Storm track variations as seen in radiosonde observations and reanalysis data

Nili Harnik

Lamont-Doherty Earth observatory
61 Route 9W
Palisades, NY, 10964
email: nili@ldeo.columbia.edu
phone: 845-365-8493
fax: 845-365-8736

and

Edmund K. M. Chang

ITPA/MSRC
Stony Brook University, SUNY
Stony Brook, NY 11794-5000
email: kmchang@notes.cc.sunysb.edu

Submitted to *J. Climate.*
Revised July 15, 2002

Abstract

The interannual variations in the northern hemisphere storm tracks between 1949-1999 based on unassimilated radiosonde data is examined, and compared to similarly derived quantities using the NCEP/NCAR reanalysis at sonde times and locations. This is done with the motivation of determining the extent to which the storm track variations in reanalysis data are real. Emphasis is placed on assessing previous findings, based on NCEP/NCAR reanalysis data, that both storm tracks intensified from the 1960's to the 1990's with much of the intensification occurring during the early 70's, and that the Atlantic and Pacific storm tracks are significantly correlated.

Sonde data suggests the Atlantic storm track intensified during the 1960's to 1990's, but the intensification was weaker than the reanalysis suggests. The larger trend in reanalysis is due to an overall decrease in biases with time. In the Pacific storm track entrance and exit regions sonde data shows notable decadal time scale oscillations, similar to the reanalysis, but no significant overall positive trend. Sonde does show a positive trend over Canada, consistent with a Pacific storm track intensification and north-eastward shift, but lack of data over the storm track peak prevents drawing any strong conclusions. The biases in the reanalysis are found to have a strong spatial pattern, with the largest biases being over the Pacific entrance region (Japan).

The correlation between the Atlantic and Pacific storm tracks in sonde data (which exists mainly over the storm track entrance and exit regions) is not as significant as in the reanalysis, with differences being mostly due to the decadal time scale variations of the storm tracks rather than their year-to-year variations.

1 Introduction and motivation

The mid-latitude storm tracks are the locations where baroclinic cyclones prevail. There are two popular ways to define the storm tracks. The first methodology takes a synoptic view and defines storm tracks based on cyclone count statistics (e.d. Pettersen, 1956; Whitaker and Horn, 1984). This certainly makes sense since these cyclones are the systems that significantly impact the weather of the regions over which they pass through, especially during the cool seasons. An alternative definition takes the storm tracks to be regions where transient eddy variance/covariance statistics are maximal (e.g., Blackmon, 1976; Lau, 1978). Such a definition is based on the fact that these cyclones are associated with deep baroclinic waves which transport heat, moisture, and momentum which significantly impact the planetary scale flow field (for a review, see Chang et al., 2002). While these two definitions are not necessarily congruent (Wallace et al., 1988, Paciorek et al., 2002), in the Northern Hemisphere they both pick out the mid- latitude regions over the oceanic basins to be the most prominent regions. Here, storm tracks will be defined based on eddy variance statistics.

In this paper we compare the storm track variations computed based on the NCEP/NCAR (National Center for Environmental Prediction/ National Center of Atmospheric Research) reanalysis data set (Kalnay et al. 1996; Kistler et al. 2001) to those computed using observations based on the unassimilated radiosonde data, in order to determine whether the decadal time scale variations of the intensity of the storm tracks, observed in reanalysis data, are real.

In particular, we want to verify recent results of Chang and Fu (2002, referred to as CF) who showed that both the Pacific and Atlantic storm tracks intensified by nearly 40% from the 1960's to the late 1980's, early 1990's. Based on the reanalysis data, the leading EOF (empirical orthogonal function) of the storm tracks (defined as the variance of high-pass filtered meridional velocity perturbation at 300 hPa level) is found to be a mutual intensification of the Pacific and Atlantic storm tracks, and the corresponding PC (Principal component) time series shows a transition from weak storm tracks before 1971 to stronger storm tracks after 1975. CF also showed a significant correlation between the individual Atlantic and Pacific storm track PCs, even on a month-to-month basis, suggesting the mutual intensification is not an artifact of the EOF analysis but rather that the two storm tracks vary to some extent together.

The need for a verification of CF's results using sonde data arises because there may be biases in the reanalysis that affect the observed variations in the storm tracks. These biases can arise for the following reasons: Observations are sparse or nonexistent near the peaks of the storm tracks, hence they do not constrain the reanalysis very strongly (e.g. Chang, 2000, showed that the storm track intensity over the Southern Hemisphere, where observations are sparse, differ significantly between the NCEP/NCAR and ECMWF reanalyses); while the reanalysis minimizes the RMS difference from sonde observations, the intensity of the storm tracks is measured by eddy variance, which is not necessarily optimally analyzed (see appendix A); large changes in sonde and aircraft measurement coverage over the years might introduce large biases in the analysis (Ebisuzaki and Kistler, 1999; Kistler et al. 2001).

Several other studies also suggested that there has been a secular trend in the Northern Hemisphere storm track activity. Based largely on cyclone counts statistics computed from the NCEP/NCAR reanalysis data, Geng and Sugi (2001) suggested that the cyclone activity over the northern North Atlantic, in terms of cyclone density, cyclone deepening rate, as well as cyclone central pressure gradient, all exhibit a significant intensifying trend along with a decadal timescale oscillation in winter during the past 40 years. Meanwhile, Graham and Diaz (2001) suggested that the frequency and intensity of extreme cyclones over the North Pacific Ocean over the past 50 years has increased markedly. In addition, Paciorek et al. (2002) also found an increase in the number of intense cyclones over both oceanic regions. On the other hand, Gulev et al. (2001), using a somewhat different cyclone tracking technique, found a significant decrease in the number of cyclones over the Pacific, with no significant trend in the number of intense cyclones over that sector, and significant decrease in both the number of cyclones and intense cyclones over the Atlantic sector, though the number of intense cyclones are found to have increased in the Arctic. However, their results did suggest an increase in the average intensity and deepening rates of cyclones. What leads to the discrepancies between the results of Gulev et al. and those of the other groups is not entirely clear, but it could be due to differences in the tracking routines (see the appendix in Gulev et al. 2001), differences in the definition of cyclone frequencies and intense cyclones, and/or differences in the geographical extent of the averaging areas.

While it is not clear how the eddy variance and covariance statistics relate to cyclone counts, we expect that any biases in eddy variances could also show up as biases in cyclone intensity and thus could also affect cyclone statistics based on counts of intense cyclones. The results of Graham and Diaz (2001), Geng and Sugi (2001), and Gulev et al. (2001) are based largely on NCEP/NCAR reanalysis data, though Graham and Diaz (2001) did provide supplementary support by comparison with some in situ observations. Clearly, more comparisons need to be made between NCEP/NCAR reanalysis data and unassimilated observations to verify the trend displayed in the reanalysis data set.

CF analyzed data from 21 sonde stations along the storm tracks. They found evidence for the intensification of the storm tracks, but weaker, suggesting there could be some biases in the reanalysis data which changed with time. In this paper we compare the interannual variations of the storm tracks in sonde and reanalysis, using all sonde observations that went into the reanalysis between 1949-99. Using EOF analysis we also examine whether the mutual variation of the two storm tracks, and the transition from a weak to strong storm track in the early 70's, are found in sonde data.

2 Data and diagnostics

We use archived NCEP/NCAR radiosonde data from all stations that reported during 1949-1999, in the latitude range of 20-80N. The data set includes ship reports and land stations. The observational data archive was produced as part of the reanalysis project (Kistler et al, 2001).

To diagnose storm track strength we use the 300 hPa meridional winds variance, computed using a 24 hour difference filter (Wallace et al. 1988), denoted by V_{1df}, and calculated as follows:

$$V_{1df} = \overline{(V_{(t+24h)} - V_{(t)})^2} \tag{1}$$

where $\overline{(\ )}$ denotes an average over all available observation times. CF showed that this measure of storm track strength is comparable to more common diagnostics. We use it because it can easily be applied to time series with observation gaps. We present results for winter mean (December-February) statistics, and note that calculations using January data give similar results. We denote each winter by the corresponding January year (meaning our data set starts December 48).

Most stations report every 12 hours, some report every 6 hours, and occasionally a station report contains missing or bad data, which is flagged in the raw data set. We filter out obviously bad sonde reports, that exceed 80m/sec and differ from the corresponding reanalysis grid point value by more than 30m/sec. Only a tiny percent of the observations are filtered, and most of these are clearly bogus reports of winds much stronger than 100m/sec. Similar calculations with different filtering limits, give similar results.

The NCEP/NCAR reanalysis pressure level data set comprises of 6 hourly data on a 2.5x2.5 grid. For comparison, we compile a gridded sonde data set (SONDE) as follows. For each reanalysis grid box we search the sonde data set for all stations that are located within it. Most of the grid boxes that contain an observing station have only one. When there are more than one station in a grid box, we use data from the station with the largest number of valid reports during the month, and fill in data gaps using other stations. We also repeated some of the calculations using an average of all reports in a grid box and found only tiny differences.

We only use winter statistics for stations that reported valid meridional winds more than half the time (90 reports in DJF, assuming 12 hour reports) and more than a quarter of the time during individual months (more than 15 reports during December and January, and more than 14 reports during February). This threshold is reasonable in comparison with other studies using sonde data (Kidson and Trenberth, 1988). Figure 1 (top) shows the number of years of sufficient sonde observations for each grid box, superposed over the mean storm track strength (1949-99 mean of V_{1df}). We see that there is very little coverage (temporal and spatial) over the peaks of the storm tracks, in particular, the Pacific one. Our results, therefore, are based on the storm track entrance and exit regions, which do have a decent coverage.

For direct comparison with the reanalysis, we repeat our calculations using the reanalysis data only at grid points and synoptic times that have SONDE data. We refer to this sampled reanalysis data set as RSAMP.

3 Results

In the following section we examine the variability of the storm tracks as seen in two kinds of analyses, one based on the full data sets (section 3.1), and the other based on their 1st EOFs and corresponding PC time series (section 3.2).

3.1 The decadal and interannual changes in the intensity of the storm tracks

CF showed that the first PC of the storm tracks computed based on reanalysis data, increased significantly from the 60's to the 90's. They also showed that this intensification is representative of the evolution of the storm tracks, by looking at the difference in storm track intensity, between the decades with strongest and weakest mean PC values (1990-1999 and 1962-1971, correspondingly). We start by verifying that these results hold for sonde data. Figure 2 shows the decadal mean V_{1df} for each of these decades, for SONDE (colored rectangles) and the reanalysis (REAN, contours), using grids that have at least 7 years of sonde data during the decade. Based on the reanalysis data, we see that the storm tracks intensified, with the Atlantic storm track shifting north by about 5° and the Pacific storm track exit shifting south by 5°. During the high decade there are no observations over the peaks of the oceanic storm tracks, and during the weak decade there is only one weather ship near the peak of the Pacific storm track, and relatively good coverage of the Atlantic storm track, but no station at the peak itself. From Fig. 2, it is apparent that the SONDE variance agrees well with reanalysis data during the 1990s, but the agreement is not as good during the earlier decade, when the reanalysis appears to be biased low.

Figure 3 shows the difference between the 1990-99 and 1962-71 decadal means (denoted by $\Delta_{10}V_{1df}$), for SONDE and the sampled reanalysis, RSAMP, for grid boxes that have V_{1df} observations for 7 or more years of each decade. We see that the reanalyses shows stronger storm tracks in the 90's compared to the 60's, while SONDE shows an intensification in Northern Europe, Western USA, Western Canada and the Northern part of Japan, but the intensification is weaker than in the reanalysis. Over Central and Eastern USA, Southern Europe along the Mediterranean coast, and the southern part of Japan, SONDE shows no significant difference in storm track strength between the two decades, while RSAMP shows some intensification. Since the two data sets, SONDE and RSAMP, have the same spatial and temporal sampling, our results suggest that there are biases in the reanalysis.

The differences between sonde and reanalysis data appear to have a spatial pattern. To get a more detailed picture, we look at area averages of V_{1df}. This also allows us to study the time evolution of the storm tracks, and determine how representative the decadal mean differences are of the full time evolution.

The averaging areas we use are shown in the bottom plot of Fig. 1, and their geographical characteristics are listed in table 1. The areas were defined to represent different regions of the storm tracks, taking into account the distribution of observations. For reference, we also mark all grid points that have more than 25 years of SONDE observations on the figure showing the averaging areas. We also tried other area divisions, and the overall results are the same. Figure 4 shows the mean number of sonde stations that had a sufficient amount of observations in each area (dashed line, in multiples of ten). We see that the number of stations changed quite a lot over the years, with a strong increase in early years and a roughly constant number afterwards, with a decrease in the 90's for some areas. We also define an area-coverage index by dividing each area into 9 (3x3) roughly equal sub-areas and counting how many sub-areas have at least one station with a sufficient number of reports, for each year. Figure 4 shows the coverage index for each of the averaging areas (dots). We see that the index changes drastically during the early 50's in most areas, before 1965 in W6, and for a few single years in W1 and W8. Since large changes of this index indicate possible trends due to changes in area coverage, we only look at area averages from 1958 onwards, and regard area 6 data with caution before 1965. Note that a substantial increase in the number of rawindsonde stations in early years occurred after the International Geophysical Year in 1957 (Kistler et al, 2001).

Area averages are calculated as a simple average of all observations in a given area. Figure 5 shows the yearly time series of the area averages of SONDE and RSAMP, for six of the areas representing the main storm track entrance and exit regions, while their difference is shown in Fig. 7 as the thick dotted line. We do not show results from areas W4, W7, and W8. Unless otherwise noted, the results from area W4 are similar to W5. W7 shows no trend in storm track strength, even in the reanalysis, and area W8 is outside the main storm track region. The linear trends for the period 1958-1999, along with the significance levels are shown in table 2[1].

We see large decadal time scale variations in the storm track strength in SONDE, as well as reanalysis data. The variance in the reanalysis is generally less than that in radiosonde data, and the difference is too large to be accounted for by the presumed observational error of rawindsonde wind observations. Overall, the biases in the reanalysis (best shown in Fig. 7) are larger in the 60s and early 70s than in the 80s and 90s. As a result, the positive trend in storm track strength is weaker in SONDE data than in the reanalysis. Figure 6 shows the 10-year running means of these time series. Comparing the decadal means around 1965 and 1995 (Fig. 6), we see that the decadal mean in the 90s is significantly stronger than that in the 60s for all areas shown, in agreement with CF, but the differences in SONDE data in areas W2, W5, and W9 are substantially smaller than those in RSAMP. Examining the full time

[1]The significance levels are computed using a two-tailed t-test assuming all data are independent.

evolution in more detail, we find that over the Atlantic entrance and Europe (regions W3, W1, W2) the storm tracks show a positive trend in reanalysis and SONDE data, but the SONDE trend is weaker and less significant than RSAMP in areas W2 and W3 and slightly weaker in area W1. In the pacific areas (W5, W6, W9) the reanalysis shows a decadal time scale oscillation with an overall positive trend while SONDE shows the oscillation but almost no trend in area W9, a somewhat weaker and less significant trend in W5, and a large positive trend in area W6.

It is also illuminating to look at the temporal evolution of the mean of V_{1df} over all averaging areas, which is some measure of the strength of the hemispheric storm track, at least over land areas. Figure 8 shows the yearly time series using a simple mean (weighting by the geographic area of the averaging regions makes no difference), for SONDE and RSAMP[2]. To examine the effects of spatial sampling on the area averages, we also calculate them from the full reanalysis data set, weighted by cosine latitude (referred to as REAN), and plot it in Fig. 8 using a dashed line. The first thing to note is the similarity in the year to year variation of the different data sets, which is also remarkably similar to the yearly variability of the first PC of the reanalysis (the correlation between REAN and the dashed line of Fig. 11 is 0.93). On longer time scales, SONDE and the reanalysis data sets differ more. The reanalysis shows a transition in the early 70's from a weaker to a stronger storm track regime. The SONDE storm track, on the other hand, only shows a hint of this transition. There is a strong intensification in the early 70's, leading to a peak in 75. SONDE values are stronger than the reanalysis, as expected, but the SONDE-reanalysis biases are largest in the 60's and early 70's, resulting in an overall trend that is much weaker in sonde data ($1.34\frac{m^2s^{-2}}{year}$ in SONDE compared to $2.52\frac{m^2s^{-2}}{year}$ in RSAMP , both with

[2]The averages shown in Fig. 8 are for an average using all areas except W8. This is done because W8 has a few years of missing data in the 90s. We get similar results, however, when we do include it in the average.

> 0.999 significance level).

Another feature we examine using sonde data is the transition, during the early 70's, from a weak to strong storm track regime, found in the reanalysis PC time series by CF. Since such a transition can result from changes in the reanalysis, it is important to verify it in sonde data. CF found evidence of an intensification in the early 70's in a small subset of the sonde stations, suggesting it is a real feature. The hemispheric storm track time series (Fig. 8) also shows an intensification during 1969-75 in SONDE data, but unlike the reanalysis time series, in SONDE it looks more like a temporary peak than a transition from weak to strong storm track regimes. We also look at the change in the storm tracks between 1971-74 (Fig. 9). Both the sonde and reanalysis data show the storm tracks intensified significantly in 1971-4 over Europe, the eastern Atlantic, USA and central Siberia, and weakened around the western and northern edges of the Pacific storm track. Over the Atlantic storm track exit (eastern Atlantic and Europe) the intensification is weaker in SONDE data.

Its hard to conclude from the existing SONDE data whether the Pacific storm track changed much during 1971-4, since there is very little data from the Pacific storm track maximum region. Over the locations of the weather ships in the eastern Pacific (145W-50N, 140W-30N), the reanalysis shows a weak intensification, but the SONDE data shows no significant intensification. In the SONDE data, the intensification is weaker along the western US coast, but is quite strong over the US.

3.2 An EOF analysis of the variability of the storm tracks

In this section, EOF analysis is used to determine whether the significant correlation between the Atlantic and Pacific storm tracks, and the transition in the early 70's from a weak to a strong storm track regime, both found in reanalysis data (CF), also exist in sonde data. We calculate the EOFs from the eigenvalues of the covariance matrix of the data, and the PC time series are then calculated from the projection of the data onto the EOFs. Each element of the covariance matrix is calculated only from years for which data exists for both of the corresponding grid points. This method does not require any time interpolation of the data. For the results to be meaningful, we only use grid points that have at least 25 years of observations. For the EOF calculations, we weigh each grid point by the square root of cosine latitude, to account for the difference in grid area, and by the square root of the number of years of existing observations (the *variances* are weighted by the cosine of latitude and the number of observations). Figures 10 and 11a show the resulting EOF's and PC time series, for SONDE and RSAMP. For comparison, we also plot the EOF and PC time series calculated by CF from the full reanalysis data set. Table 3 shows the percent of variance accounted for by the first EOFs, and their linear trends.

Comparing RSAMP and the full reanalysis, we see that our analysis does produce meaningful results. The overall EOF pattern that emerges from the scattered grid points of RSAMP is similar to the full reanalysis, with the main differences being somewhat smaller loadings for RSAMP over Eastern Europe, China, and Japan. Over the location of the two weather ships near the peak of the Atlantic storm track, the EOF loadings have a smaller value than the full reanalysis, probably because they only existed in earlier years, when the storm track was weaker. The year-to-year variability of the PC of RSAMP is very similar to the full reanalysis, with their correlation being 0.80. The reanalysis time series is smoother, and correspondingly, RSAMP has more extreme values, most notably the much stronger peak in 1974. In addition, the trend in the reanalysis PC is significantly stronger than the RSAMP one, suggesting the storm track increased more over the ocean than over land[3]. Nevertheless, The

[3]We verify that the differences are not due to temporal sampling by repeating our calculations using the reanalysis at all synoptic time scales, which gives results very similar to the RSAMP PCs.

similarity between RSAMP and the reanalysis also suggests that an EOF analysis based only on land data is indicative of the variability of the full storm tracks. We also performed a series of sensitivity tests that show the robustness of our results. We describe these in appendix B.

Comparing SONDE and RSAMP (in Fig. 10 and 11a) we see that the EOF patterns are remarkably similar, with the largest difference over Eastern Europe, China, and Japan, where the SONDE EOF loading is close to zero, while the reanalysis is small but mostly positive. The SONDE EOF accounts for less of the total variance, compared to the reanalysis (table 3), but the largest differences are in the PC time series. Consistent with our previous findings of an overall smaller positive trend (see table 2), SONDE values are larger than the reanalysis before the mid 70's and smaller after that, with the trend clearly apparent only when looking at the 10 year running means (see Fig. 13). One of the striking features in all the PCs is the strong intensification of the storm tracks during the early 70's, but while in the reanalysis (especially using the full dat set) this coincides with a transition from weak to strong storm track regimes, in SONDE data most of the intensification is temporary, with a small overall intensification that is clearly noticeable only when looking at longer time scales (e.g. Fig. 13).

A more detailed picture is obtained from looking at the individual oceanic storm track EOFs. Figure 12 shows the first EOFs for the Atlantic (ATL) and Pacific (PAC) regions. These are calculated separately, using all grid points between 120°E to 100°W and 90°W to 50°E for PAC and ATL respectively. Table 3 shows the percent variance accounted for by the corresponding PCs, as well as the linear trends. Comparing to the full EOF (denoted by NH), the EOF loadings over Japan are stronger in PAC than in NH, for all data sets. The Atlantic entrance, on the other hand, is weaker in ATL compared to NH. These differences are largest for SONDE and are quite small for RSAMP. The PC time series, including for the full reanalysis, are shown in Fig. 11b,c, and the corresponding linear trends are shown in table 3 (in parentheses). We see again that SONDE values are generally larger than the reanalysis before 1976, and smaller afterwards, resulting in an overall weaker positive trend. We see an intensification of the Atlantic storm track after the early 70's, resulting in an overall trend that is clear even in SONDE data. In the Pacific, however, there does not seem to be much of a trend in SONDE data (a linear fit gives 0.007 standard deviations per year, with a significance level of 0.53, compared to 0.022 standard deviations per year and a significance level of 0.99 for RSAMP). The SONDE time series looks more like a very strong peak in 1974 superposed on an overall very weak positive trend. On the other hand, the trend in RSAMP is only about half of that in the full reanalysis, suggesting a large part of it comes from ocean regions, which are not covered by the SONDE network. An examination of aircraft data, which exists over the oceans, should be useful for resolving this issue. This is currently being done.

This leads us to examine the relation between the Atlantic and Pacific storm tracks. Table 4 shows the correlation between various PC time series. We find that while the Pacific and Atlantic storm tracks are significantly correlated in reanalysis data, with a correlation of 0.39 for RSAMP (0.36 is the 99% confidence level), they are only marginally significantly correlated in SONDE data (0.27, with 0.28 being the 95% significance level). This difference in correlation is a bit surprising, given that the year to year variability of SONDE and RSAMP are quite similar (with SONDE-RSAMP correlations higher than 0.91 for NH, ATL and PAC). The differences become clear when we look at longer time scales. Figure 13 shows the 10 year running means of the NH, ATL, and PAC PC's, for the two data sets. We see that on long time scales, ATL and PAC differ much more in SONDE than in the reanalysis data sets, with an ATL-PAC correlation of the 10 year running means of 0.06 for SONDE, compared to 0.55 for RSAMP[4]. This suggests the reanalysis introduces a spurious correlation between the storm tracks on long time scales. Indeed, when we detrend the PC time series before correlating them, the ATL-PAC correlations drop to 0.28 and 0.25 for RSAMP and SONDE respectively.

One possible source of a spurious correlation on long time scales is the low bias in the early years in the reanalysis, which introduces a spurious "jump" in the time series. We test this by removing a step function from the ATL and PAC RSAMP PC time series, to make the difference between the means of 1949-1972 and 1976-1999 be the same as in the corresponding SONDE PC time series. The resulting correlations are in the ranges of 0.26-0.28 for RSAMP, with the actual values depending on how we vary the time series between 1972-1976, suggesting the time evolution of the reanalysis biases introduces much of the spurious correlation in RSAMP[5].

We note, however, that EOF analysis does not maximize correlations. A better approach is to perform an SVD analysis of the cross covariance matrix of the two storm track regions, since this maximizes the covariance between the two (Bretherton et. al, 1992, see appendix of CF). The resulting SVD patterns (not shown) look similar to the corresponding EOF of each storm track region, with the exit regions being stronger than the entrance regions, especially in the Pacific. The SVD time series are highly correlated with the corresponding NH first PC time series. These correlations, along with the main results of this analysis, are summarized in table 5. We see a significant correlation between the storm tracks for all data sets of around

[4]Note that while ATL RSAMP and PAC SONDE are marginally significantly correlated (0.25, see table 4), PAC RSAMP and ATL SONDE are significantly correlated (0.36 and 0.40), suggesting the correlation in the reanalysis is introduced through the Pacific storm track

[5]These results appear to suggest that the two storm tracks are not significantly correlated – even in the reanalysis data – after the "spurious jump" in the reanalysis bias is removed. However, note that CF computed the correlation between the leading ATL and PAC monthly mean PCs based only on DJF months after 1976 and found significant correlation between the two storm tracks in both NCEP/NCAR and ECMWF reanalysis data – due to significant correlation between the two storm tracks over the oceanic regions, over which we do not have sufficient data to verify.

0.6, with SONDE being slightly larger than RSAMP. The leading SVD mode, however, explains a smaller percentage of the squared covariance in SONDE data (44.6% for SONDE vs 66.5 % for RSAMP respectively). Correspondingly the individual SVD patterns in SONDE data explain a smaller percent of the variance in each individual basin (17.2/23.0 % ATL/PAC for SONDE, compared with 29.2/32.4 % for RSAMP). This supports our earlier conclusion, that the reanalysis introduces some spurious correlation between the storm tracks. Finally, in accordance with our previous findings, the SVD analysis shows a positive long term linear trend for both storm tracks but stronger in the Atlantic, and a significantly weaker trend in SONDE data compared with the reanalysis.

3.3 The effects of spatial and temporal sampling

The comparison of SONDE and RSAMP data gives us an idea of the biases in the reanalysis, but it does not account for all of the differences between observations based on the full reanalysis and raw SONDE data, which may also arise due to the effects of spatial and temporal sampling. We discuss these in the following section. We examine the effects of temporal sampling by repeating our calculations using the reanalysis data at SONDE stations, and all synoptic time scales, and comparing to RSAMP. We also examine the effects of spatial sampling on the area averages, by comparing RSAMP to area averages using the full reanalysis data, weighted by cosine latitude. Their difference, which is plotted as the thin line in Fig. 7, includes the effects of temporal and spatial sampling. We see from Fig. 7 that the effects of spatial and temporal sampling are on the whole smaller than the biases between SONDE and RSAMP, with a few exceptions.

Temporal and spatial sampling result in an over-estimation of the linear trends in area W1 (suggesting the trend may be even smaller than indicated by SONDE), and an under-estimation of the trend in area W3. In area W4 (central US, not shown) the reanalysis shows an overall positive trend, while SONDE does not. Some of this difference is due to spatial and temporal sampling. Temporal sampling, in particular, results in a large under estimation of the storm tracks in the 80's and 90's, suggesting actual SONDE values may be larger than shown during these years, and the linear trend might be larger. One source of bias may be the fact that the reanalysis data is calculated using 4 daily reports, while sonde observations were usually taken only twice a day. A comparison of reanalysis variances based on 00 and 12 Z, vs 06 and 18 Z suggests this does indeed introduce a bias of the observed sign, but it can only account for about a third of the difference. The distribution of the observations between the three winter months is quite constant over the US during most of the observation period, suggesting the source of this bias is not straightforward, and requires further investigation. In area W6, before 1965, both the spatial and temporal sampling were poor (see Fig. 4), resulting in a large under estimation of the storm tracks during early years, which suggests the SONDE trend might be an overestimate.

We also find that temporal sampling results in an under-estimation of the intensification in the early 70s (the difference 1974-1971).

The effects of temporal sampling on the EOF and SVD analyses is mostly to decrease the percentage of variance that the various EOFs and SVD patterns explain, by about 3-4% variance (e.g. from 29% for the first NH EOF using the data set with all synoptic times, to 25% for RSAMP). In addition, sampling decreases the correlation between the Atlantic and Pacific PC time series. For example, we get correlations of 0.39 and 0.46 (0.36 is the 99% confidence level) for the yearly PC time series using RSAMP and the corresponding reanalysis data set using all synoptic times, and correlations of 0.55 and 0.76 for the ten year running mean time series (as in Fig. 13).

4 Discussion and conclusions

To summarize, we find that the intensification of the Atlantic storm track found in the reanalysis data is also found in sonde data from the Atlantic exit and entrance regions, but weaker. The weaker trend is due to the fact that the biases between the reanalysis and sonde data have on the whole decreased with time. In the Pacific regions, the reanalysis shows an interdecadal oscillation superposed on a positive trend. There is strong evidence of the decadal time scale oscillations, but not of the positive trend, in sonde data from the Pacific storm track entrance and exit regions. It is possible that an intensification of the Pacific storm track did occur, but the sonde data is too sparse to say anything about it. A north eastward shift along with the intensification could, in that case, explains the fact that a positive trend is not observed over Japan and the central US (W4), is observed to be weak over the west coast of the US (W5), and is quite strong over Canada (W6).

In this paper, we find a major change in the difference in variances between the sonde data and reanalysis data occurring during the early 1970s. During that time, there was no major change in the observational network itself (apart from increasing number of satellite observations, which does not have a significant impact on the analysis over the Northern Hemisphere. See, e.g. Mo et al. 1995), but in 1973, a new scheme of encoding rawindsonde data (the "NMC Office Note 29" scheme, see Kistler et al. 2001) was introduced, enabling better quality control as well as more efficient error detection and correction (of temperature and height data) based on hydrostatic consistency. Note that this change does not directly affect the SONDE dataset, but it affects the reanalysis in the way the SONDE data are assimilated. Thus any changes in the biases between the two datasets over this time period basically reflects changes in the quality of the reanalysis data. The fact that the forecasts using the reanalysis data set as initial conditions show a jump in skill around that time (Kistler et al. 2001, their Fig. 7) suggest that the quality of the reanalysis is improved after that time, which is consistent with our result that the difference between radiosonde and reanalysis variances is much smaller after the early 1970s (Fig. 7). Thus, our results indicate that the storm track intensity in the reanalysis data set is biased low before the early 1970s, and part of the jump in storm track intensity seen in the PC time series based on the reanalysis data at that time is probably due to this bias.

Nevertheless, the sonde data still display a marked increase in storm track intensity during the early 1970s, especially over the Atlantic sector. This increase is suggestive of a transition between weak and strong storm track regimes (especially in the reanalysis data and less so in the sonde observations), which could be an indication of a larger scale climate regime change. In fact, the NAO index (the normalized pressure difference between Portugal and Iceland) has been anomalously high since the early 1970's, compared to the 1864-1994 time series (Hurrell, 1995). Seager et al. (2000) performed and SVD analysis on SST and surface winds for the Atlantic basin, and the SST time series showed a transition from low to high values in the early 1970's (their Fig. 6c). CF also found that the Atlantic storm track intensity is well correlated with the NAO index (see also Geng and Sugi, 2001) as well as the Arctic Oscillation Index (Thompson et al. 2000). There is also much evidence of a climate regime change in 1977, seen most strongly over the Pacific basin, in various indices based on pressure, SST, and ocean subsurface temperature (e.g. Trenberth, 1990, Mantua et al. 1997, Desser et al. 1999, and references therein), but the few year time lead of the storm tracks needs to be explained if these transitions are to be related. Graham and Diaz (2001) suggested that the Pacific storm track intensification could be related to decadal trend observed in ENSO (e.g. Zhang et al. 1997); but CF found that the Pacific storm track intensity is only weakly correlated with the Southern

Oscillation Index (SOI), with much of the decadal variability and trend remaining even when storm track variations which are linearly congruent with the high and low frequency components of the SOI have been taken out. Another possibility is that the intensification in the early 70's led to a temporary peak in storm track strength, which is not associated with any regime transition, and an observed intensification over the past 30 years could be an unrelated trend. Based on the SONDE data, the latter possibility appears to be the more likely.

With regard to the correlation between the Pacific and Atlantic storm track anomalies, our analysis of the sonde data do show some correlation between the two storm tracks, especially over the storm track exits (where the SVD patterns of the two storm tracks are largest) but the correlation appears to be somewhat weaker than that shown in the reanalysis data. No definitive conclusion can be drawn for this case, since much of the covariance between the two storm tracks found by CF occur over the oceans, where we do not have enough data to address.

Our results once again demonstrate how the analysis of long term climate variability can be hampered by the lack of consistent, uninterrupted data records, especially over the oceanic regions. The decommissioning of the weather ships over the past few decades made it impossible to verify the trend in storm track intensity over the oceanic storm track peaks based on radiosonde data alone. The large change in biases between the radiosonde and reanalysis variance around the early 1970s, which has a magnitude dependent on the geographical location, casts serious doubts on the magnitude of climate trends determined based on reanalysis data alone. In view of these uncertainties, more efforts have to be made to better determine the variability of the storm tracks over the oceanic regions. Graham and Diaz (2001) made use of some in situ observations as well as wave height measurements and hind casts to infer changes in storm track intensity. While those measurements do suggest an overall increasing trend in Pacific storm track activity, it is difficult to quantify the actual amount that storm track activity has changed based on those observations since they do not directly measure either cyclone activity or eddy variance statistics near the storm track peak. We are in the process of analyzing aircraft observations over the Pacific and Atlantic storm track regions to see whether the trend in variances can be observed in aircraft observations.

Acknowledgments: Much of the work this paper is based on was conducted when the authors were at the Florida State University, where they were supported by NSF grant ATM-0003136, and NOAA grant NA06GP0023. EC was also supported by a CRC award at FSU, and NSF grant ATM-0296076 at SUNYSB. NH would also like to acknowledge hospitality provided by Professor R. Lindzen at MIT.

Appendix

A Possible biases in objectively analyzed variances

The spectral statistical interpolation (see Parrish and Derber, 1992) method used in the NCEP/NCAR reanalysis procedure is based on the principle of least squares estimation (see, e.g. Daley 1993). In this appendix, we will consider least squares estimation of a single variable at a single point based on combining a single observation with a background first guess, in this case assumed to be a forecast from a model with known biases in the mean value which have been removed. Parts of the following discussions may seem trivial, but they are presented here for the sake of completeness. Let t be the unknown truth, a denote the analysis, b the background guess, and o the observation value to be assimilated with b to produce the analysis, such that

$$a = b + w(o - b) \tag{2}$$

where w is the weight to be determined. In least squares estimation, w is chosen a priori such that the expected r.m.s. error in a,

$$rmse = < (a - t)^2 >^{1/2} \tag{3}$$

is minimized. Here $< >$ represents the expectation operator. Under the assumptions that observational and forecast errors are uncorrelated, and that observational and forecast mean biases have been removed, simple manipulations (see, e.g. Daley 1993) give the well known result

$$w = \frac{< (b - t)^2 >}{< (b - t)^2 > + < (o - t)^2 >} \tag{4}$$

Assuming that we know the magnitude of the expected observational and forecast error variances, the analysis for a using equation 2, making use of the weight w from equation 4, is optimized in the sense that the expected analysis error variance, $<(a-t)^2>$, is minimized by this procedure, and is guaranteed to be less than or equal to both the observational and forecast error variances.

However, this procedure does not guarantee that the analyzed variance is necessarily a better estimate of the true variance than the observed variance is. Assuming that biases (in the mean) have been removed, we can subtract the mean value of each quantity from both sides of equation 2, squaring the equation, and taking the expectation values on both sides, we get

$$<a'^2>=(1-w)^2<b'^2>+w^2<o'^2>+2w(1-w)<o'b'> \tag{5}$$

Here, primed quantities denote deviation from the mean value. Let us assume that the observational errors are random, i.e.,

$$o'=t'+r \tag{6}$$

where r is uncorrelated with t' or b'. Further assume that the first guess error is partly correlated with the truth, and partly random – in other words, the forecast has an amplitude error on top of random errors – i.e.,

$$b'=\alpha t'+s \tag{7}$$

where s is uncorrelated with t' (or o'). If α equals 1, the forecast errors and the truth are uncorrelated, but when α differs from 1, forecast errors are correlated with the truth. While one can argue that such a bias can be easily corrected for in this example, in reality, if such a bias occurs only over some limited regions in a global model forecast, or only for some modes of variability out of many modes, this kind of bias is not corrected for (or even easily recognizable) in the current objective analysis procedures.

Using equations 6 and 7, we get

$$< o'^2 > = < t'^2 > + < r^2 >$$

$$< b'^2 > = \alpha^2 < t'^2 > + < s^2 >$$

$$< o'b' > = \alpha < t'^2 > \tag{8}$$

Substituting equation 8 into 5 we get, for this example,

$$< a'^2 > = [w + \alpha(1 - w)]^2 < t'^2 > + w^2 < r^2 > + (1 - w)^2 < s^2 > \tag{9}$$

If α equals 1, equation 9 becomes

$$< a'^2 > = < t'^2 > + w^2 < r^2 > + (1 - w)^2 < s^2 > \tag{10}$$

and, in this case, using equations 4, 6 and 7, the weight w is simply given by

$$w = \frac{< s^2 >}{< s^2 > + < r^2 >} \tag{11}$$

and it is clear from equations 10 and 11 that the expected error in the analyzed variance, $< a'^2 > - < t'^2 >$, is smaller than either $< r^2 >$ or $< s^2 >$.

However, if α is not 1, the expected error in the analyzed variance is no longer guaranteed to be smaller than $< r^2 >$, the expected error in the observed variance.

As a counter example, assume that α equals 1.1, such that $b' = 1.1t' + s$. Hence $< (b'-t')^2 >= 0.01 < t'^2 > + < s^2 >$. Also, assume that $< s^2 >= 0.01 < t'^2 >$, hence the total forecast error variance, $< (b' - t')^2 >= 0.02 < t'^2 >$. Further, assume that the observational error variance equals 8% of the true variance, i.e. $< (o' - t')^2 >=< r^2 >= 0.08 < t'^2 >$. From equation 4, for this example, $w = 0.2$, and we can calculate that the expected analysis error variance, $< (a' - t')^2 >= 0.016 < t'^2 >$, indeed smaller than both the observational or forecast error variances, as expected. However, from equation 9, we find that for this example, $< a'^2 >= 1.176 < t'^2 >$, whereas $< o'^2 >= 1.08 < t'^2 >$, hence the error in the analyzed variance is larger than the error in the observed variance. This is due to the fact that while $< (b'-t')^2 >= 0.02 < t'^2 >$, $< b'^2 >= 1.22 < t'^2 >$, because of non zero correlation between the forecast error and the truth value.

The simple example above shows that the variance estimated from least squares estimation is certainly not guaranteed to be better than the variance estimated based on observations alone, even though the observations are used in the production of the analyses. This is true even in the absence of any changes in the analysis system. In the presence of changes in the analysis system (e.g. changes in observational network), the errors and biases in variances (and covariances) introduced by the analysis system will change with time, thus possibly distorting the temporal evolution of, or even completely masking, climate change signals if changes in variances or covariances are sought.

B Sensitivity tests of the EOF calculations

We repeated the EOF calculations described in section 3.2 with different thresholds for the minimum number of years of observations, ranging from 1 (using all the available data) to 45 years (out of 51 years total). The resulting PC time series are almost identical to each other for thresholds up to 35 years. Higher thresholds result in too many stations being dropped out of the analysis. The corresponding EOFs are also very similar, with the additional grid points for lower thresholds having small values. The largest difference is for SONDE data over Japan, where the EOF loading is small, with a mix of negative (blue) and positive (green) values. As the threshold is increased, the EOF loading over Japan shifts from zero (a roughly even mix of negative and positive points) to a small but positive value. We also calculated the EOFs and PC time series without weighing by the number of years of observations, and with weighing each covariance matrix element by the number of observation pairs that contribute to its calculation, and got very similar results. Finally, we repeated the EOF analysis with a coarse grid version of the data set. The coarse grids used were 5°x5°, 15°x5°, and 30°x10° longitude by latitude, with the corresponding V_{1df} grid values calculated as a simple average over all the existing 2.5°x2.5° V_{1df} values in each coarse grid box. The resulting PCs are surprisingly insensitive to coarse graining the data. The corresponding EOFs have similar overall features, with positive values along the storm track latitude band, with peak values in the oceanic storm track exit regions. The largest sensitivity is over Japan in SONDE data, where the EOF loadings are small, and the relative distribution of negative and positive values changes slightly with the choice of coarse grid. We conclude that our EOF estimates are sufficiently meaningful and robust to use for comparing sonde and reanalysis data.

C References

Bretherton C. S., C. Smith, and J. M. Wallace, 1992: An intercomparison of methods for finding coupled patterns in climate data. *J. Climate* **5**, 541-560.

Blackmon, M.L., 1976: A climatological spectral study of the 500 mb geopotential height of the Northern Hemisphere. *J. Atmos. Sci.* **33**, 1607-1623.

Chang, E. K. M., 2000: Wave packets and lifecycles of troughs in the upper troposphere. Examples from the Southern Hemisphere summer season of 84/85. *Mon. Wea. Rev.* , **128**, 25-50.

Chang, E.K.M., and Y. Fu, 2002: Inter-decadal variations in Northern Hemisphere winter storm track intensity. *J. Clim.* **15**, 642-658.

Chang, E.K.M., S. Lee, and K.L. Swanson, 2002: Storm track dynamics. *J. Climate* **15**, 2163-2183.

Daley, R. 1993: *Atmospheric data analysis.* Cambridge University Press.

Desser, C., M. A. Alexander, and M. S. Timlin, 1999: Evidence for a wind-driven intensification of the Kuroshio current extension from the 1970s to the 1980s. *J. Climate* **12**, 1697-1706.

Ebisuzaki W., and R. and Kistler, 1999: An examination of data-constrained assimilation. Proceedings of the Second WCRP International conference on reanalyses, Wokefield Park, UK, 23-27 August 1999, WMO-TD/No 985. pp 14-17.

Geng, Q, and M. Sugi, 2001: Variability of the North Atlantic cyclone activity in wonter analyzed from NCEP-NCAR reanalysis data. *J. Climate* **14**, 3863-3873.

Graham, N. E, and H. F. Diaz, 2001: Evidence for intensification of north Pacific winter cyclones since 1948. *Bull. Am. Meteorol. Soc.* **82**, 1869-1893.

Gulev, S. K, O. Zolina, and S. Grigoriev, 2001: Extratropical cyclone variability in the northern hemisphere winter from the NCEP/NCAR reanalysis data. *Clim. Dyn.* **17**, 795-809.

Hurrell J. W, 1995: Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation. *Science* **269**, 676-679.

Kalnay, E., and coauthors, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* **77**, 437-471.

Kidson J. W, and K. E. Trenberth, 1988: Effects of missing data on estimates of monthly mean general circulation statistics. *J. Climate* **1**, 1261-1275.

Kistler, R., and coauthors, 2001: The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bull. Am. Meteorol. Soc.* , **82**, 247-267.

Lau, N.-C., 1978: On the three-dimensional structure of the observed transient eddy statistics of the Northern Hemisphere wintertime circulation. *J. Atmos. Sci.* **35**, 1900-1923.

Mantua, N. J., S. R. Hare, Y. Zhang, J. M. Wallace, and R. C. Francis, 1997: A Pacific interdecadal climate oscillation with impacts on salmon production. *Bull. Am. Meteorol. Soc.* **78**, 1069-1079.

Mo, K.C., X.L. Wang, R. Kistler, M. Kanamitsu, and E. Kalnay, 1995: Impact of satellite data on the CDAS-reanalysis system. *Mon. Wea. Rev.* , **123**, 124-139.

Paciorek, C.J., J.S. Risbey, V. Ventura, and R.D. Rosen, 2002: Multiple indices of Northern Hemisphere cyclone activity, winters 1949-1999. *J. Climate* , in press.

Parrish D. F, and Derber J. C, 1992: The National Meteorological Centers's statistical interpolation analysis system. *Mon. Wea. Rev.* **120**, 1747-1763.

Pettersen, S., 1956: *Weather Analysis and Forecasting*, Vol. 1., 2d ed. McGraw-Hill, 422 pp.

Seager, R., Y. Kushnir, M. Visbeck, N. Naik, J. Miller, G. Krahmann, and H. Cullen, 2000: Causes of Atlantic ocean climate variability between 1958 and 1998. *J. Climate* **13**, 2845-2862.

Thompson D. J, J. M. Wallace, and Hegrel, 2000: Annular modes in the extratropical circulation. Part II: trends. *J. Climate* **13**, 1000-1016.

Trenberth, K. E., 1990: Recent observed interdecadal climate changes in the northern hemisphere. *Bull. Am. Meteorol. Soc.* **71**, 988-993.

Wallace, J. M., G. H. Lim, and M. L. Blackmon, 1988: Relationship between cyclone tracks, anticyclone tracks, and baroclinic wave guides. *J. Atmos. Sci.* **45**, 439-42.

Whitaker, L.M., and L.H. Horn, 1984: Northern Hemisphere extratropical cyclone activity for the midseason months. *J. Climatol.* **4**, 297-310.

Zhang Y. M., J. M. wallace and D. S. Battisti, 1997: ENSO-like interdecadal variability: 1900-93. *J. Climate* **10**, 1004-1020.

Table Captions:

Table 1. Geographic characteristics of the averaging areas used in this study (shown in Fig. 1).

Table 2. The linear trend ($\frac{m^2s^{-2}}{year}$) of the area mean time series of V_{1df} from 1958-1999, for SONDE and RSAMP data. Trends with a significance level > 0.95 and < 0.80 are marked by *'s and parentheses, respectively.

Table 3. The percent of variance explained by the first EOF calculated using SONDE, RSAMP, and the full reanalysis data sets, for NH, ATL and PAC, as well as the linear trend of the corresponding PC time series (in parentheses, in units of standard deviation per year).The significance levels of the trends are very high (all are >0.97, with most exceeding 0.99) except for SONDE PAC, for which the significance level is 0.53.

Table 4. The correlations between various PC time series, for the full data set (NH), and the Atlantic (AT) and the Pacific (PC) storm track regions, for SONDE (S), and RSAMP (R_S). Parentheses denote values that are not statistically significant, based on a 95% confidence level, but all other values are above the 99% percent confidence level.

Table 5. Statistics of the SVD analysis of the cross covariance matrix of the Pacific an Atlantic storm track regions. The significance levels of the trends exceed 0.989 for RSAMP and are 0.97/0.62 for SONDE ATL/PAC.

Figure Captions:

Figure 1. Top: The number of years with observations (colored rectangles) plotted at the observation location, and the reanalysis 1949-99 mean V_{1df} (contours). Bottom: The averaging areas used in this study (see text for details), and grids with more than 25 years of SONDE observations (maroon dots).

Figure 2. The decadal means of V_{1df} for the high (1990-99) and low (1962-71) decades, for SONDE (colored rectangles) and the full reanalysis data (contours). SONDE data shown only for grids that have at least 25 years of observations. To facilitate comparison, the same color scheme is used for the reanalysis contours and SONDE rectangles.

Figure 3. The difference between V_{1df} of the high and low decades of Fig. 2, for SONDE (top), and RSAMP (bottom). Data is shown only for grids with at least 7 years of observations in each of the decades.

Figure 4. The yearly time series of the total number of stations with sufficient observations, in multiples of 10 (dashed), and the area coverage index for the areas shown in Fig. 1 (solid dots). See text for definition of the index.

Figure 5. Yearly time series of the area averages of 300hPa V_{1df}, for areas W1, W2, W3, W5, W6, and W9 of Fig. 1. Shown are SONDE (dotted line) and RSAMP (thin line).

Figure 6. The 10 year running means of the time series in Fig. 5 (using the same line types).

Figure 7. SONDE minus RSAMP yearly time series (the difference between the curves in Fig. 5, dotted line) and REAN minus RSAMP (thin line).

Figure 8. The average of all area mean V_{1df} time series except for W8, for SONDE and RSAMP as well as for the corresponding full reanalysis curve (REAN). SONDE (dotted line), RSAMP (thin line), and REAN (dashed).

Figure 9. The relative change in V_{1df} during the early 70's - (1974-1971)/1971, for SONDE (top), and RSAMP (bottom).

Figure 10. The first EOF of SONDE (top) and RSAMP (bottom) data, only using grid points with more than 25 years of observations (colored rectangles). Also shown in the bottom plot is the first EOF of the full reanalysis data (contours). The EOFs are normalized to represent one standard deviation. To facilitate comparison, the same color scheme is used for SONDE and RSAMP and the reanalysis.

Figure 11. The first PC time series corresponding to the EOFs shown in figures 10 (top) and 12 (Atlantic in the middle and Pacific at the bottom). SONDE (dotted line), RSAMP (thin line), and the full reanalysi (dashed). PCs are in units of standard deviation.

Figure 12. The same as in Fig. 10, but for the first EOFs calculated separately for the Atlantic and Pacific storm track regions.

Figure 13. The 10 year running means of the PC time series of Fig. 11, for SONDE (top) and RSAMP (bottom). Curves are for NH (+), ATL (o), and PAC ($*$).

	Region	Storm track part	Longitude range	Latitude range
W1	East. Europe	Atl exit - downstream	16.25E-63.75E	43.75-71.25
W2	West. Europe	Atl exit	11.25W -16.25E	43.75-61.25
W3	East. USA	Atl entrance	83.75W-56.25W	33.75-53.75
W4	Cent. USA	Pac exit - downstream	106.25W-83.75W	31.25-48.75
W5	West. USA	Pac exit	126W-106.25W	31.25-48.75
W6	Canada	Pac exit North branch	141.25W-83.75W	51.25-68.75
W7	Alaska	Pac. North edge	166.25W-141.25W	28.75-51.25
W8	Siberia	–	88.75E-131.25E	51.25-68.75
W9	Japan	Pac entrance	126.25E-151.25E	28.75-51.25

Table 1: Geographic characteristics of the averaging areas used in this study (shown in Fig. 1).

Area	SONDE	RSAMP
W1	*3.04*	*4.10*
W2	1.73	*3.57*
W3	1.31	*2.17*
W5	1.27	*2.52*
W6	*3.03*	*3.82*
W9	(0.66)	*2.58*

Table 2: The linear trend ($\frac{m^2 s^{-2}}{year}$) of the area mean time series of V_{1df} from 1958-1999, for SONDE and RSAMP data. Trends with a significance level > 0.95 and < 0.80 are marked by *'s and parentheses, respectively.

Areas used	SONDE	RSAMP	full reanalysis
NH	16 (0.022)	25 (0.034)	29 (0.051)
ATL	25 (0.024)	36 (0.032)	35 (0.041)
PAC	27 (0.007)	36 (0.022)	43 (0.043)

Table 3: The percent of variance explained by the first EOF calculated using SONDE, RSAMP, and the full reanalysis data sets, for NH, ATL and PAC, as well as the linear trend of the corresponding PC time series (in parentheses, in units of standard deviation per year).The significance levels of the trends are very high (all are >0.98, with most exceeding 0.99) except for SONDE PAC, for which the significance level is 0.53.

	NH-S	NH-R_S	AT-S	AT-R_S	PA-S	PA-R_S
NH-S	1	-	-	-	-	-
NH-R_S	0.94	1	-	-	-	-
AT-S	0.90	-	1	-	-	-
AT-R_S	-	0.95	0.96	1	-	-
PA-S	0.65	-	(0.27)	(0.25)	1	-
PA-R_S	-	0.65	0.36	0.39	0.95	1

Table 4: The correlations between various PC time series, for the full data set (NH), and the Atlantic (AT) and the Pacific (PC) storm track regions, for SONDE (S), and RSAMP (R_S). Parentheses denote values that are not statistically significant, based on a 95% confidence level, but all other values are above the 99% percent confidence level.

Data set	ATL-PAC correlation	% squared covariance	% variance ATL/PAC	linear trend ATL/PAC ($\frac{\sigma}{yr}$)	corr. w/1st PC ATL/PAC
SONDE	0.61	44.6	17.2/23.0	0.029/0.010	0.95/0.73
RSAMP	0.58	66.5	29.2/32.4	0.051/0.028	0.98/0.70

Table 5: Statistics of the SVD analysis of the cross covariance matrix of the Pacific an Atlantic storm track regions. The significance levels of the trends exceed 0.989 for RSAMP and are 0.97/0.62 for SONDE ATL/PAC.

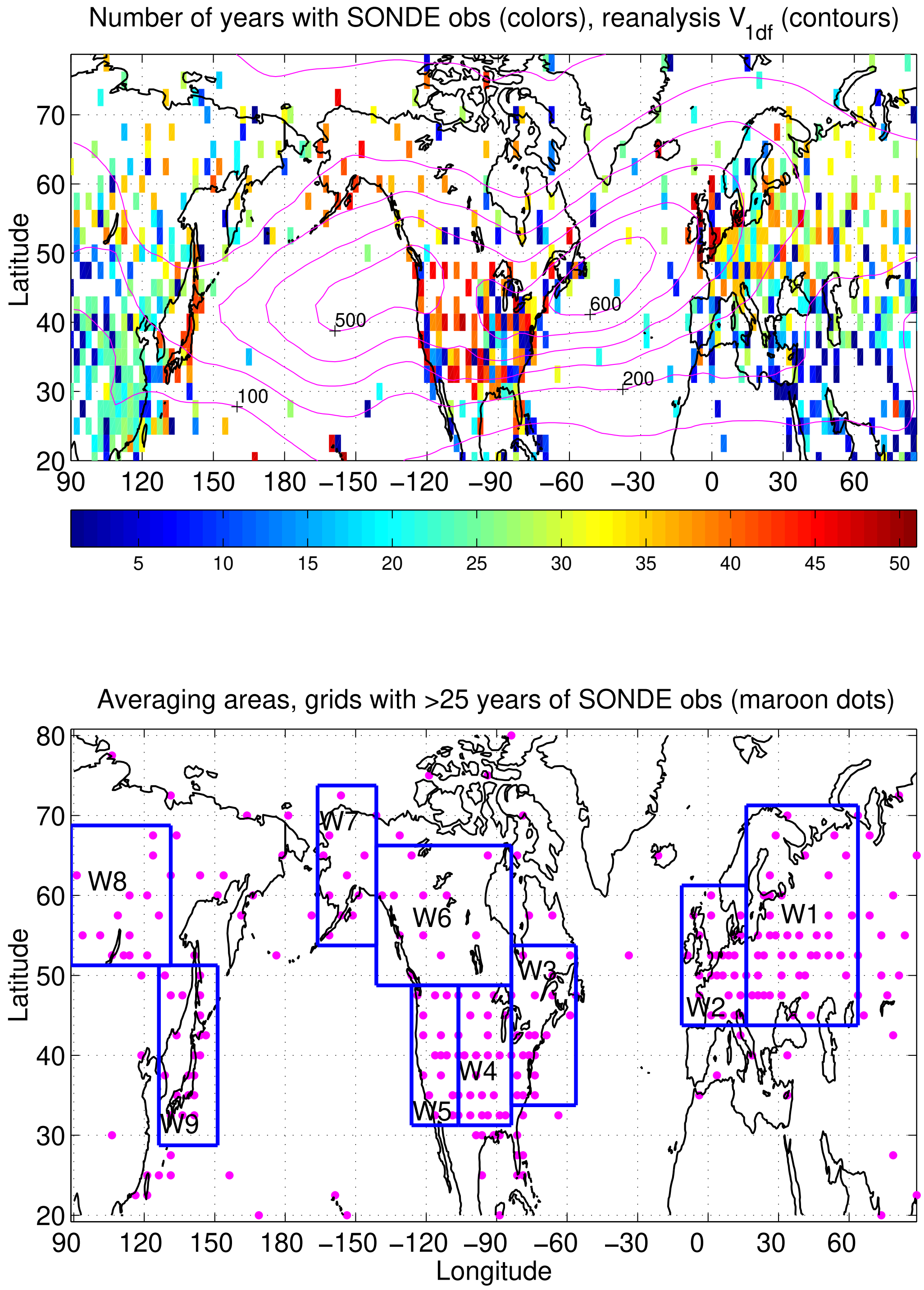

Figure 1: Top: The number of years with observations (colored rectangles) plotted at the observation location, and the reanalysis 1949-99 mean V_{1df} (contours). Bottom: The averaging areas used in this study (see text for details), and grids with more than 25 years of SONDE observations (maroon dots).

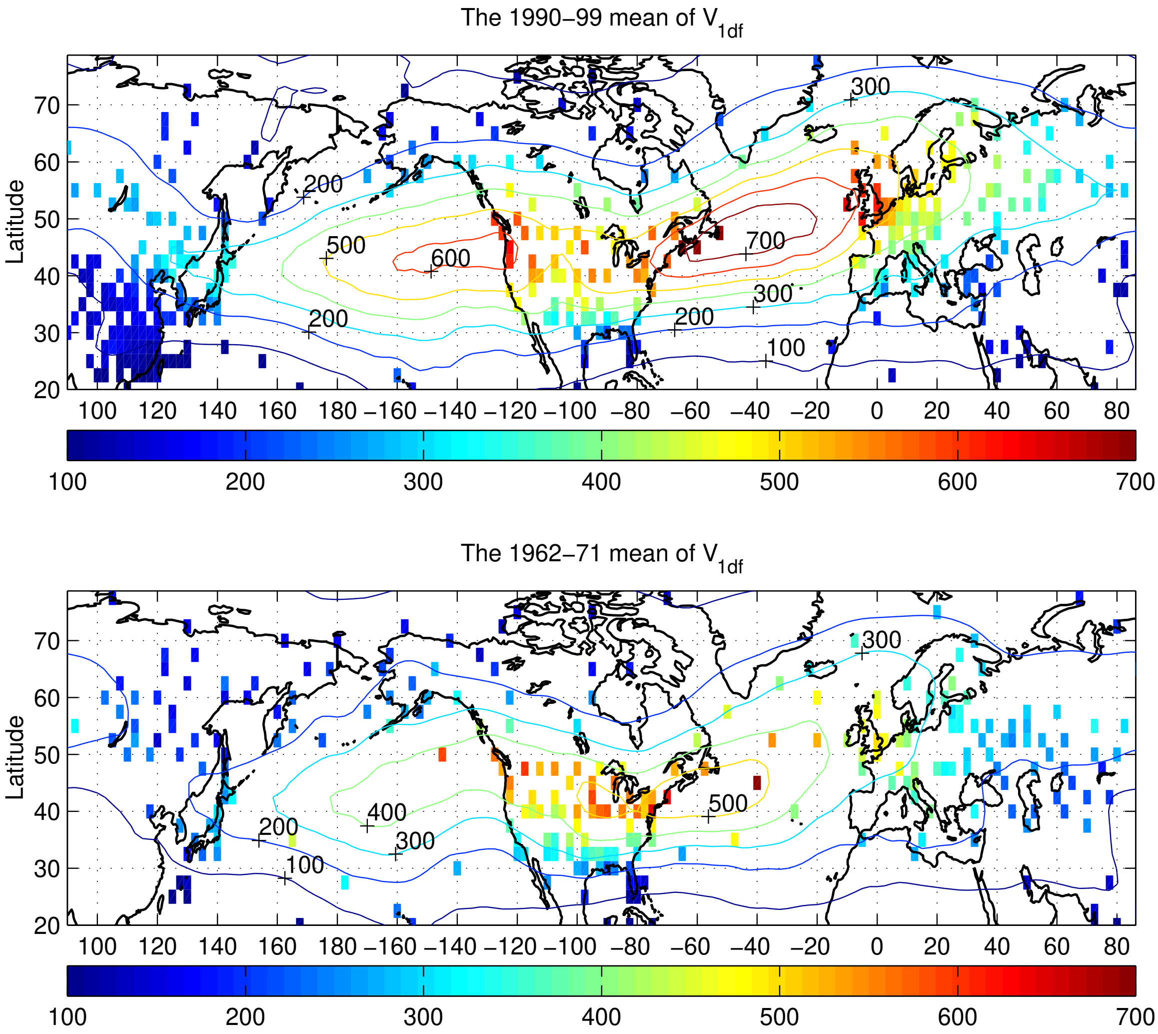

Figure 2: The decadal means of V_{1df} for the high (1990-99) and low (1962-71) decades, for SONDE (colored rectangles) and the full reanalysis data (contours). SONDE data shown only for grids that have at least 25 years of observations. To facilitate comparison, the same color scheme is used for the reanalysis contours and SONDE rectangles.

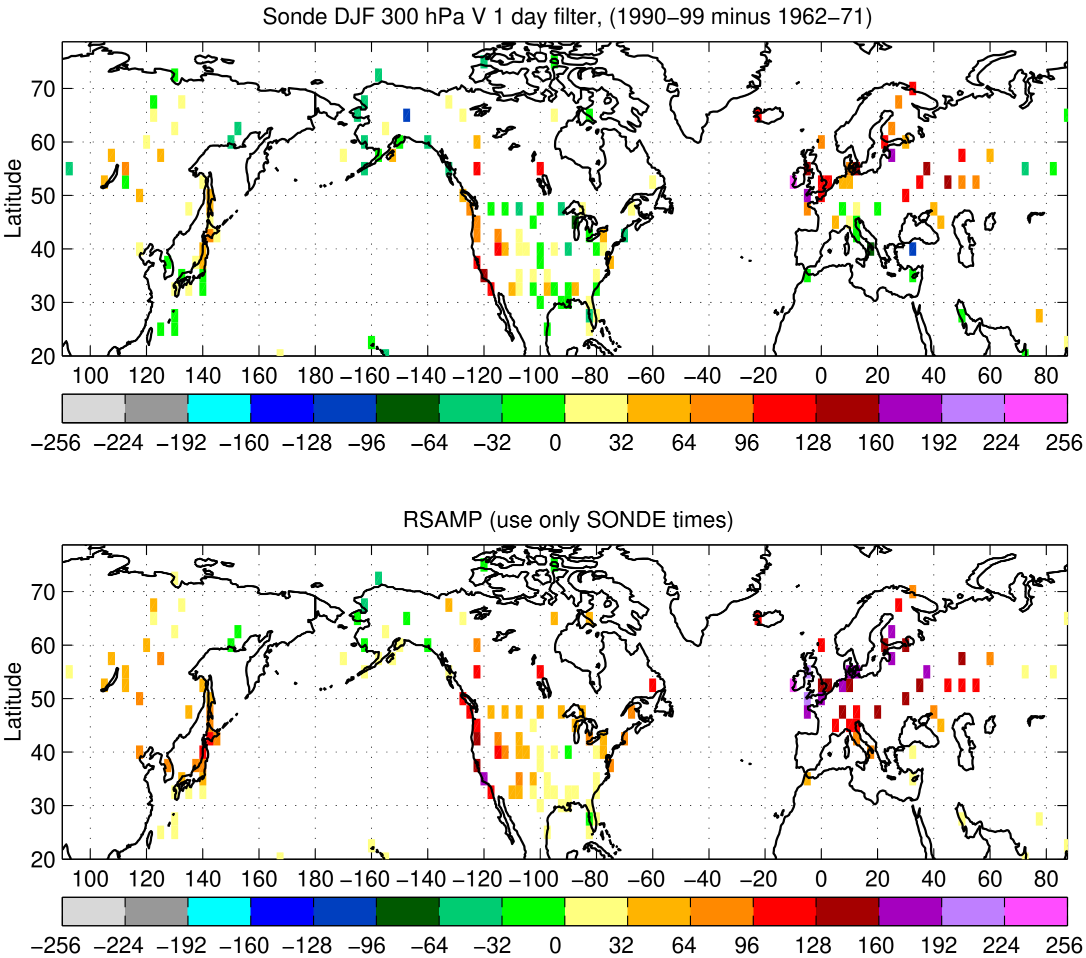

Figure 3: The difference between V_{1df} of the high and low decades of Fig. 2, for SONDE (top), and RSAMP (bottom). Data is shown only for grids with at least 7 years of observations in each of the decades.

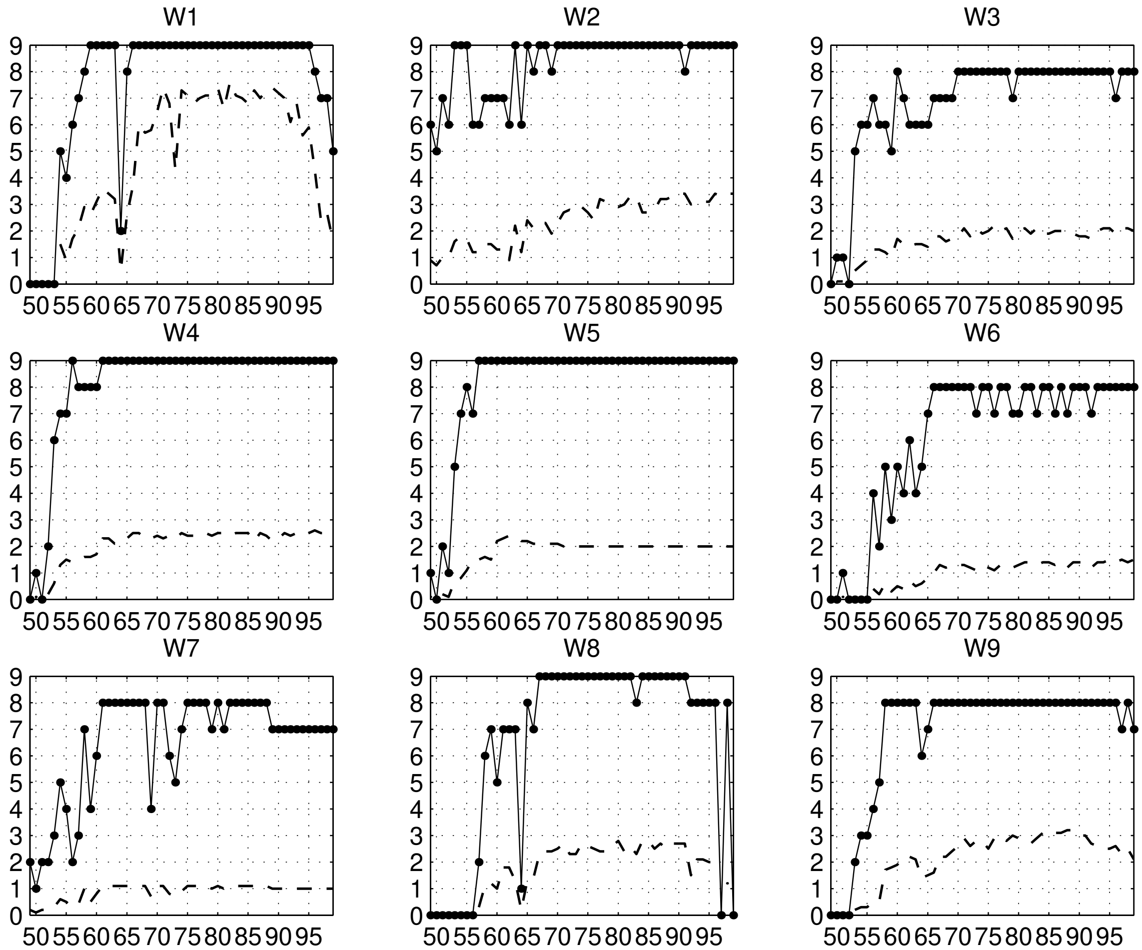

Figure 4: The yearly time series of the total number of stations with sufficient observations, in multiples of 10 (dashed), and the area coverage index for the areas shown in Fig. 1 (solid dots). See text for definition of the index.

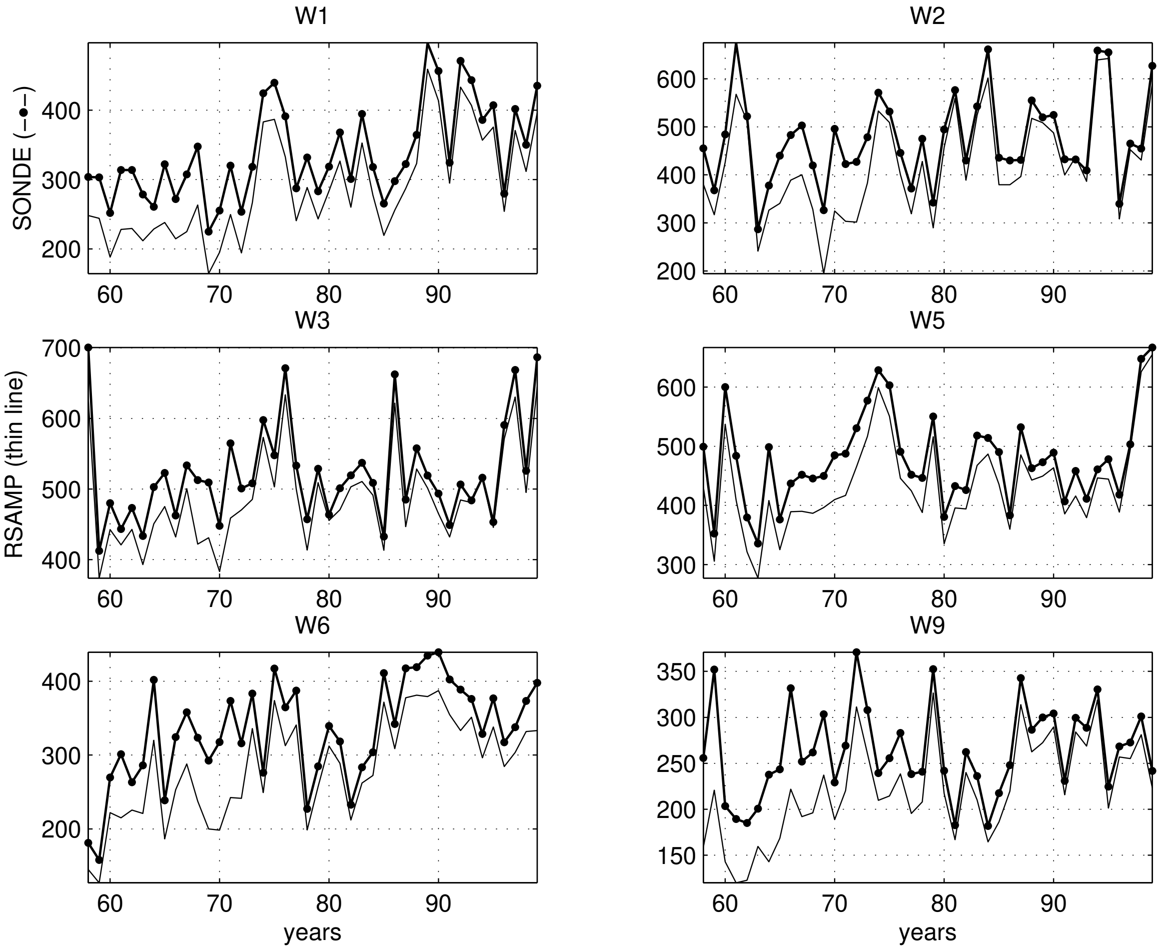

Figure 5: Yearly time series of the area averages of 300hPa V_{1df}, for areas W1, W2, W3, W5, W6, and W9 of Fig. 1. Shown are SONDE (dotted line) and RSAMP (thin line).

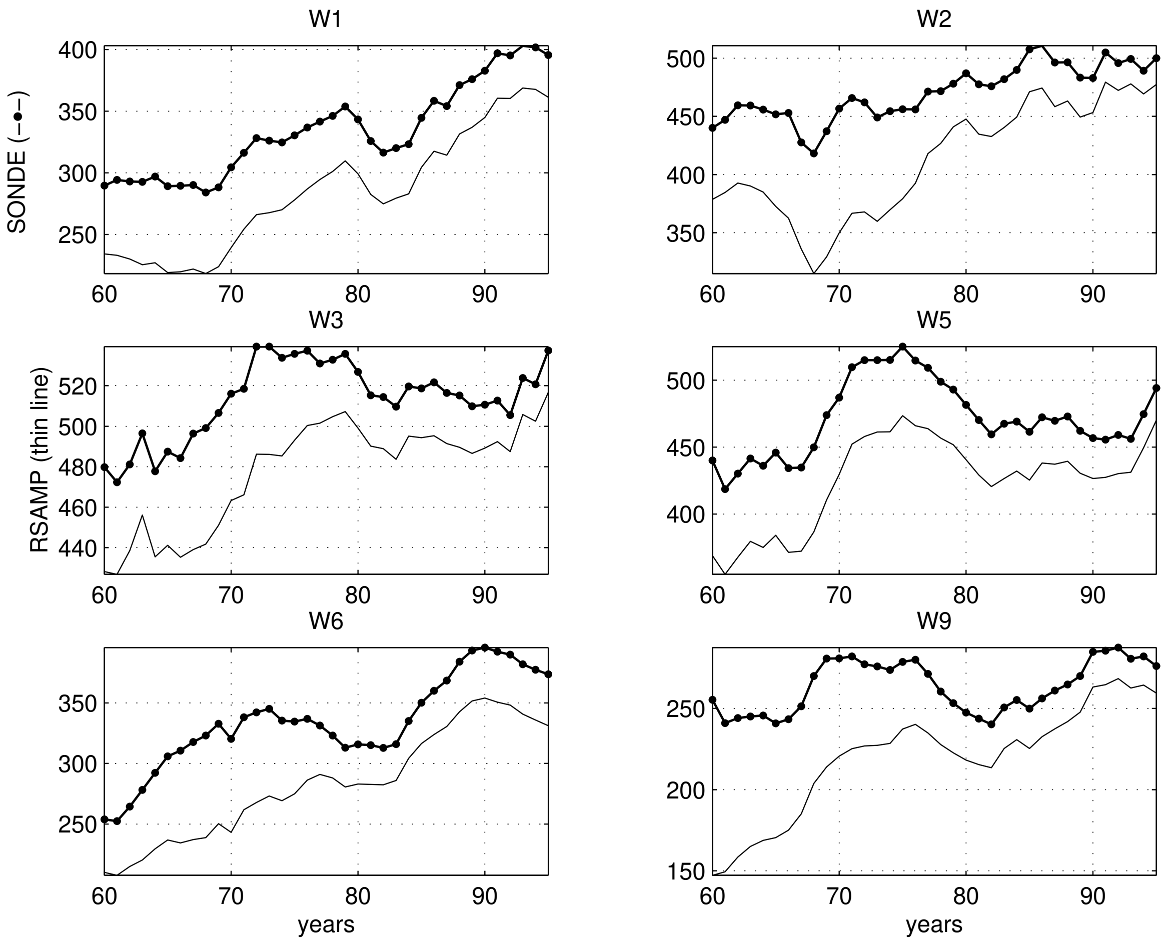

Figure 6: The 10 year running means of the time series in Fig. 5 (using the same line types).

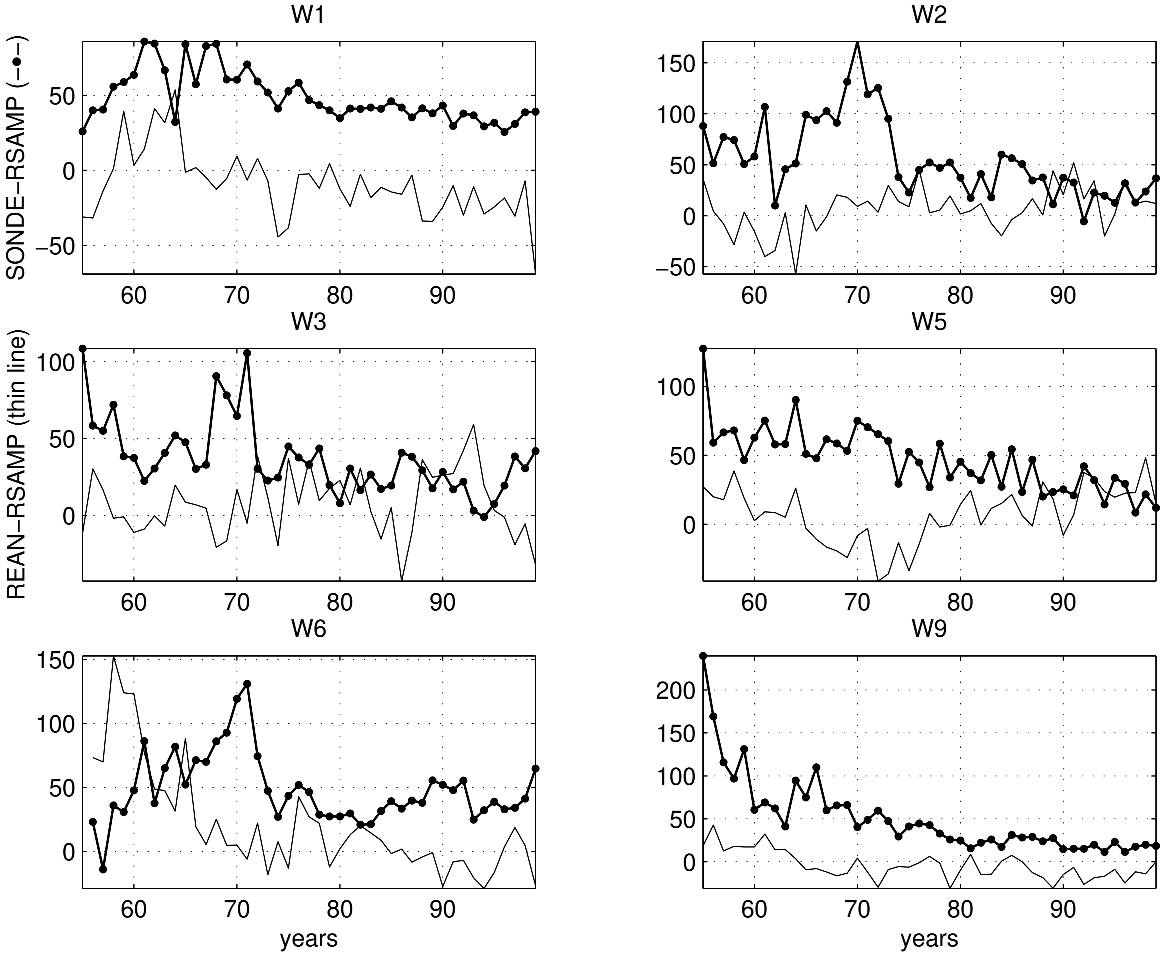

Figure 7: SONDE minus RSAMP yearly time series (the difference between the curves in Fig. 5, dotted line) and REAN minus RSAMP (thin line).

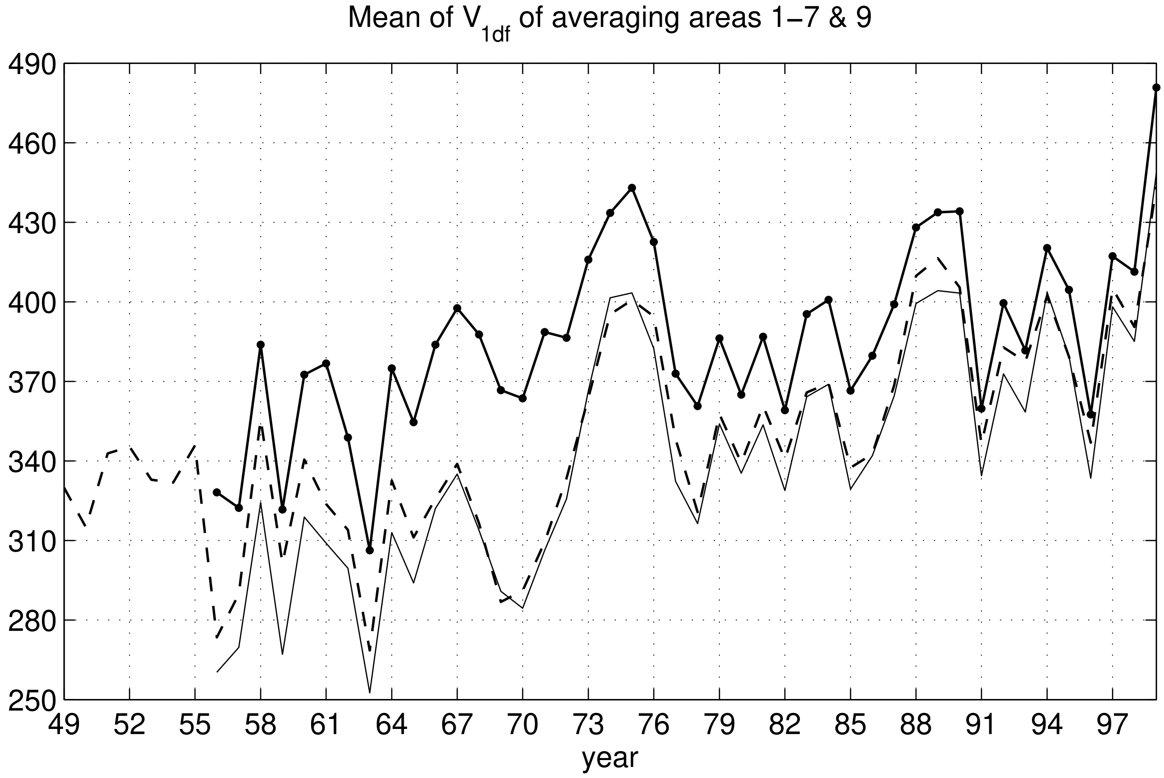

Figure 8: The average of all area mean V_{1df} time series except for W8, for SONDE and RSAMP as well as for the corresponding full reanalysis curve (REAN). SONDE (dotted line), RSAMP (thin line), and REAN (dashed).

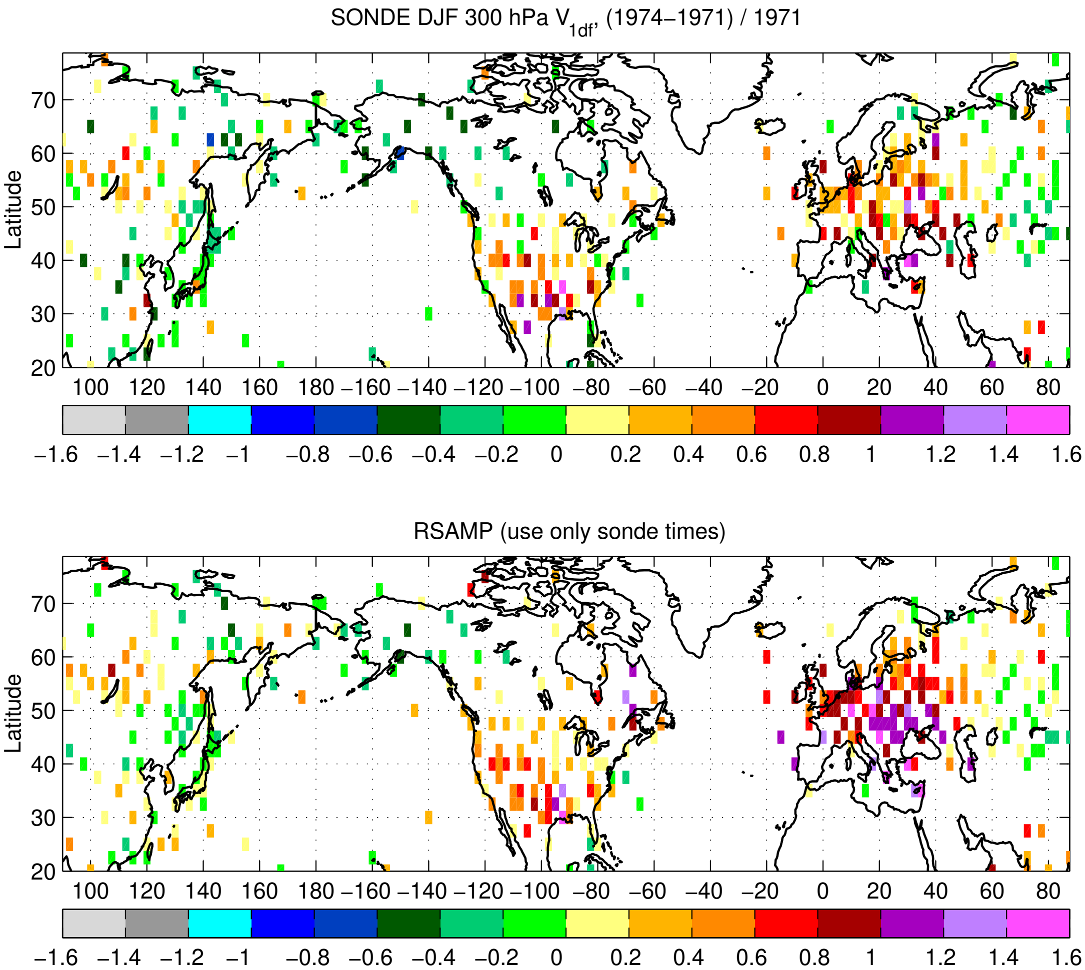

Figure 9: The relative change in V_{1df} during the early 70's - (1974-1971)/1971, for SONDE (top), and RSAMP (bottom).

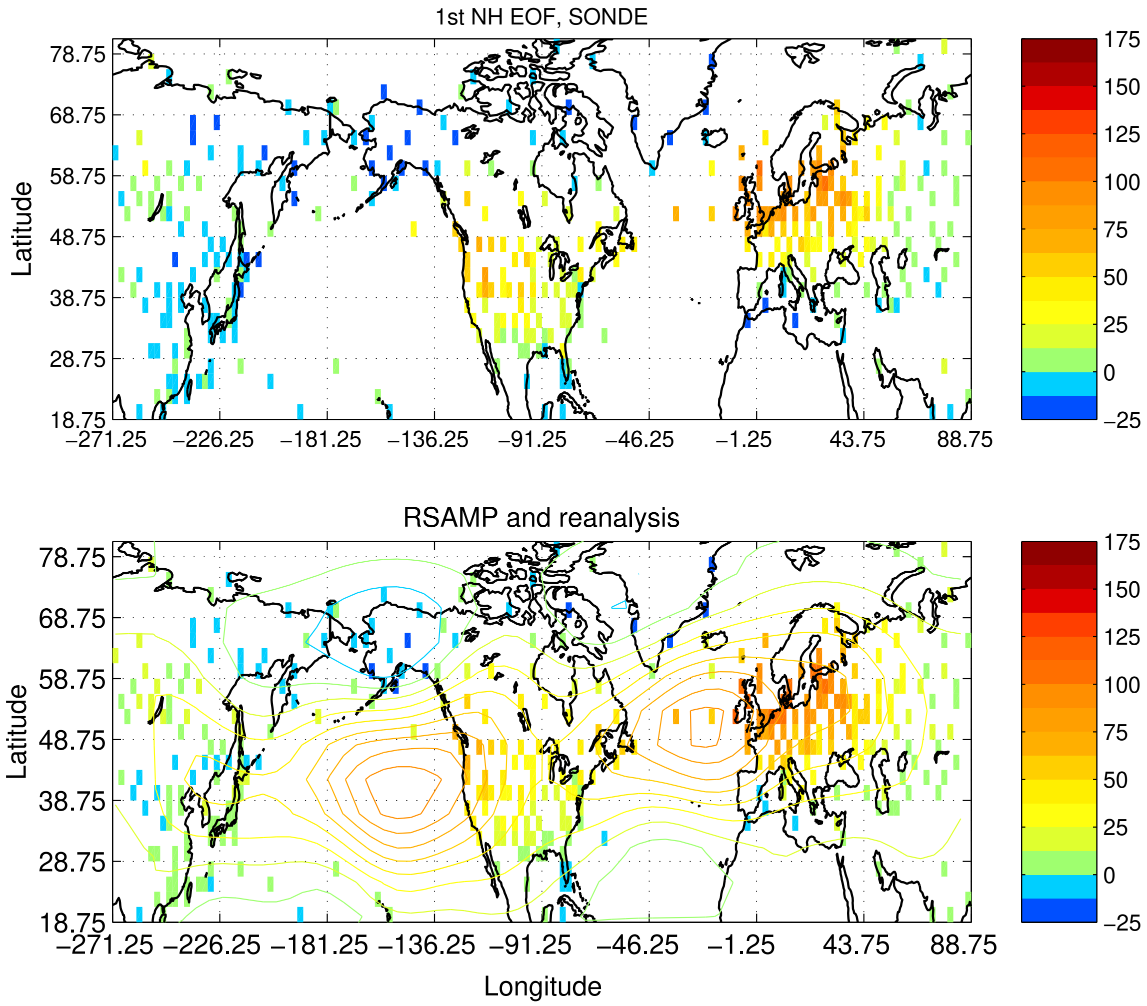

Figure 10: The first EOF of SONDE (top) and RSAMP (bottom) data, only using grid points with more than 25 years of observations (colored rectangles). Also shown in the bottom plot is the first EOF of the full reanalysis data (contours). The EOFs are normalized to represent one standard deviation. To facilitate comparison, the same color scheme is used for SONDE and RSAMP and the reanalysis.

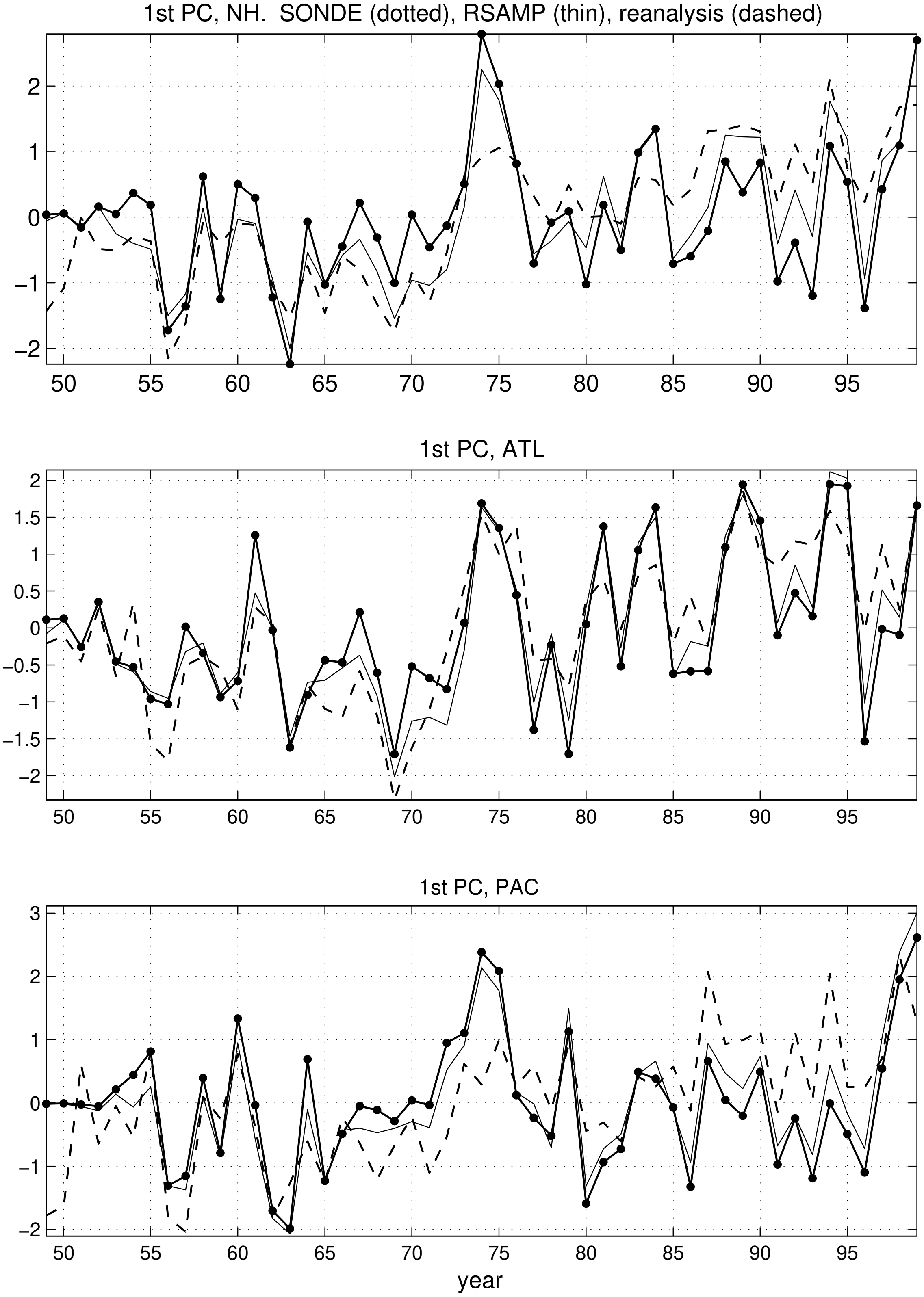

Figure 11: The first PC time series corresponding to the EOFs shown in figures 10 (top) and 12 (Atlantic in the middle and Pacific at the bottom). SONDE (dotted line), RSAMP (thin line), and the full reanalysi (dashed). PCs are in units of standard deviation.

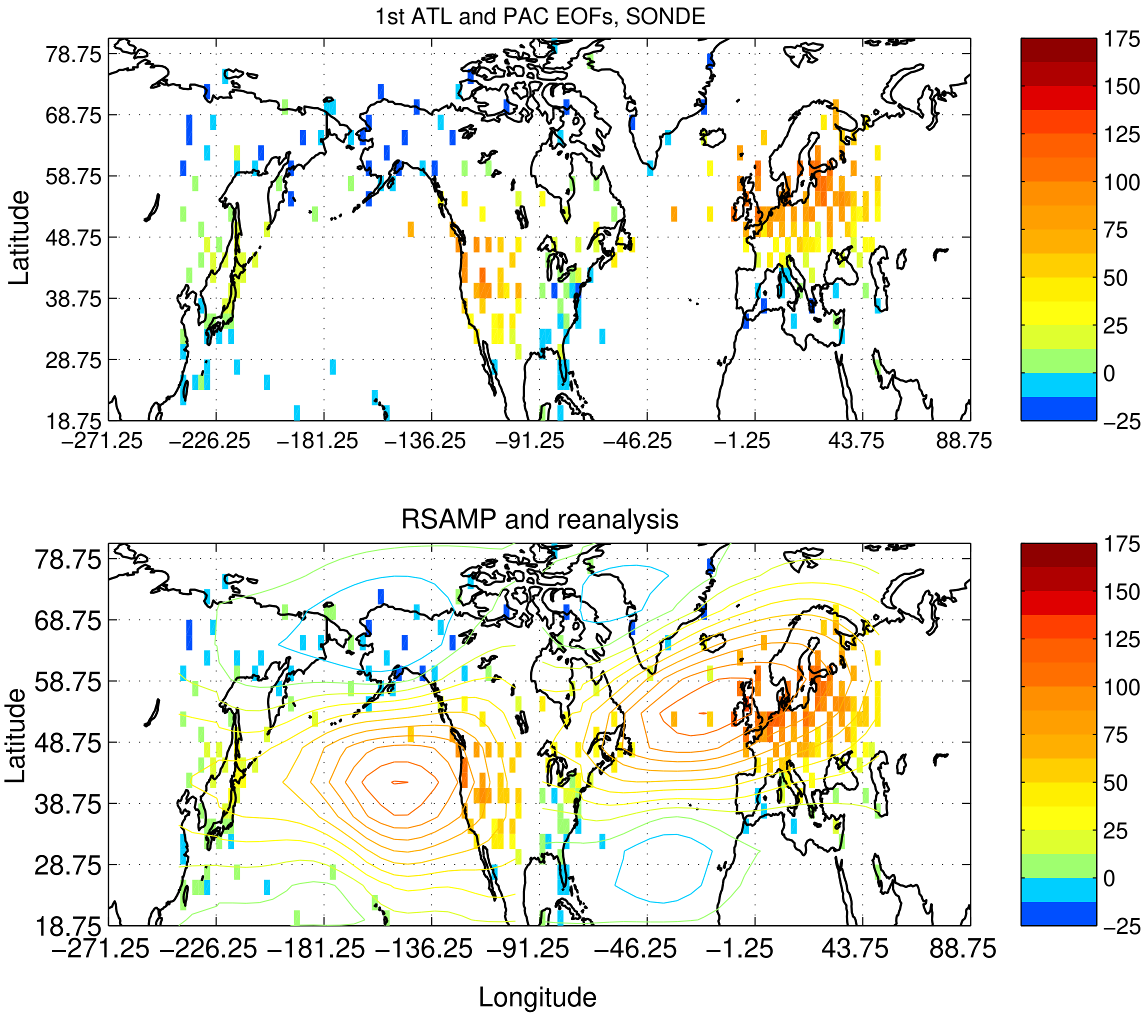

Figure 12: The same as in Fig. 10, but for the first EOFs calculated separately for the Atlantic and Pacific storm track regions.

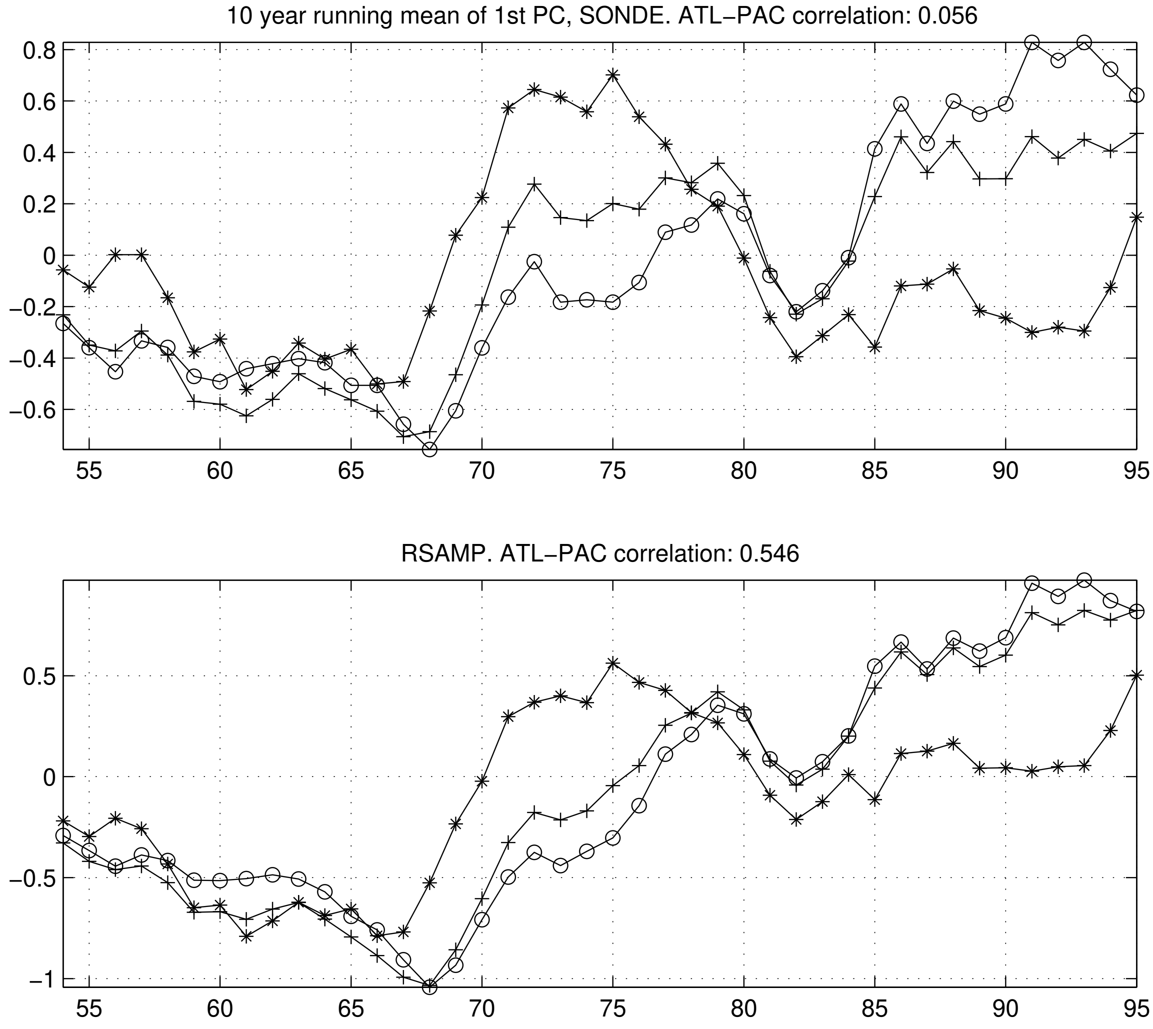

Figure 13: The 10 year running means of the PC time series of Fig. 11, for SONDE (top) and RSAMP (bottom). Curves are for NH (+), ATL (o), and PAC (∗).