# JP2.9 STRONG TRENDS IN THE SKILL OF THE ERA-40 AND NCEP/ NCAR REANALYSES IN THE HIGH AND MIDDLE LATITUDES OF THE SOUTHERN HEMISPHERE, 1958-2001

David H. Bromwich[1,2*] and Ryan L. Fogt[1,2]
[1] Polar Meteorology Group, Byrd Polar Research Center, The Ohio State University, Columbus, OH
[2] Atmospheric Sciences Program, Department of Geography, The Ohio State University, Columbus, OH

## 1. Introduction

The Southern Ocean and the Antarctic continent represent perhaps the largest spatial 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. The National Centers for Environmental Prediction / National Center for Atmospheric Research 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, 1957-1958).

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. Hines et al. 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. Marshall (2002) also noted 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-year global reanalysis, ERA-40, spanning September 1957- August 2002 (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. Further, ERA-40 is a second generation reanalysis, having been preceded by the 15-year reanalysis (1979-1993) completed earlier by ECMWF, ERA-15 (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 examines the change of reanalysis skill throughout time. Section 4 extends the comparisons beyond Antarctica and the Drake Passage. Section 5 provides a discussion, and conclusions are drawn Section 6.

## 2. Data and Methodology

To validate each reanalysis, monthly mean observations from the several stations in the mid to high southern latitudes were compared with the reanalysis values. Observations of MSLP were compared against bilinearly interpolated reanalyses values for all stations. The stations were chosen based on data completeness, yet some comparisons are unavoidably affected from poor data quality and are identified when needed. The station observations for Antarctica were obtained from the British Antarctic Survey READER project website (http://www.antarctica.ac.uk/met/READER/). The MSLP data for the remaining stations were obtained until 1998 through the NCAR ds570.0 data set.

## 3. Time Evolution

Here the correlation, bias, and root mean squared error (RMSE) are presented using five-year windows. In this method, the statistics are calculated as before over a five-year 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. The time series are constructed for only the JJA data since this season appears as the most problematic in the annual cycle (not shown). 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 1958-1972 and represents the 15 years before any satellite data were assimilated into ERA-40. The second era, 1973-1978, represents the years when the Australian surface bogus pressure data, PAOBS, were assimilated into ERA-40 from gridded Australian surface pressure analyses (A. Simmons, personal communication, 2004). More importantly, during this period satellite sounder data were first assimilated into ERA-40. The Vertical Temperature Profile Radiometer (hereafter, VTPR) sounding data were assimilated into ERA-40 starting on 1

---

[*] *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.1@osu.edu

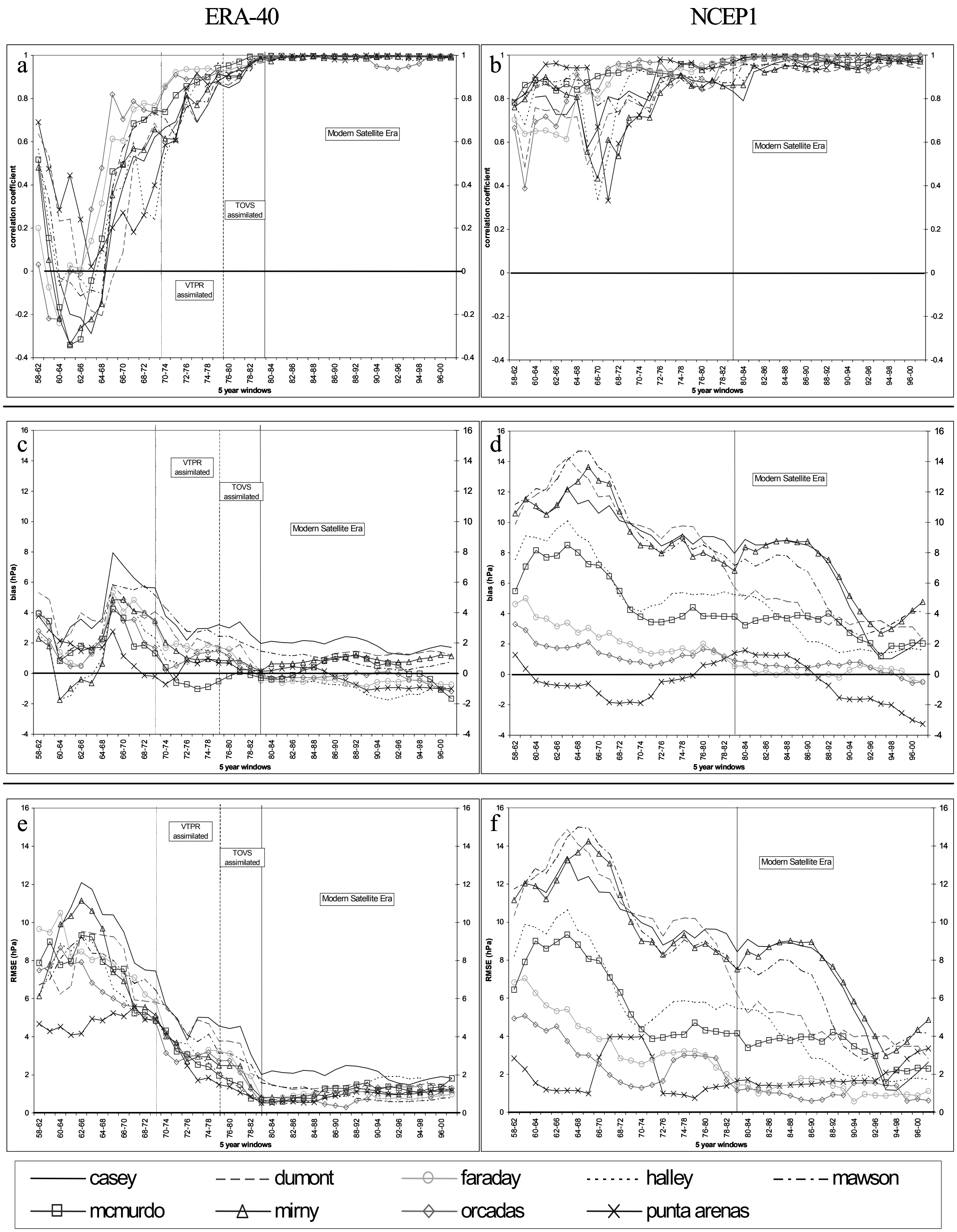


Figure 1. JJA 5-yr running window comparison between observed and reanalysis MSLP values (1958-2001) for (a),(c),(e) ERA-40 and (b),(d),(f) NCEP1. (a),(b) correlation; (c),(d) bias; (e),(f) rmse.

January 1973, while the TIROS Operational Vertical Sounder (hereafter, TOVS) data entered in late 1978. The final era, 1979-2001 represents the years 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.

### *3.1 MSLP in Antarctica and High Southern Latitudes*

The five-year running statistics for MSLP at Antarctic and high latitude stations for both ERA-40 and NCEP1 in JJA are presented in Fig. 1. The correlations, biases, and RMSE are plotted side-by-side for ERA-40 and NCEP1 to facilitate comparison of the two reanalyses.

From Figs. 1a and b one can clearly see the problem with ERA-40's ability to capture the monthly variability during JJA. Correlations during the mid-1960s 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 sounding data begin to be assimilated into the five-year 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. During the modern satellite era, ERA-40 is superior to NCEP1, although both have high correlations > 0.9.

Examining the biases in Figs. 1c and 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 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. 1d, 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 two 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. 1e and 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 pre-satellite 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 six 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 which demand precision, such as those involving the Antarctic Oscillation. Similar trends exist for other dynamic variables such as 2-m temperature and the 500 hPa geopotential heights, which lies above the vastly varying Antarctic topography.

### *3.2. MSLP in Southern Mid-Latitudes*

Up to this point, ERA-40 has shown to have some major shortcomings before the assimilation of satellite data with problems around Antarctica where there are known deficiencies in NCEP1. Due to the strong convergence towards 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. 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 deficiency in ERA-40, but are instead related to the quantity of the observational data assimilated. Due the fact that stations over the continental mainland (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 were selected based on their location (farthest south, global representation) and data quality / completeness. Statistics (using 5-year running windows as before) are displayed in Fig. 2.

Clearly there is a large improvement in skill at these selected stations. Although the correlations of ERA-40 vs. observations in Fig. 2a suggest a similar problem as in the Antarctic and Southern Ocean stations, Fig. 2a is plotted on a different scale; correlations are consistently above 0.5. Further, 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 lack of nearby observations. Additionally, the problems at Macquarie Island in Fig. 2a in the early stages also appear in NCEP1 (Fig. 2b), and could be thus 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

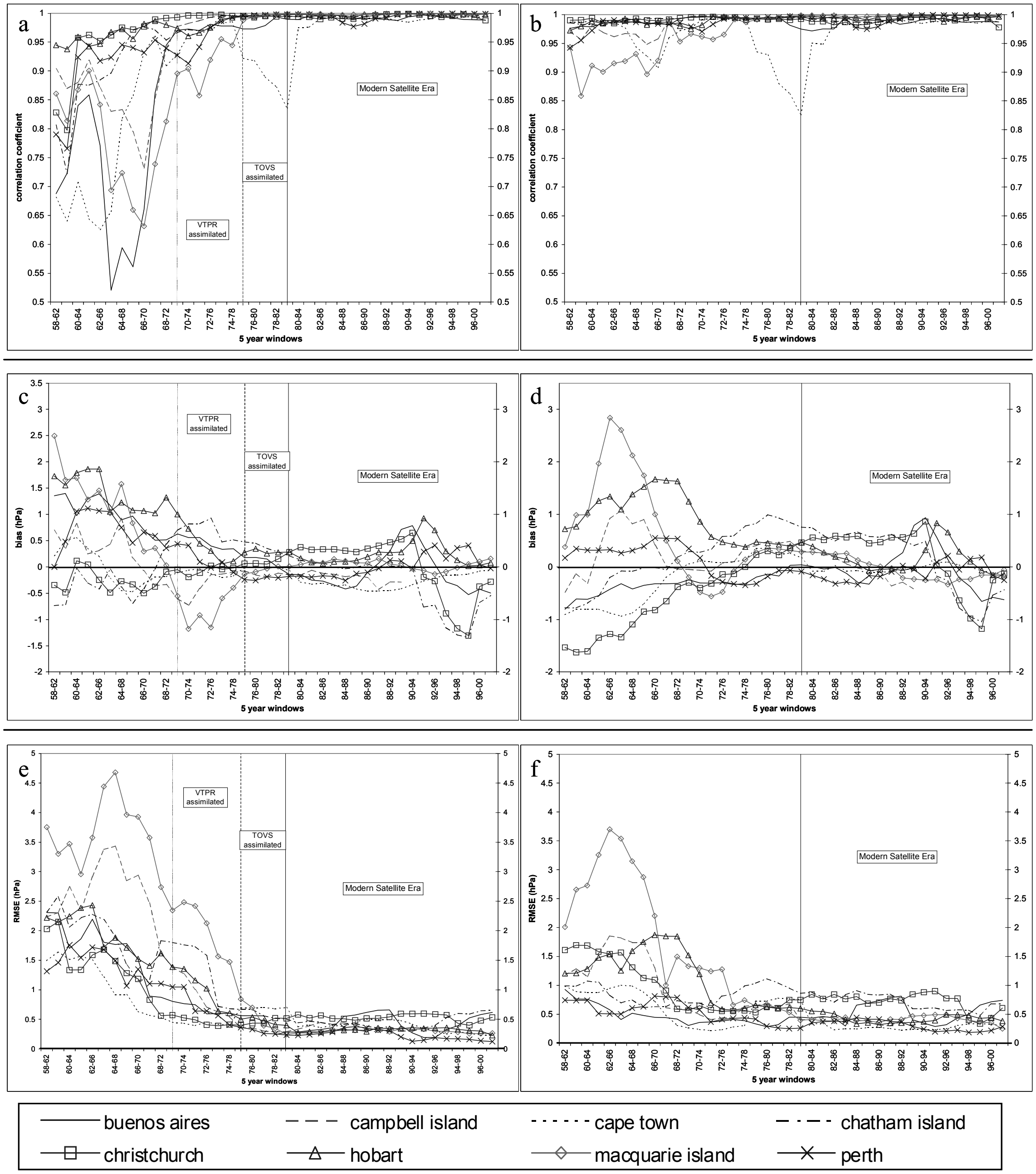


Figure 2. Same as in Fig.1, but for southern mid-latitude stations.

aren't readily explained, especially since correlations at Hobart, Tasmania, and Christchurch, New Zealand—not on major continents—are consistently at or above 0.95 (Fig. 2a). The noticeable dip in correlation values in Figs. 2a and b at Cape Town right around 1979-1983 is also likely due to the quality of the observational data.

The biases are small throughout and comparable between ERA-40 and NCEP1. 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 five years or so not necessarily reflecting the true skill of the reanalyses. The trend in skill is seen in the RMSE plots (Fig. 2e and 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. 2e), 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. 2f).

## 4. Discussion

It is important to note that the comparisons are only made at single points in the Southern Hemisphere where observational data sets are available; comparisons

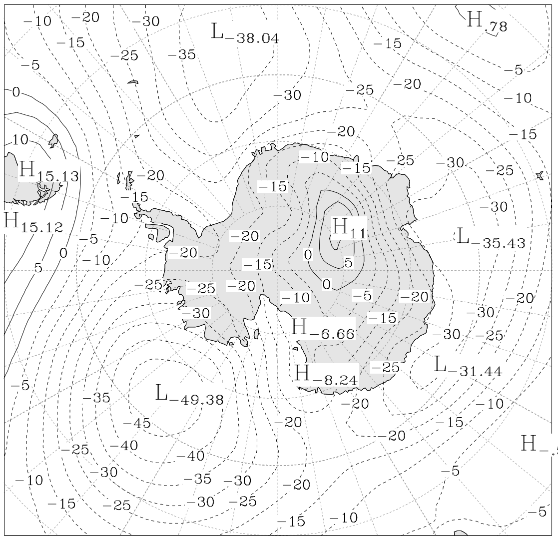

Figure 3. Average ERA-40 minus NCEP1 500-hPa geopotential height difference for all months (1958-78) in gpm.

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 oceans 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. 3). 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 (60º-70ºS, 130º-150ºW, i.e., the center of greatest difference in Fig. 3), a time series of bias between ERA-40 and NCEP1 is produced (Fig. 4). 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).

## 5. 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 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 mid-troposphere above the Antarctic continent, and thus

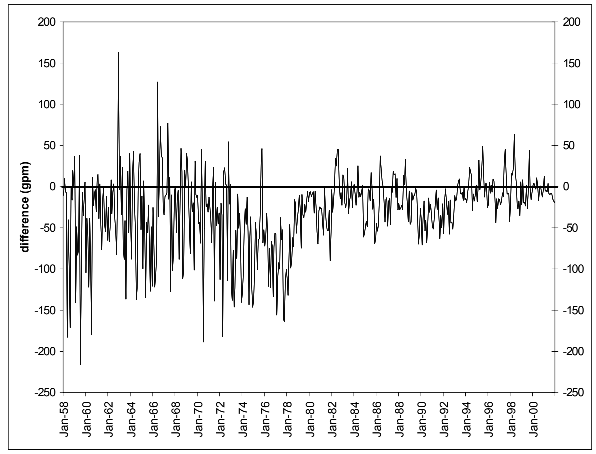


Figure 4. ERA-40 minus NCEP1 differences for 500-hPa geopotential height averaged in a South Pacific Ocean box for all months (60º-70ºS,130º-150ºW).

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, personal communication, 2004). 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. 2). 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) data base and in tuning the assimilation schemes before reliable data assimilation can be conducted for the pre-satellite era during the non-summer months.

*Acknowledgments.* This research was funded in part by NSF Grant OPP-0337948. Discussions with Adrian Simmons of ECMWF were very helpful and appreciated.

**References**

Gibson, J.K., P. Kållberg, S. Uppala, A. Hernandez, A. Nomura, and E. Serrano, 1997: ERA Description. ECMWF Re-analysis Project Report 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.

Kalnay, E. et al., 1996: The NCEP-NCAR 40 Year Reanalysis Project. *Bull. Amer. Meteor. Soc*., **77**, 437-471.

Kistler, R. et al., 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.

Marshall, G. J. and S. A. Harangozo, 2000: An appraisal of NCEP/NCAR reanalysis MSLP data viability for climate studies in the South Pacific. *Geophys. Res. Letts.*, **27**, 3057-3060.