corrected in the new ERA-40 reanalysis. Further, ERA-40 is a second-generation reanalysis, having been preceded by the 15-yr Re-Analysis (ERA-15:1979–93) completed earlier by ECMWF (see Gibson et al. 1997, and references therein).

This paper compares the temporal skill in the new ERA-40 system compared to the widely used NCEP1 system. Section 2 describes the data and methods employed. Section 3 details the annual evolution of the overall (1958–2001) skill, while section 4 examines the change of skill throughout time. Section 5 extends the comparisons beyond Antarctica and the Drake Passage. Observation counts in NCEP1 are presented in section 6 to compare the temporal changes in skill with the trend in the quantity of observations assimilated into NCEP1. Section 7 provides a discussion, and conclusions are drawn in section 8.

2. Data and methodology

To validate each reanalysis, monthly mean observations from the stations in Table 1 (Fig. 1 displays the locations) were compared with the reanalysis values. Observations of MSLP were compared for all stations, whereas surface temperature was compared only for Antarctic stations (1–9 in Table 1), and 500-hPa geopotential height compared for three Antarctic stations (Casey, Halley, and McMurdo). The stations in Table 1 were chosen based on data completeness, yet some comparisons are unavoidably affected by poor data quality and are identified when needed. The station observations (both surface and upper air) for Antarctica were obtained from the British Antarctic Survey READER project Web site (<http://www.antarctica.ac.uk/met/READER/>). The MSLP data for the remaining stations were obtained until 1998 through the NCAR ds570.0 dataset, with the recent years completed from data available through the National Climatic Data Center (NCDC; see online at <http://www.ncdc.noaa.gov/oa/ncdc.html>).

Both NCEP1 and ERA-40 are available on a 2.5° by 2.5° grid every 6 h, although both are run at higher resolutions (T-159/125 km for ERA-40 and T-62/209 km for NCEP1) and downgraded to a 2.5° resolution. ERA-40 contains 60 vertical levels (23 standard pres-

TABLE 1. Coordinates of all stations used in the study. Number corresponds to the number above the plotted stations in Fig. 1. Horizontal lines separate groups of stations as outlined in the text: 1–10 are Antarctic and Drake Passage stations, 11–17 island stations in the Southern Ocean, and 18–25 stations on or near major continents.

| Station name | Lat | Lon | No. |
|------------------|--------|---------|-----|
| Amundsen–Scott | 90.0°S | 0.0° | 1 |
| Casey | 66.3°S | 110.5°E | 2 |
| Dumont D’Urville | 66.7°S | 140.0°E | 3 |
| Faraday | 65.4°S | 64.4°W | 4 |
| Halley | 75.5°S | 26.4°W | 5 |
| Mawson | 67.6°S | 62.9°E | 6 |
| McMurdo | 77.9°S | 166.7°E | 7 |
| Mirny | 66.5°S | 93.0°E | 8 |
| Orcadas | 60.7°S | 44.7°W | 9 |
| Punta Arenas | 53.0°S | 70.8°W | 10 |
| Easter | 27.2°S | 109.4°W | 11 |
| Gough | 40.4°S | 9.9°W | 12 |
| Grytviken | 54.3°S | 36.5°W | 13 |
| Kerguelen | 49.3°S | 70.2°E | 14 |
| Marion | 46.9°S | 37.9°E | 15 |
| New Amsterdam | 37.8°S | 77.5°E | 16 |
| Stanley | 51.7°S | 57.9°W | 17 |
| Buenos Aires | 34.6°S | 58.5°W | 18 |
| Campbell Island | 52.6°S | 169.2°E | 19 |
| Cape Town | 34.0°S | 18.6°E | 20 |
| Chatham Island | 44.0°S | 176.6°W | 21 |
| Christchurch | 43.5°S | 172.5°E | 22 |
| Hobart | 42.8°S | 147.5°E | 23 |
| Macquarie Island | 54.5°S | 159.0°E | 24 |
| Perth | 31.9°S | 116.0°E | 25 |

sure levels) compared to the 28 vertical levels (17 standard pressure levels) of NCEP1. Monthly means were calculated from the reanalysis data and were bilinearly interpolated to the observational station location to within a tenth of a degree as a means of validating the performance of the two reanalyses. The use of monthly mean data allows for a basis of comparison between this study and the appraisal studies identified earlier. Very similar findings are produced using the 6-hourly data although the statistics are more influenced by the observational data completeness at this shorter time scale.

Statistics from the monthly averaged reanalysis and surface data for the MSLP field were examined to compare the evolution of overall (1958–2001) skill in ERA-40 and NCEP1. In each case, the correlation coefficient, bias, and root-mean-square error (rmse) are calculated from the station observations and the reanalysis values. Here bias refers to the mean reanalysis value over the given period minus the mean observed value. Rmse is defined as the square root of the mean-squared difference between the extracted reanalysis values and the observations, and effectively combines the errors of low correlation and high bias into one statistic.

3. Overall (1958–2001) comparison between ERA-40 and NCEP1 MSLP

Annual cycles of these statistics are presented in Fig. 2 for the high-latitude stations (mostly Antarctic) listed

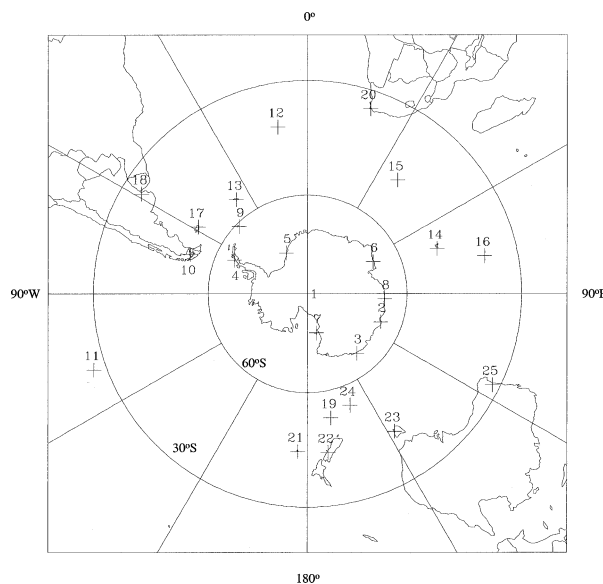


FIG. 1. Locations of the stations listed in Table 1.

in Table 1 (stations 2–10), with the exception of Amundsen–Scott station, which lies well above sea level and therefore does not record MSLP. The time series for the correlations show a marked decline in ERA-40 during the austral winter (Fig. 2a). At nearly every station, the lowest correlations are observed roughly from June to August, although there are a few periods of low correlation at other times. The correlations are lower than 0.6 at four of the eight stations, namely, Casey, Mirny, Mawson, and Punta Arenas. With 13 degrees of freedom, correlations exceeding ± 0.45 , ± 0.52 , and ± 0.65 are significant at the 90%, 95% and 99% levels, respectively. One might assume that the low correlation values are associated with problems in the handling of seasonal sea ice since the correlations show the lowest skill when sea ice is extensive. Deficiencies due to the handling of sea ice can arise because early sea ice data are based on the model climatology; there are few reliable sea ice observations before the modern satellite era. However, the low correlation problem extends to Punta Arenas, north of the maximum sea ice extent, suggesting that sea ice alone is not the main factor.

The correlations in NCEP1 (Fig. 2b) remain fairly high (above 0.8) throughout the year and do not show as strong a seasonal dependence for all stations excluding the East Antarctic stations. At these locations (Dumont D’Urville, Mawson, Casey, Mirny), correlations are at some points lower than those observed in ERA-40. This is particularly true in September, when the correlations at all four East Antarctic stations are at their respective lowest points. The findings are in agreement with Hines et al. (2000) who found large biases in the East Antarctic stations, largely due to the fact that the Antarctic station surface data were not assimilated into NCEP1 until the GTS data were made available in 1967.

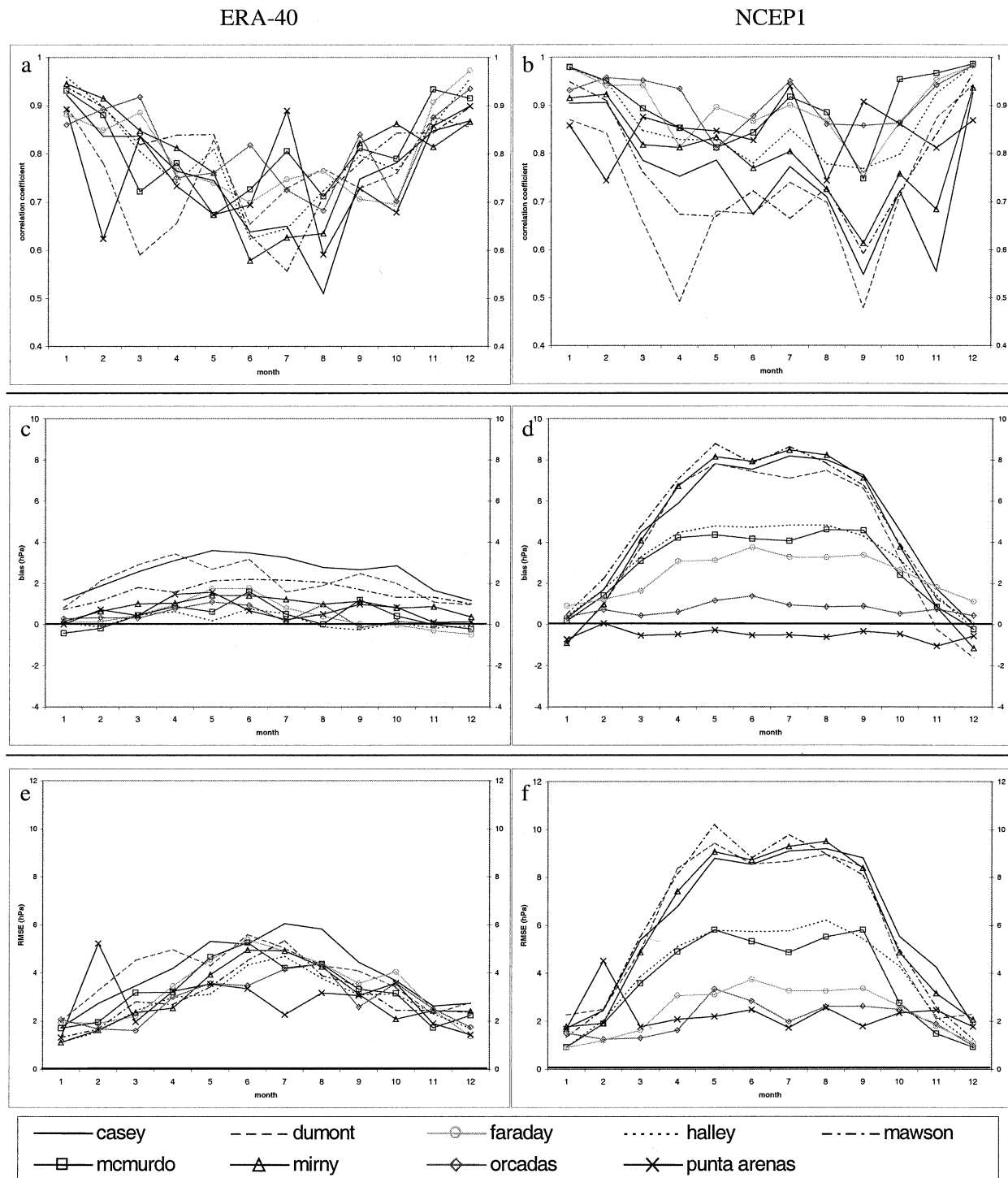


FIG. 2. Annual cycle of comparison between observed and reanalysis MSLP values (1958–2001) for (a), (c), (e) ERA-40 and (b), (d), (f) NCEP1 for stations 2–10 of Table 1. (a), (b) Correlation; (c), (d) bias; and (e), (f) rmse.

Notably, NCEP1 performs better at the other stations throughout the year, particularly in June–July–August (JJA) when ERA-40 has the most distinct problems in capturing the variability.

A slightly different outlook is presented in ERA-40

when looking at the long-term biases (Fig. 2c). Although the biases tend to reach their peak around JJA, consistent with the period of low correlation, they are small: on average less than 2 hPa too high. Relatively large deviations occur at the problem stations noted earlier, al-

though ERA-40 reasonably captures the magnitude of the MSLP at Punta Arenas, evident from the comparably low bias at this station.

NCEP1 produces a drastically different picture (Fig. 2d). Here, at the four East Antarctic stations, the long-term biases on average are around 8 hPa too high in JJA. These four stations indicate a strong seasonal cycle in skill, as the biases are near zero during the austral summer, in agreement with previous studies (Marshall and Harangozo 2000; Hines et al. 2000). There is still a seasonal cycle of the biases at McMurdo and Halley, however the magnitude of the error at these two stations is roughly half that compared to the East Antarctic stations (~ 4 hPa). For the stations in the vicinity of the Drake Passage, NCEP1 has consistently low biases near zero, and is actually performing with a higher degree of skill in this location than ERA-40.

The rmse highlights the apparent problems in both reanalyses (Figs. 2e and 2f), showing that the season of lowest skill is in the austral winter. Overall, the long-term rmse in ERA-40 is roughly half of that in NCEP1 during JJA. During DJF, the two reanalyses are comparable and closely follow the observations. Notably, the equinoctial seasons show a much stronger degradation in NCEP1 brought about by the very large long-term biases across Antarctica that peak in austral winter.

It is noteworthy that in the stations near the Drake Passage region (Orcadas and Punta Arenas) ERA-40 is generally outperformed by NCEP1. In this region, NCEP1 overall has a higher or comparable correlation, lower bias, and lower rmse. This result is somewhat surprising due to the fact that the Drake Passage region is the area with the greatest density of station observations poleward of 50°S . This would perhaps indicate that a relatively large density of station observations does not constrain ERA-40 to the extent that it controls NCEP1; this topic will be addressed further in section 7.

Altogether, Fig. 2 would favor ERA-40 over NCEP1. Although the correlations are lower during the winter in ERA-40, they are quite comparable throughout the rest of the year, and much better in September. The bias and rmse are generally lower in ERA-40. However, as these statistics are calculated using the entire years of complete overlap, 1958–2001, they can fail to capture the changes in skill with time. Particularly, nearly 50% of the 1958–2001 time interval lies in the modern satellite era (post 1978), and as such the statistics presented in Fig. 2 are weighted by the performance during this time period.

4. Time evolution

Here the statistics are presented using 5-yr windows. In this method, each parameter is calculated as before over a 5-yr span, and then edged forward a year and recalculated. This allows one to directly observe the temporal evolution in the skill as well as the impacts of assimilating satellite data. Because the austral winter

was shown to be the most problematic season in terms of skill, these time series are constructed for only the JJA data, thus giving 15 individual months within a 5-yr window. The effect of assimilating satellite data is visualized by dividing the comparisons into three distinct eras of data assimilation noted in ERA-40. The first era spans 1958–72 and represents the 15 yr before any satellite data were assimilated into ERA-40. The second era, 1973–78, represents the years when the Australian surface bogus pressure data (PAOBS) were assimilated into ERA-40 from gridded Australian surface pressure analyses (A. Simmons 2004, personal communication). More importantly, during this period satellite sounder data were first assimilated into ERA-40. The Vertical Temperature Profile Radiometer (VTPR) sounding data were assimilated into ERA-40 starting on 1 January 1973, while the Television Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS) data entered in late 1978. The final era, 1979–2001 represents the complete years of overlap between ERA-40 and NCEP1 when a vast array of satellite and conventional data including drifting buoys and commercial aircraft observations were available to both reanalyses, and is hereafter identified as the modern satellite era. We also observe the skill for other variables apart from MSLP.

a. MSLP

The 5-yr running statistics for MSLP for both ERA-40 and NCEP1 in JJA are presented in Fig. 3. The correlations, biases, and rmse are plotted side by side for ERA-40 and NCEP1 to facilitate comparison of the two reanalyses.

From Figs. 3a,b one can clearly see the problem with ERA-40's ability to capture the monthly variability during JJA. In fact, the values presented in Fig. 2 obscure a period of weak-positive to weak-negative correlations during the mid-1960s. Correlations during this time period reach a minimum at about -0.2 , showing a slight anticorrelation with observed values. The values increase rapidly and in 1973, after the VTPR data are first assimilated (indicated by the vertical line), the range of correlations drops significantly and is comparable to the values obtained by NCEP1. After the TOVS data begin to be assimilated into the 5-yr windows starting in 1979, the range of correlation values again decreases; the correlations are beginning to converge to near 1.0. Notably, throughout the modern satellite era, the correlations are all near perfect (1.0). NCEP1, on the other hand, does not show nearly such large temporal changes in skill. As expected from Fig. 2, during the modern satellite era, ERA-40 is superior to NCEP1, although both have high correlations >0.9 .

Examining the biases in Figs. 3c,d, a nearly opposite picture is found compared to the correlation values. This time, ERA-40 is performing with a greater degree of accuracy; the biases in ERA-40 are roughly half of those

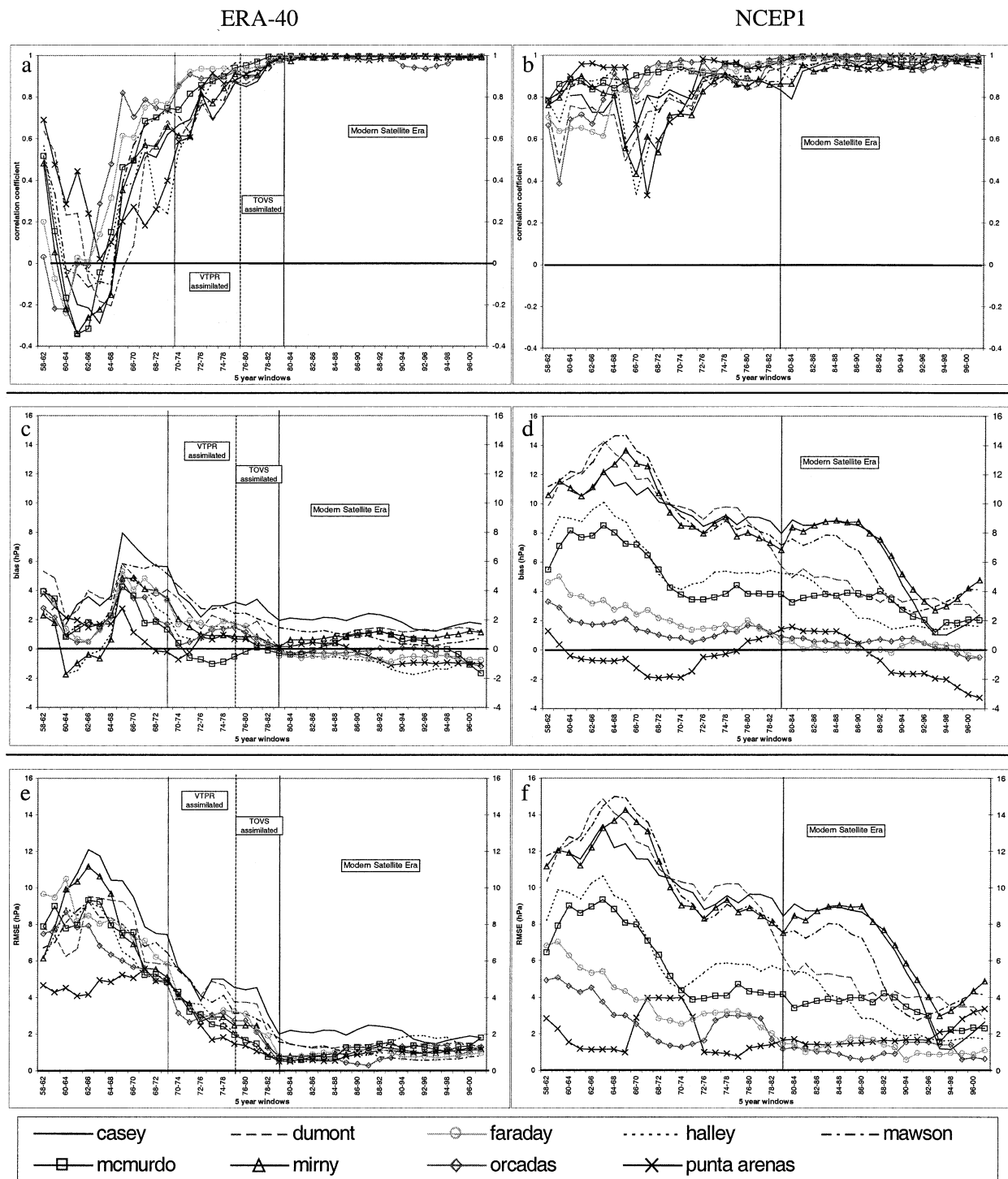


FIG. 3. JJA 5-yr running window comparison between observed and reanalysis MSLP values (1958–2001) for (a), (c), (e) ERA-40 and (b), (d), (f) NCEP1 for stations 2–10 of Table 1. (a), (b) Correlation; (c), (d) bias; and (e), (f) rmse.

in NCEP1, except for the Drake Passage stations addressed earlier. The large linear trend addressed by Marshall and Harangozo (2000) and Hines et al. (2000) is readily obvious in Fig. 3d, with improvements continuing until the 1990s. The improvement noted in the mid-

1990s is a direct result of the inclusion of the Australian AWS data (Marshall 2002), which provided observations over most of the East Antarctic interior, although it created a sudden drop in the geopotential heights. Yet even at this stage the biases in NCEP1 are still more

than 2 times those found in ERA-40. The differences in the assimilation schemes between ERA-40 and NCEP1 are evident; ERA-40 is strongly guided by satellite observations whereas NCEP1 shows considerable constraint by a relatively large density of station observations due to the maintained skill in the Drake Passage region. ERA-40 MSLP biases are sporadic, although spatially consistent, before the assimilation of the VTPR data; after this period the range begins to converge, falling between ± 2 hPa in the modern satellite area.

The rmse plots in Figs. 3e,f show the same general picture, with ERA-40 being spatially consistent and covering a much smaller range of rmse values. NCEP1, largely affected by the high biases in excess of 12 hPa during some periods, has a much greater range of rmse. The stations with the highest rmse are the East Antarctic stations mentioned earlier, with substantial improvements occurring in conjunction with the improvements in the bias during the mid-1990s.

Clearly the problems in the presatellite years limit the reliability of both ERA-40 and NCEP1. With correlations low despite monthly averaging, ERA-40's MSLP fields before roughly 1972 are of limited value at high southern latitudes, especially when MSLP variability is an issue, as in cyclone tracking (since these problems also appear in the 6-hourly data). However, the large MSLP biases throughout much of NCEP1 create erroneous linear trends that make its use highly questionable as well, especially for studies that demand precision, such as those involving the Antarctic Oscillation.

b. MSLP time series

Time series of the observed versus ERA-40 JJA MSLP are presented in Fig. 4 for the three stations demonstrating the lowest correlations at their minimums, namely, McMurdo (Fig. 4a), Mirny (Fig. 4b), and Casey (Fig. 4c). Looking at the two series plotted simultaneously reveals the problems of ERA-40 in the presatellite era. The moderate biases at all stations during the 1960s are readily distinguished in Fig. 4, with an isolated event where ERA-40 is over 20 hPa too high during June 1959 at McMurdo. The large bias noted by Marshall (2003) during July 1964 is seen in all three plots, especially at Mirny, where the magnitude of the bias is greater than 20 hPa (too low). The biases, although high in isolated situations, are not the main limitation. During this period, correlations between observed and ERA-40 MSLP are negative. This suggests that the observations are not constraining ERA-40. Rather, ERA-40 relies on a better model climatology that produces overall lower biases (Fig. 3c). This is quite different than NCEP1 (Fig. 3d), whose inadequate model climatology creates high biases before the surface observations were assimilated via the GTS. Additionally, NCEP1 appears more constrained by the obser-

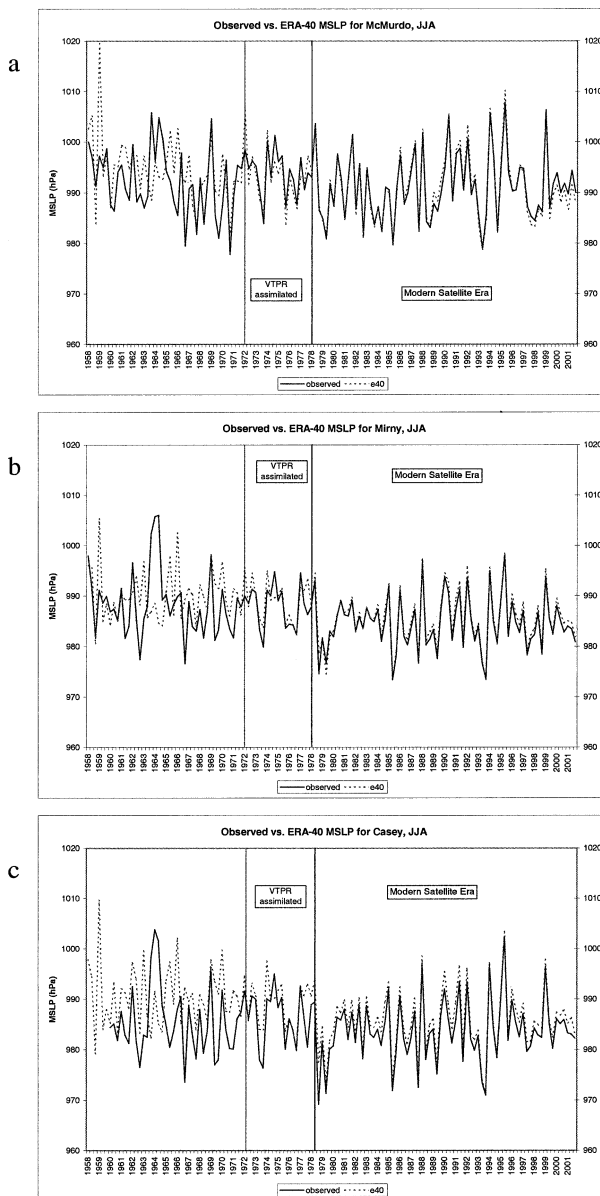


FIG. 4. JJA observed vs ERA-40 MSLP for (a) McMurdo, (b) Mirny, and (c) Casey.

vations, yielding the higher correlations seen in the presatellite era.

Apart from the problems before the mid-1970s, the other main feature observed in Fig. 4 is the improvement of skill over the last two decades. For each station, the observed and the ERA-40 values converge and nearly trace each other beginning around 1979. This corresponds to the year when TOVS data were assimilated into the reanalysis, after which an abundance of various satellite and other conventional data became available. This high level of skill is unique to ERA-40, indicated by the high correlations, low biases, and low rmse displayed in Fig. 3.

c. 2-m temperature

It is important to extend the comparisons to other variables to verify whether these problems exist throughout the whole reanalysis or if they are simply confined to one variable. Here, we observe the 5-yr running window statistics for the 2-m temperatures in the same fashion as for MSLP, with the exception that Amundsen–Scott Station is included in the analysis instead of Punta Arenas. These results are presented in Fig. 5 and in general exhibit very similar characteristics to the MSLP fields.

The ERA-40 correlations (Fig. 5a) contain the same drop-off around the mid-1960s to negative values as in the MSLP field. An exception is at McMurdo, whose correlation decreases in the early 1970s right before the VTPR data were assimilated. NCEP1 has lower and more sporadic correlations with the observed temperatures than it does with the MSLP (Fig. 5b). Although these correlations improve with time, they do not converge as close to 1 or carry the same spatial continuity seen in the ERA-40 correlations during the modern satellite era. Clearly the effects of adding satellite data to ERA-40 are readily observed again for the 2-m temperature correlations, as the skill changes significantly with time.

There is a marked negative bias for both reanalyses at most stations (Figs. 5c,d). However, these are not true systematic biases as one can see throughout the three time periods that they do not greatly improve, even in ERA-40. Rather, the biases noted in Fig. 5 are most likely a result of the sharp changes in the terrain that occur near the Antarctic coastal stations. Due to a relatively coarse horizontal resolution, both ERA-40 and NCEP1 greatly smooth the terrain at the coastal stations. Thus, the station locations in ERA-40 are much more likely to be at a higher height than they are in reality, a problem observed by other authors, even for higher-resolution models (e.g., Bromwich et al. 2005). Table 2 demonstrates the actual station heights and the model station heights for ERA-40 and NCEP1. Here the ERA-40 station heights were extracted from the higher-resolution model output (regular Gaussian grid) since this is the only format the heights were archived. The higher-resolution of ERA-40 accounts for the model heights being closer to the actual heights compared to NCEP1. By assuming a dry adiabatic lapse rate, we can project how the surface temperatures are affected by the differing elevations between the observed and reanalysis station locations; the average model minus observed temperature biases for the modern satellite era, 1979–2001, are also listed in Table 2 for comparison. At the coastal stations, the greater elevation of these stations in the reanalyses accounts for a large portion of the cold biases observed here for both ERA-40 and NCEP1. Where the orographic gradient is gentle, such as at Amundsen–Scott and Orcadas (on the South Orkney Islands), the actual station height and the model station

heights are in much closer agreement and a smaller magnitude in the bias is seen in Figs. 5c,d. Thus the strong biases shown here are not as large as the statistics would indicate, but are exaggerations due to the reanalyses' smoothing of the sharp changes in the terrain at nearly every Antarctic coastal station.

Over the interior, there is an apparent warm bias, which is seen at Amundsen–Scott in the modern satellite era, but can also be observed at Vostok (not shown due to gaps in the observational record), some of which is due to the reanalyses underestimating the actual station height (Table 2). Although smaller than the MSLP fields, there is still improvement in the ERA-40 biases with time, starting in the period when VTPR, and especially TOVS data, were assimilated. The improvement is a decrease of about 2–3 K (up to ~5 K) in the magnitude of the bias during these 10 yr. In agreement with other plots (e.g., Fig. 3), there is little change during the modern satellite era, except for a warm bias that continues to increase at Halley. The bias in NCEP1, however, does not improve with time as much as in ERA-40. Improvements in NCEP1 are on the order of 3 K at Mawson. The lack of improvement in NCEP1 may be due to the fact that it uses a different terrain than in ERA-40, although this is likely to only be part of the explanation for the difference. The impact of the relatively poor model climatology in NCEP1 is likely to also reduce the improvement.

The rmse reflects these changes, with decreases occurring during the assimilation years of satellite data in ERA-40, and rather gradually in NCEP1 (Figs. 5e,f). Even though the improvements are not as clear as they were for the MSLP statistics, the fact that ERA-40 is largely guided and improved by the satellite data still is apparent from the 2-m temperature statistics. As such, precautionary measures should be taken when using these data from ERA-40 and NCEP1 prior to 1970.

d. 500-hPa geopotential height

Clearly the surface circulation has significant problems in both ERA-40 (low correlations in presatellite era) and NCEP1 (large biases that have a strong trend in time). We look next to see if these problems persist above the surface. Because of large data gaps that can strongly influence the statistics presented here, only three stations (Amundsen–Scott, Casey, and Halley) were chosen from Table 1 based on their data completeness. We present the statistics using the 5-yr running window method for only the 500-hPa geopotential heights since there are even more gaps in the data records at higher levels; these results are shown in Fig. 6.

Not surprisingly, the problems with the ERA-40 correlations are again seen in the 500-hPa geopotential height fields, although to somewhat lesser of a degree; only at Casey do the correlations become negative (Fig. 6a). The trends in the 500-hPa height biases are consistent with the trends seen in MSLP (Fig. 3c), indicating

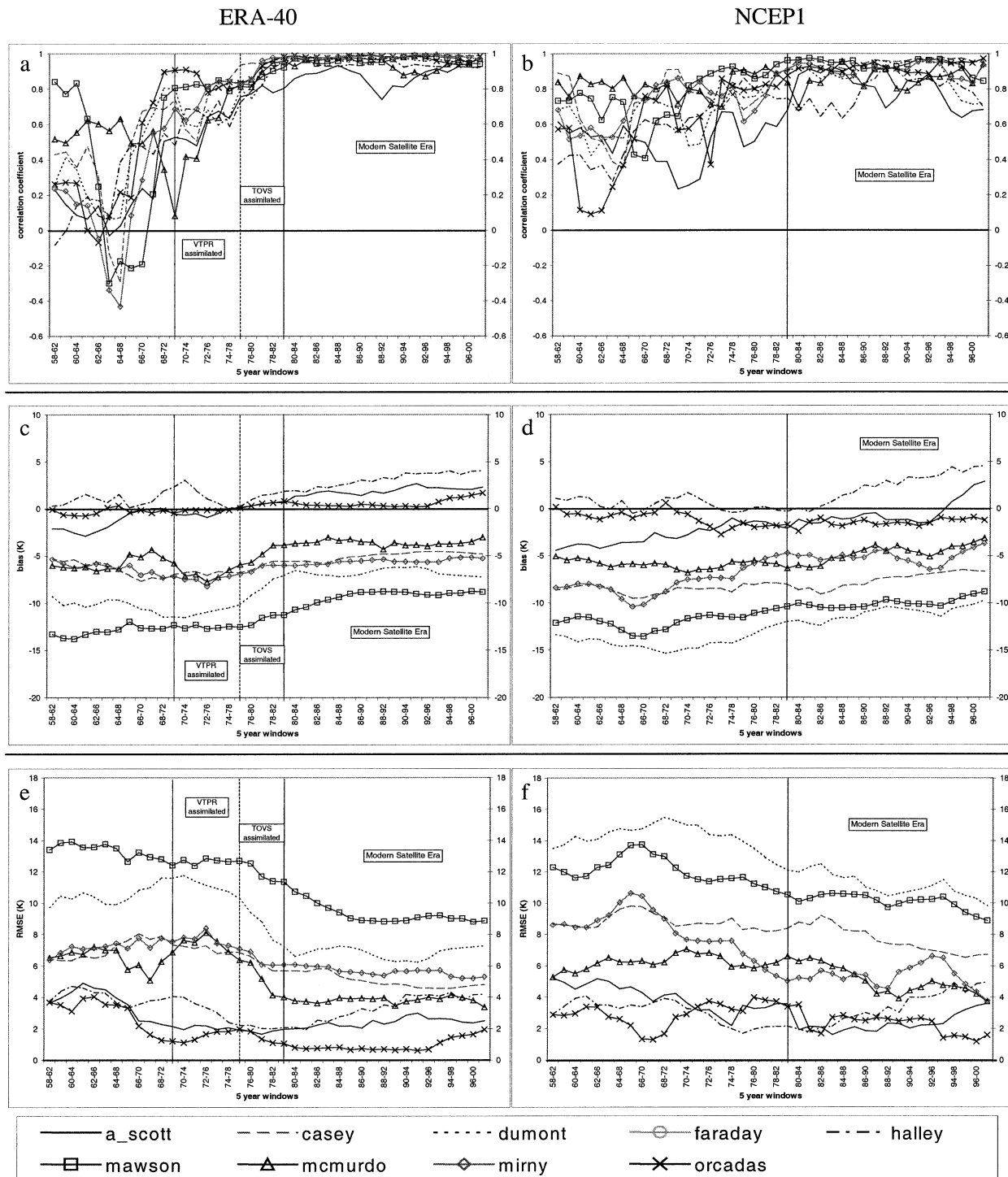


FIG. 5. Same as in Fig. 3, but for 2-m temperature and stations 1–9.

the equivalent barotropic nature of the errors. There is a high level of skill during the modern satellite era, however, and is of equal if not greater value than NCEP1. Hardly any noticeable improvements in the upper-air correlations with NCEP1 are noted throughout the 42-yr comparison here (Fig. 6b). The places where

a problem in a specific series (e.g., Amundsen–Scott during 1979–83) is mimicked in both the ERA-40 and NCEP1 plots (Figs. 6a,b), indicate that it is the gaps in the data that are affecting the statistics, and not deficiencies in the individual reanalyses.

Interestingly, the biases between observations and ex-

TABLE 2. Actual and model station heights for stations 1–9 of Table 1 (in m). Corresponding temperature biases (K) resulting from large actual/model height differences are listed for ERA-40 and NCEP1, compared with the average model-observed temperature biases for ERA-40 and NCEP1 in the modern satellite era (1979–2001). See text for details.

| Station name | Actual station height | ERA-40 | | | | NCEP1 | | | |
|------------------|-----------------------|--------|------------|-------------------------|----------------------------------|--------|------------|-------------------------|---------------------------------|
| | | Height | Difference | Height temperature bias | ERA-40 observed temperature bias | Height | Difference | Height temperature bias | NCEP1-observed temperature bias |
| Amundsen–Scott | 2835 | 2738 | 97 | 1.0 | 1.8 | 2903 | –68 | –0.7 | –0.5 |
| Casey | 42 | 348 | –306 | –3.1 | –5.0 | 577 | –535 | –5.5 | –7.6 |
| Dumont D’Urville | 43 | 511 | –468 | –4.8 | –6.8 | 730 | –687 | –7.0 | –11.1 |
| Faraday | 11 | 275 | –264 | –2.7 | –2.7 | 203 | –192 | –2.0 | –6.2 |
| Halley | 30 | 160 | –130 | –1.3 | 3.1 | 341 | –311 | –3.2 | 2.3 |
| Mawson | 16 | 581 | –565 | –5.8 | –9.3 | 760 | –744 | –7.6 | –10.0 |
| McMurdo | 24 | 249 | –225 | –2.3 | –3.6 | 270 | –246 | –2.5 | –4.8 |
| Mirny | 30 | 441 | –411 | –4.2 | –5.6 | 676 | –646 | –6.6 | –5.1 |
| Orcadas | 6 | 2 | 4 | 0.0 | 0.6 | 21 | –15 | –0.2 | –1.5 |

tracted reanalysis values are highly comparable in terms of magnitude (Figs. 6c,d; both are $\sim \pm 40$ gpm in the 1960s, and ~ -50 gpm in the 1970s). Yet, a noticeable difference is the fact that the biases are quite inconsistent (cycling from positive to negative, back to positive) in ERA-40 prior to the 1970s, after which they become negative (Fig. 6c). The negative height bias in ERA-40 was an issue identified by Bromwich et al. (2002) in their report to ECMWF in their study of preliminary ERA-40 data from 1989 to 1991. However, the biases approach zero quite rapidly during the mid- to late 1970s when satellite data were assimilated into ERA-40, and for most of the modern satellite era are slightly better than NCEP1. The sudden switch to the negative bias at Casey in NCEP1 (Fig. 6d) is again related to the assimilation of the Australian AWS data as seen in Marshall (2002).

Despite the relative ability to compare ERA-40 and NCEP1’s magnitude of the biases, ERA-40 is still quite inferior to NCEP1 in the 1960s as indicated by the rmse plots (Figs. 6e,f). In fact, ERA-40’s rmse is nearly double that of NCEP1, but quickly drops in the mid-1970s, and is similar to NCEP1 during the modern satellite era (Fig. 6e). Nonetheless, an rmse of >120 gpm at Casey and over 80 gpm at both Amundsen–Scott and Halley substantially diminish the quality and usefulness of the 500-hPa geopotential height data prior to the modern satellite era. The inconsistencies in the biases (switching from negative to positive and then back again) also reduce the reliability of ERA-40 prior to 1970 unlike in the MSLP data, which had a consistent positive bias that was many times smaller than in NCEP1.

5. ERA-40 performance in southern midlatitudes

Up to this point, ERA-40 has been shown to have some major shortcomings before the assimilation of satellite data around Antarctica where there are known deficiencies in NCEP1. Due to the strong convergence toward higher skill with the increasing assimilation of satellite data, the projected reasoning for the observed errors is most likely due to the dependence of ERA-40 on satellite observations. For this conjecture to hold true, ERA-40 would need to demonstrate the same patterns in skill across the island stations in the Southern Ocean, where large spatial gaps in data exist as in Antarctica. To verify the claim, MSLP statistics for seven southern midlatitude stations (11–17) from Table 1 are presented in Fig. 7 using the 5-yr running window method as above. Unfortunately, many of the records for these stations are incomplete, and therefore the statistics for these stations are only displayed when over 80% of the observations are included in each 5-yr window. For the most recent decades, this limits the stations to only Easter, Gough, and Marion Islands.

Overall Fig. 7 shows the same structure noted in Fig. 3. The correlations in ERA-40 (Fig. 7a) are very low in the presatellite era and are negative at some locations

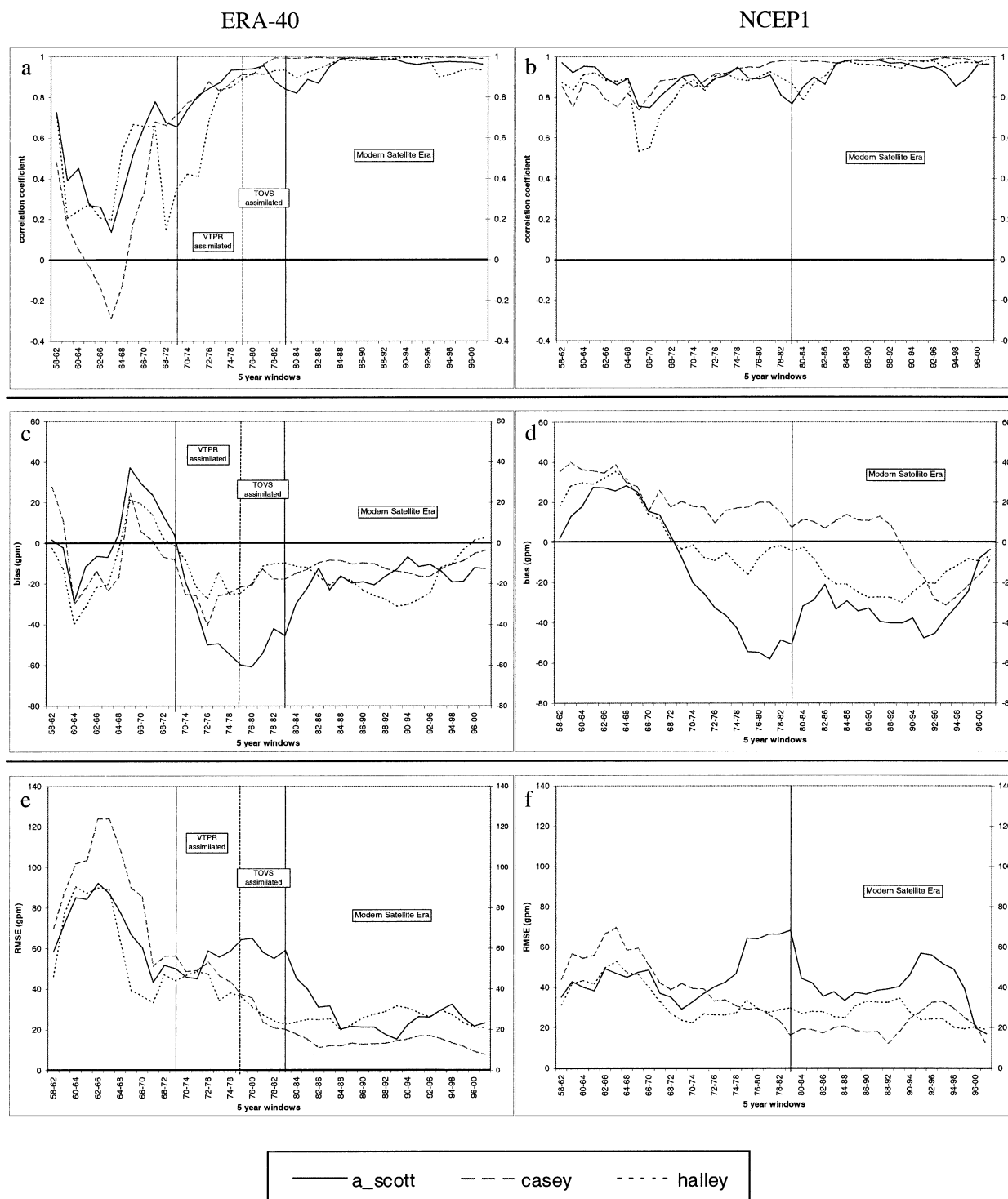


FIG. 6. Same as in Fig. 3, but for 500-hPa geopotential height.

in the mid-1960s; however, NCEP1 also has quite low correlations (Fig. 7b) unlike those noted at the Antarctic stations. A bit surprisingly, the biases in ERA-40 (Fig. 7c) are larger and cover a greater range than in NCEP1 (Fig. 7d), a problem hinted at by Marshall (2003) when

reconstructing the zonal pressure index at 40°S. The larger biases and low correlations lead to higher rmse errors in ERA-40 overall (Figs. 7e,f). Again, the trend in the skill is largely a factor of the assimilation of satellite data in ERA-40, with values converging within

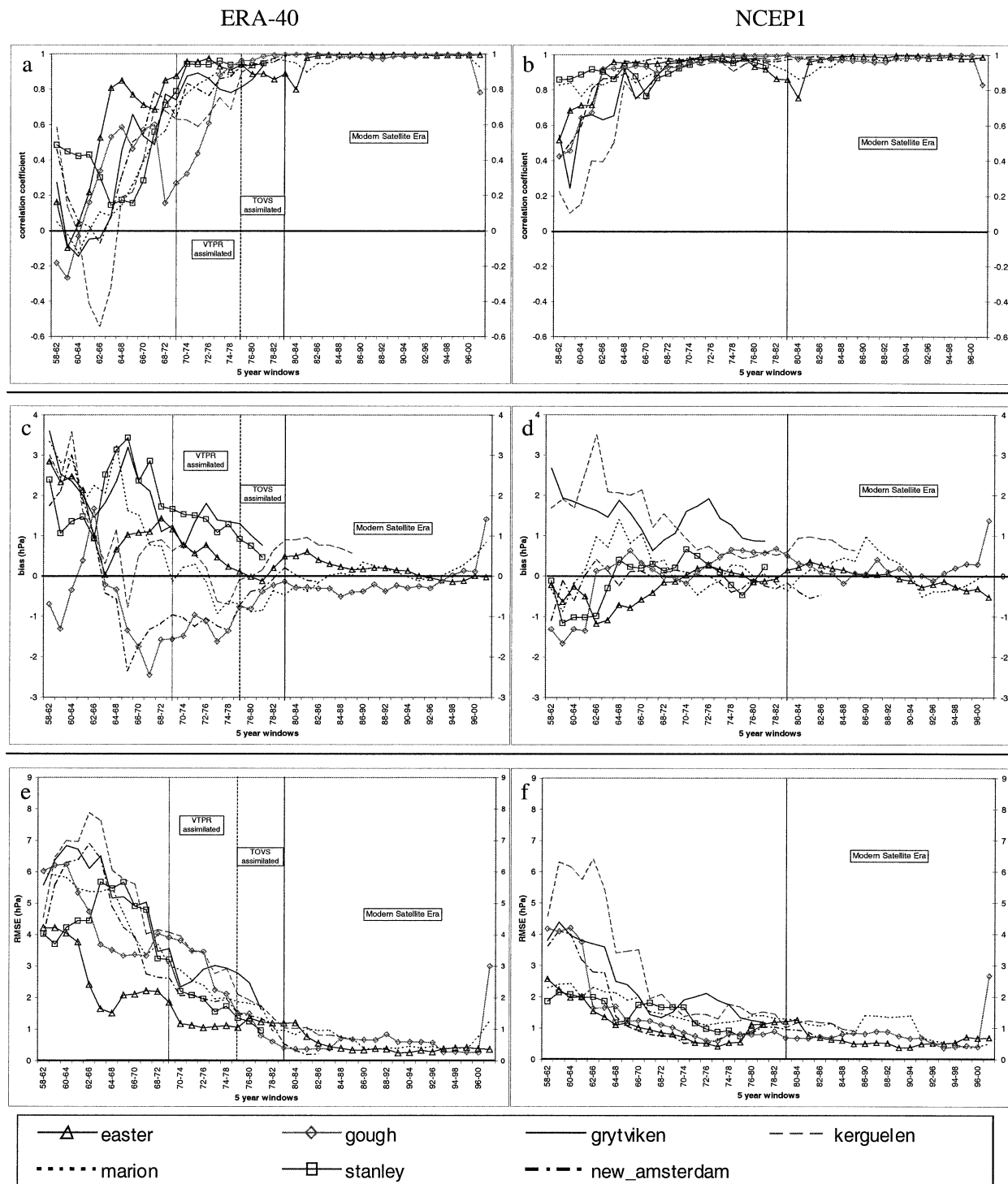


FIG. 7. Same as in Fig. 3, but for stations 11–17 listed in Table 1.

acceptable ranges (within the range of measurement error) almost exactly at 1979. Problems at Easter Island in 1980–84 and at Gough Island in 1997–2001 are apparent in both reanalyses and are therefore likely to be an observational data-quality issue rather than a problem with the reanalyses.

Examination of stations on or near the continental mainland or island stations with frequent ship or air traffic (i.e., Macquarie Island, Campbell Island, Chatham Island are such less isolated stations) across the Southern Hemisphere is warranted to ensure that the problems during the austral winter are not a gross de-

iciency in ERA-40, but are instead related to the quantity of the observational data assimilated. Due the fact that stations over the continental mainlands (Australia, New Zealand, South America) have a greater spatial density of observations than the Drake Passage region, it is expected that there will be significant improvements in the overall skill of ERA-40. Eight continental stations (18–25) were selected based on their location (farthest south, global representation) and data quality/completeness and are listed in Table 1. Statistics (using 5-yr running windows as before) are displayed in Fig. 8.

Clearly there is a large improvement in skill at these selected stations. Although the correlations of ERA-40 versus observations in Fig. 8a suggest a similar problem as in the Antarctic and Southern Ocean stations, Fig. 8a is plotted on a different scale; correlations are consistently above 0.5. Furthermore, of the four stations with the lowest correlations, two are islands south of New Zealand (i.e., Campbell Island and Macquarie Island) and are thus more influenced by the lack of nearby observations as at stations 10–16 from Table 1 (previous section). Additionally, the problems at Macquarie Island in Fig. 8a in the early stages also appear in NCEP1 (Fig. 8b), and thus could be compromised by observational data-quality issues independent of the reanalyses. However, ERA-40 still seems to have problems with the correlation at Cape Town, South Africa, and Buenos Aires, Argentina, which are not readily explained, especially since correlations at Hobart, Tasmania, and Christchurch, New Zealand—not on major continents—are consistently at or above 0.95 (Fig. 8a). The noticeable dip in correlation values in Figs. 8a,b at Cape Town right around 1979–83 is also likely due to the quality of the observational data.

The biases are small throughout and comparable. It can be argued that ERA-40 has a lower bias in the modern satellite era, with the discrepancies observed in both reanalyses in the last 5 yr or so not necessarily reflecting the true skill of the reanalyses. The trend in skill is seen in the rmse plots (Fig. 8e,f); with values in both reanalyses converging to less than 0.5 hPa, well within measurement error. Even in the initial stages of ERA-40, the rmse values are still reasonable (Fig. 8e), suggesting that it is really the lack of data that is negatively affecting ERA-40 before 1973 and not some gross error. However, NCEP1 does appear to have a slight edge in performance before the mid-1970s (Fig. 8f).

During the early years and in data-sparse regions such as Antarctica and the Southern Ocean island stations, the biases are much smaller in ERA-40 than in NCEP1, however, the correlations are generally higher in NCEP1. The small biases and simultaneous low correlations in ERA-40 demonstrate that ERA-40 relies on a more representative model climatology than NCEP1. NCEP1 captures more of the variability, but the model climatology is not representative of the observed conditions, thus creating large biases. The contrasts seen

here between ERA-40 and NCEP1 are thus not only a factor of the differing assimilation schemes, but also a result of the better background fields in ERA-40 compared to NCEP1.

6. Observation counts

The NCEP–NCAR reanalysis Web site (<http://wesley.wwb.noaa.gov/reanalysis.html>) contains diagnostic software that allows a user to display the average number of observations assimilated into the reanalysis per 2.5° by 2.5° grid box, a product that is currently unavailable for ERA-40. Although the reanalysis projects use fixed assimilation schemes, the quantity of observed data changes significantly with time. To show the temporal and seasonal dependence of the number of observations available, plots were made over the coastal Antarctic domain (all longitudes between 60° and 80° S). Figure 9 shows the plots for austral winter (Fig. 9a) and austral summer (Fig. 9b). ERA-40 should be based on approximately the same observations as NCEP1, with the addition of the surface Antarctic stations that were absent in NCEP1 prior to the inclusion of the GTS data in 1967. It is likely that ERA-40 might also have had extra radiosonde data compared to NCEP1; additionally, concerted efforts were made to assimilate more of the early satellite data (VTPR) into ERA-40 compared to NCEP1.

As expected, there is both a strong seasonal and temporal dependence to data availability around Antarctica. During the polar winter, there were essentially no observations except the radiosonde data that were assimilated into NCEP1 before the mid-1960s (Fig. 9a). Here, ship data refer to both ship observations and buoy observations. Thus, the large spike seen in 1979 corresponds to the First Global Atmospheric Research Program (GARP) Global Experiment (FGGE) release of many drifting buoys transmitted on the GTS across the high southern latitudes. Except for this peak, the JJA ship data occur much less than radiosonde observations until the most recent decades. In the polar summer, ship observations accounted for over 80% of the total observations assimilated into NCEP1 prior to the inclusion of the surface data, after which ship observations still accounted for roughly 20%–30% of the total assimilated data (Fig. 9b). The large portion of ship data is most likely due to the continued presence of Antarctic buoy data through the World Climate Research Programme (WCRP) International Programme for Antarctic Buoys (IPAB) (WCRP IPAB 2002). Overall, it is evident that the majority of assimilated data in this region comes from the surface station data, especially in winter when surface observations account for $\sim 80\%$ of the total assimilated data (Fig. 9a).

The plots in Figs. 9a,b were presented without the inclusion of the satellite data. In ERA-40, satellite-derived temperatures from the VTPR were first assimilated in January 1973, with the inclusion of TOVS in 1978;

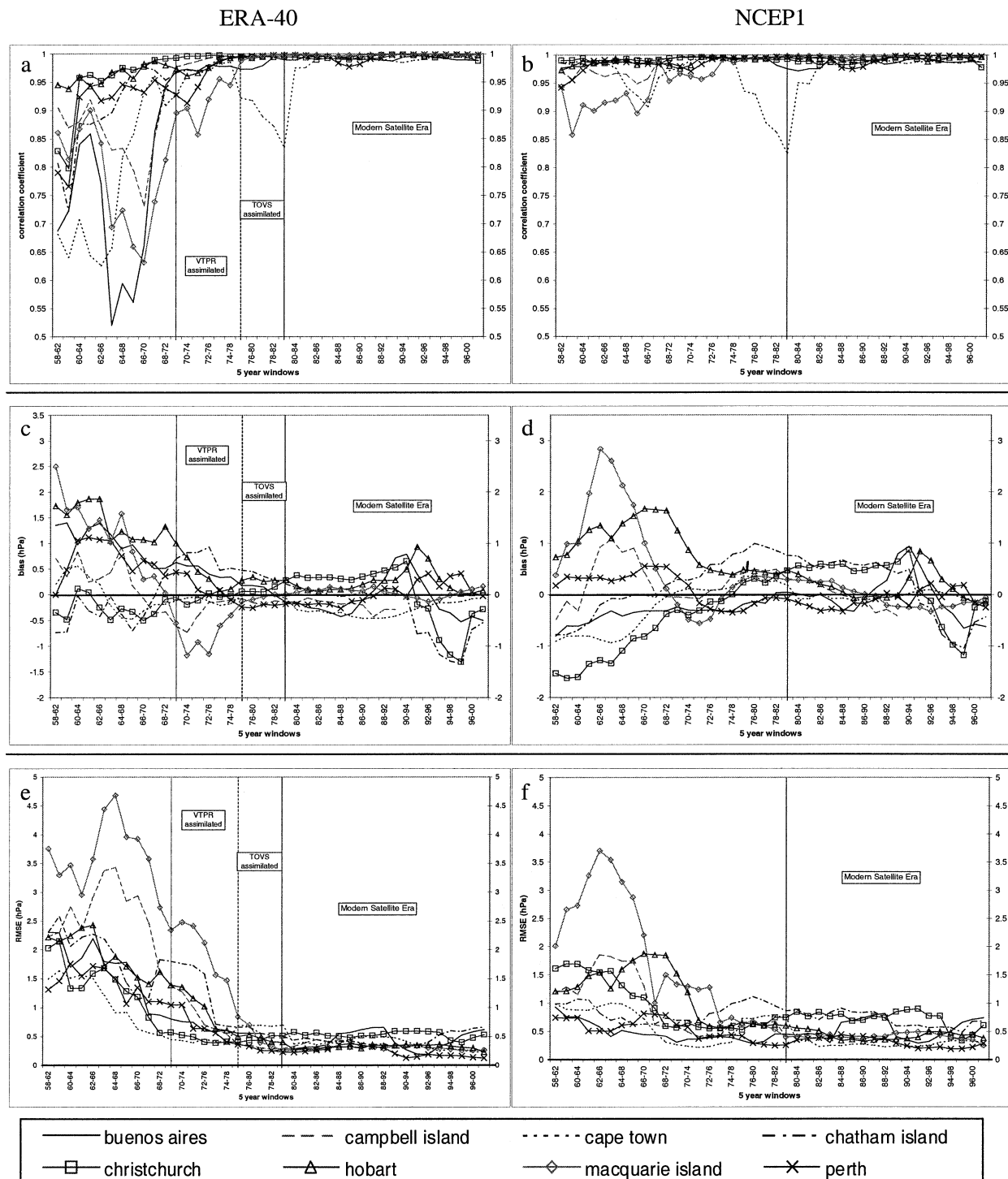


FIG. 8. Same as in Fig. 3, but for stations 18–25 listed in Table 1.

some satellite data entered NCEP1 prior to TOVS. Figure 9c displays the sum of the total observations (including the inferred surface pressure data from satellite images, known as PAOBS) assimilated into the NCEP1 for both JJA and December–January–February (DJF)

plotted simultaneously for the sums with and without the satellite data. In the early 1970s, the total observation curve breaks away from each season, indicating that at least some of the VTPR data were assimilated into NCEP1. In the late 1970s the two significantly di-

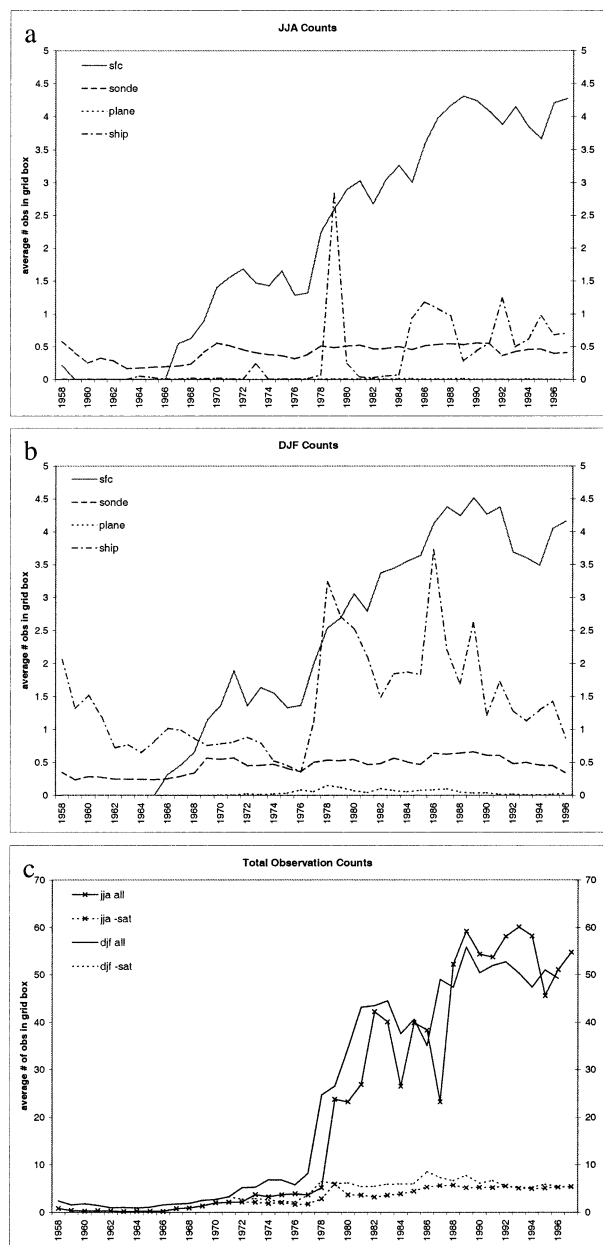


FIG. 9. Total observation counts in NCEP1 per 2.5° by 2.5° grid boxes over coastal (60° – 80° S) Antarctica. Counts for (a) JJA and (b) DJF. (c) Counts for JJA and DJF with (solid line; all) and without (dashed line; -sat) the inclusion of the satellite data.

verge, showing that in the modern satellite era over 85% of the assimilated data is information from the satellites. Such a dramatic increase in observation counts should be reflected into the reanalysis system, as many spatial gaps previously unmeasured are now filled through satellite data. As shown in Figs. 3 and 5–8, NCEP1 does not show a shift in improvement associated with the abundance of satellite data. However, ERA-40 performance reflects this change in observation density, and is well-guided and constrained by the assimilation of

the satellite data during the last two decades, to the point at which it maintains a higher level of skill than NCEP1. The large distinctions in data quantity between winter and summer disappear in the modern satellite era (Fig. 9c), as the satellite inclusive series are similar in magnitude; isolated periods even exist when there are more observations in JJA than DJF. As expected, both ERA-40 and NCEP1 perform comparably well in both winter and summer over the last few decades.

Although the plots in Fig. 9 show counts produced over Antarctica for NCEP1, ERA-40 does produce global spatial plots and radiosonde time series plots for select geographical locations on the ERA-40 project Web site (<http://www.ecmwf.int/research/era>) under the section on monitoring. A project report series by Simmons et al. (2004) clearly shows the strong increase in observations through the modern satellite era as seen here for NCEP1. Their study also shows a decrease in observation counts in the Southern Hemisphere from 1958 to 1966. This may help to explain the decrease seen here in ERA-40 correlations during these times. Similar to this study, Simmons et al. (2004) find that the skill of the 2-m temperature analyses in ERA-40 in the Southern Hemisphere drastically increases after 1978. This increase is related not only to the inclusion of satellite data, but also to new surface observations from drifting buoys and increased data from commercial aircraft.

7. Discussion

It is important to note that the comparisons are only made at single points in the Southern Hemisphere where observational datasets are available; comparisons for the Northern Hemisphere in data-sparse areas are beyond the scope of the current study. Yet, this study still neglects the large differences occurring between NCEP1 and ERA-40 over the data-sparse Southern Ocean where large data gaps also exist. For example, average differences for all months between the 500-hPa geopotential height fields in NCEP1 and ERA-40 in the South Pacific region prior to the 1980s can be as large as ~ 50 gpm (Fig. 10). However, the differences fluctuate quite drastically. By averaging the ERA-40 minus NCEP1 500-hPa geopotential height difference in a box in the South Pacific for all months (60° – 70° S, 130° – 150° W, i.e., the center of greatest difference in Fig. 10), a time series of bias between ERA-40 and NCEP1 is produced (Fig. 11). There are many events when the 500-hPa level in ERA-40 is over 100 gpm lower than in NCEP1 and in July 1959 the difference is >200 gpm. Without data to verify either situation projected by the reanalyses, there is no objective way to discern whether ERA-40 or NCEP1 is producing the more accurate representation. These problems persist throughout the depth of the troposphere, indicated by similar but even larger differences in the geopotential height field at 300 hPa (not shown).

Clearly, the large deficiencies noted here at the sta-

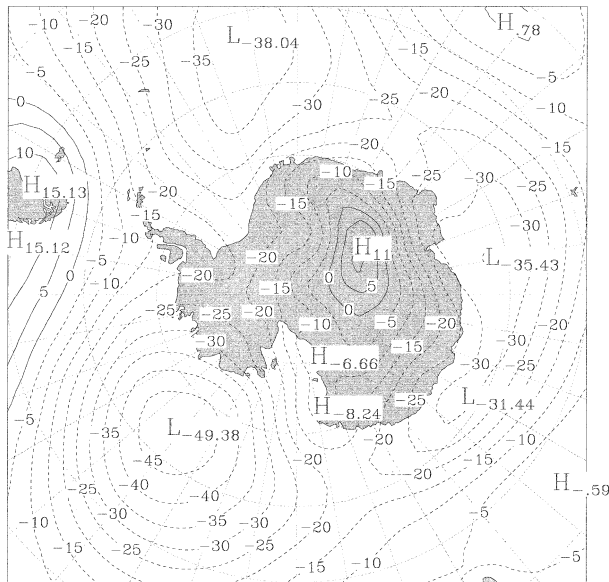


FIG. 10. Average ERA-40 minus NCEP1 500-hPa geopotential height difference for all months (1958–78) in gpm.

tions before the satellite era reflect even larger deficiencies in the data-void South Pacific and other oceans. This problem was alluded to in Bromwich et al. (2000), who demonstrated that over the island stations in the southern midlatitudes the differences between the ERA-15 reanalysis and the ECMWF operational analyses are generally less than half of that over the observation-sparse Southern Ocean regions, such as the South Pacific. Cullather et al. (1996), using sea level pressure observations from independent ship data in the Southern Ocean between 120°W and 180° from 1980 to 1994, demonstrated that the ECMWF operational analyses closely follow the observations in this data-sparse region. More recently, King (2003) demonstrated that the ECMWF operational analyses show very close agreement with independent surface pressure observations in the Bellingshausen Sea region for February–May 2001. Both Cullather et al. (1996) and King (2003) indicate again that the ECMWF operational analyses (that are very similar to ERA-40) are well constrained by the satellite data over otherwise data-sparse locations. However, their comparisons say nothing of the quality of these fields prior to the modern satellite era, which from Fig. 10 are significantly different than NCEP1.

8. Conclusions

The results here demonstrate significant shortcomings in both ERA-40 and NCEP1 before ~1970. In ERA-40, very low and even negative correlations severely limit the reliability of ERA-40's surface and upper-air fields, although there is a marked improvement in the bias and rmse of these fields with time. NCEP1, however, has large biases in the MSLP fields and artificial

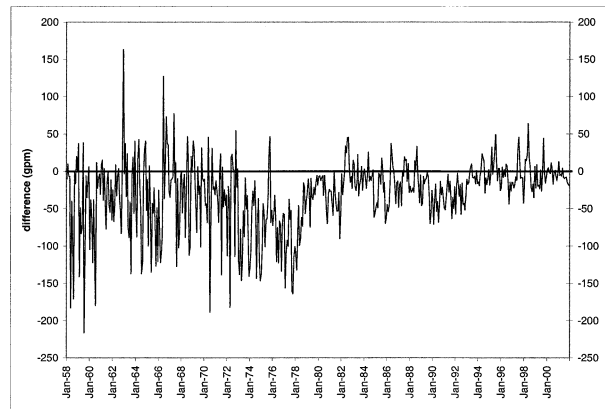


FIG. 11. ERA-40 minus NCEP1 differences for 500-hPa geopotential height averaged in a South Pacific Ocean box for all months (60°–70°S, 130°–150°W).

trends in the high-latitude time series. It is shown that these problems are the largest during JJA, coincident with the small quantity of assimilated observations into both reanalyses. The problems noted here extend into the midtroposphere above the Antarctic continent, and thus are not solely a result of the reanalyses' ability to adequately resolve the Antarctic surface topography.

Although ERA-40 shows low skill in its early years, the improvements with the assimilation of satellite data are remarkable. Of all the statistics used here, ERA-40 shows impressive adjustments as the quantity of assimilated satellite data increases, converging to a skill level during the modern satellite era (post 1978) that puts its performance level above NCEP1. NCEP1, on the other hand, appears to be more constrained by the abundance of surface and radiosonde (conventional) data. Its higher performance than ERA-40 in the relatively dense data areas of the Drake Passage, along the southern extents of the major continents (Cape Town and Buenos Aires), and the southern islands indicate this fact. Differences in the assimilation schemes between ERA-40 and NCEP1 likely account for the large portion of the changes noted in this study. It is clear how the assimilation system handles the satellite data in ERA-40; with NCEP1, the inference as to what the assimilation system is doing is less obvious as little improvement occurs at the start of the modern satellite era. Also, the model climatology is clearly better in ERA-40 than in NCEP1. This is shown throughout this study by the low biases in ERA-40 in the early years, when ERA-40 is strictly following the background fields, therefore explaining the low correlations (A. Simmons 2004, personal communication). Conversely, NCEP1 had much larger biases when observations were sparse, and thus the statistics reflect a poorer model climatology.

The reliability of NCEP1 and ERA-40 before the early 1970s is questionable, however, there is no doubt that ERA-40 does an excellent job after ~1978. Also, most of the results shown here detail the problems during the

austral winter. In DJF, ERA-40 and NCEP1 perform with much higher skill, comparable to the skill seen in the more observationally dense continental stations during winter (Fig. 8). As such, austral summer studies can be extended back into the earlier years of these reanalyses (with a working knowledge of the limitations); however, care must be exercised in using the early data across the high southern latitudes and Antarctica for the other seasons, especially winter. The improvements in the assimilation scheme observed in ERA-40 by its large adjustment to the satellite data clearly indicate that reanalysis projects are taking steps in the right direction. However, additional efforts are needed in enhancing the observational (both conventional and satellite) database and in tuning the assimilation schemes before reliable data assimilation can be conducted for the presatellite era during the nonsummer months.

Acknowledgments. This research was funded in part by NSF Grant OPP-0337948. The authors would like to thank the three anonymous reviewers for their valuable comments and suggestions. Discussions with Adrian Simmons of ECMWF were also very helpful and appreciated.

REFERENCES

- Bromwich, D. H., A. N. Rogers, P. Kållberg, R. I. Cullather, J. W. C. White, and K. J. Kreutz, 2000: ECMWF analysis and reanalysis depiction of ENSO signal in Antarctic precipitation. *J. Climate*, **13**, 1406–1420.
- , S.-H. Wang, and A. J. Monaghan, 2002: ERA-40 representation of the Arctic atmospheric moisture budget. *Proc. ECMWF Workshop on ReAnalyses*, ERA-40 Project Rep. Series 3, Reading, United Kingdom, ECMWF, 287–298.
- , A. J. Monaghan, K. W. Manning, and J. G. Powers, 2005: Real-time forecasting for the Antarctic: An evaluation of the Antarctic Mesoscale Prediction System. *Mon. Wea. Rev.*, in press.
- Comiso, J. C., 2000: Variability and trends in Antarctic surface temperatures from in situ and satellite infrared measurements. *J. Climate*, **13**, 1674–1696.
- Cullather, R. I., D. H. Bromwich, and M. L. VanWoert, 1996: Interannual variations in Antarctic precipitation related to El Niño–Southern Oscillation. *J. Geophys. Res.*, **101D**, 19 109–19 118.
- Genthon, C., G. Krinner, and M. Sacchettini, 2003: Interannual Antarctic tropospheric circulation and precipitation variability. *Climate Dyn.*, **21**, 289–307.
- Gibson, J. K., P. Kållberg, S. Uppala, A. Hernandez, A. Nomura, and E. Serrano, 1997: ERA Description. ERA-40 Project Rep. Series 1, ECMWF, 72 pp.
- Hines, K. M., D. H. Bromwich, and G. J. Marshall, 2000: Artificial surface pressure trends in the NCEP–NCAR reanalysis over the Southern Ocean and Antarctica. *J. Climate*, **13**, 3940–3952.
- Jones, J. M., and M. Widmann, 2003: Instrument- and tree-ring-based estimates of the Antarctic Oscillation. *J. Climate*, **16**, 3511–3524.
- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. 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. Meteor. Soc.*, **83**, 1631–1643.
- Key, J. R., D. Santeck, C. S. Velden, N. Bormann, J. N. Thepaut, L. P. Riishojgaard, Y. Q. Zhu, and W. P. Menzel, 2003: Cloud-drift and water vapor winds in the polar regions from MODIS. *IEEE Trans. Geosci. Remote Sens.*, **41**, 482–492.
- King, J. C., 2003: Validation of ECMWF sea level pressure analyses over the Bellingshausen Sea, Antarctica. *Wea. Forecasting*, **18**, 536–540.
- Kistler, R., and Coauthors, 2001: The NCEP–NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.*, **82**, 247–267.
- Marshall, G. J., 2002: Trends in Antarctic geopotential height and temperature: A comparison between radiosonde and NCEP–NCAR reanalysis data. *J. Climate*, **15**, 659–674.
- , 2003: Trends in the Southern Annular Mode from observations and reanalyses. *J. Climate*, **16**, 4134–4143.
- , and S. A. Harangozo, 2000: An appraisal of NCEP/NCAR reanalysis MSLP data viability for climate studies in the South Pacific. *Geophys. Res. Lett.*, **27**, 3057–3060.
- Parish, T. R., and J. J. Cassano, 2003: The role of katabatic winds on the Antarctic surface wind regime. *Mon. Wea. Rev.*, **131**, 317–333.
- Pavolonis, M. J., and J. R. Key, 2003: Antarctic cloud radiative forcing at the surface estimated from the AVHRR Polar Pathfinder and ISCCP D1 datasets, 1985–93. *J. Appl. Meteor.*, **42**, 827–840.
- Renfrew, I. A., J. C. King, and T. Markus, 2002: Coastal polynyas in the southern Weddell Sea: Variability of the surface energy budget. *J. Geophys. Res.*, **107**, 3063, doi:10.1029/2000JC000720.
- Simmonds, I., 2000: Size changes over the life of sea level cyclones in the NCEP reanalysis. *Mon. Wea. Rev.*, **128**, 4118–4125.
- Simmons, A. J., and Coauthors, 2004: Comparison of trends and variability in CRU, ERA-40, and NCEP/NCAR analyses of monthly-mean surface air temperature. ERA-40 Project Rep. Series 18, ECMWF, 42 pp.
- Thompson, D. W. J., and S. Solomon, 2002: Interpretation of recent Southern Hemisphere climate change. *Science*, **296**, 895–899.
- Trenberth, K. E., D. P. Stepaniak, and J. M. Caron, 2002: Accuracy of atmospheric energy budgets from analyses. *J. Climate*, **15**, 3343–3360.
- WCRP, IPAB, 2002: Report on the Third Meeting of Programme Participants. WCRP Informal Rep. 5/2002, 1–5 and appendixes.
- Zwally, H. J., J. C. Comiso, C. L. Parkinson, D. J. Cavalieri, and P. Gloersen, 2002: Variability of Antarctic sea ice 1979–1998. *J. Geophys. Res.*, **107**, 3041, doi:10.1029/2000JC000733.