Intercomparison of Global Marine Climatologies

Stanley L. Grotch
Lawrence Livermore National Laboratory
Livermore, CA

Introduction

One of the major goals of the Program for Climate Model Diagnosis and Intercomparison at LLNL is the development and application of statistical/graphical methods for the intercomparison of GCMs among themselves and with observational data. Consequently, it is important to have available datasets that represent "current climate" and a range of tools with which to perform detailed intercomparisons. This work reports on an intercomparison of four historical climatologies for surface winds (focusing on zonal u-wind) and illustrates a variety of statistical/graphical tools that can be used to quantitatively determine the differences/similarities among these datasets.

The four climatologies used here and their reference periods are those of: COADS (1950-1979), ECMWF (1980-1989), NMC (1979-1988), and Oort (1963-1973). In addition to having different reference periods, the data also are gridded differently: COADS (2° boxes), ECMWF and NMC (2.5° x 2.5°), Oort (2.5° x 5°). To permit detailed quantitative spatial intercomparison, it is essential that the data be available on a common grid. Because the COADS data are *exclusively* ocean-based whereas the others are global (although substantially land-based in terms of observation) the non-COADS datasets were regridded using an inverse squared weighting procedure to a 2° x 2° grid corresponding to the centers of the COADS boxes. Only those locations where COADS data were available were considered (approximately 7650 gridpoints globally). Although other variables have been examined, the primary focus here will be on the eastward component of the surface wind, u, averaged seasonally or annually.

Some of the major questions considered are: which of these datasets are "closest" and which are "furthest" apart? Quantitatively, what are the magnitudes of the largest differences and what is their spatial distribution?

Large Scale Intercomparisons

The cross correlations between the seasonal (DJF or JJA) and annually averaged surface zonal wind fields of these data sets are extremely high (> 0-9) for all pairs of intercomparisons. For the ECMWF, NMC pairing the cross correlation is nearly perfect (0.98) in both seasons. In Fig. 1 the zonally averaged annual zonal winds are displayed for these four datasets. The agreement is generally good, except at low SH latitudes and in the NH mid-latitudes. Two pairings of the data that are generally "closest" to one another are evident: ECMWF, NMC and COADS, Oort.

The zonal correlation between a pair of datasets is the cross correlation for all longitudinal gridpoints at a fixed latitude. A plot of the zonal correlation of annually averaged zonal wind for the three datasets vs. COADS is shown in Fig. 2. For latitudes equatorward of 25°, the zonal correlations for all three datasets vs. COADS are generally high (> 0.9). For most of this range, the Oort dataset has the highest zonal correlation with the COADS data.

Distributional Statistics

Another useful intercomparison diagnostic is the area-weighted histogram. Rather than using the more conventional number frequency based histogram, here the ordinate represents the *fractional global area* containing a given range (abscissa) of the field being represented. In Fig. 3, four pairs of area-weighted histograms are displayed for the annually averaged surface meridional wind component (V). The histograms in the right two panels (Oort, COADS and NMC, ECMWF) are generally very similar, whereas the left two panels (ECMWF, COADS and NMC, COADS) show substantial differences. Although the area weighted global averages for v wind for ECMWF and COADS are quite close (0.20 and 0.33 m/sec) their more detailed histograms show obvious differences.

With the data available on a common grid, it is also possible to compare the distributions of pointwise differences between different pairings of the data. Figure 4 displays boxplots showing these point-wise differences for annually averaged zonal wind at all common COADS grid points (unweighted by area) for the differences: Dataset - COADS. In this representation, the shaded box contains the central 50% of the data, the middle horizontal line is the median, and the upper and lower groupings of small circles represent the 10% extrema. In examining the ranges of both the central and the 80% limits, it can be seen that the order of “closeness” to the COADS data-for these datasets is: Oort, ECMWF, and NMC, respectively. The global distribution of point differences between COADS and Oort is markedly smaller than that between COADS and either ECMWF or NMC.

The differences in velocity for two datasets, x and y, can also be represented as % differences [100(x-y)/((x+y)/2)]. To avoid possible singularities where x+y is zero, consider only grid points where the average u=(x+y)/2 of the datasets under comparison is >1 m/sec. Fig. 5 shows the percentage differences in annually averaged u-wind as boxplots for all pairwise comparisons of the four datasets. Here, only the 80% limits are shown: (e.g., the 10th to the 90th percentile of the grid points). From both the widths of the central boxes and the 80% limits it is evident that the COADS vs. Oort and the ECMWF vs. NMC percentage differences are generally the smallest, whereas the NMC vs. Oort and the COADS vs. NMC are the greatest. Generally positive biases of the three datasets vs. Oort are also evident.

The box plot technique can also be used to display side-by-side the detailed distribution of the measured fields as a function of latitude, as in Fig. 6 where 10 cuts at different 10° latitude sections are shown for the four datasets. The longitudinal distributions of annually averaged u-wind for the four datasets are displayed side-by-side in each subpanel. These representations serve as very useful diagnostics by readily showing commonalties as well as differences between these distributions. For example, at 45°S the Oort u-winds are quite different than those of NMC, there being only slight overlap in their magnitudes. However, although offset in value, the ranges and the central 50% limits at most other latitudes are very similar among these datasets.

Locations of Maximum Difference Points

In intercomparisons it is also important to know where the largest (and the smallest) pointwise differences arise. With generally available software, it is possible to examine these issues in a highly interactive manner. In the top panel of Fig. 7, the pointwise distribution of the differences in annually averaged u-wind are displayed as a histogram. By merely touching any desired bars on this distribution with a cursor, the corresponding spatial locations on a map “light up”, as shown in the lower panel. For the COADS and Oort intercomparison, many of the maximal differences are seen to lie along the continental margins.

In the same manner, by merely circling any desired area or range of latitudes on the maps, a corresponding sub-histogram darkens on the upper panel histogram. Alternatively, if other pairwise difference histograms are simultaneously displayed (see Fig. 8), they too show the same selected points. Because their sub-histograms (shaded) are more centrally located in terms of their respective differences with COADS, the points highlighted as the maximal Oort - COADS differences (left sub panel) are *not* the corresponding maximal differences for the other two datasets. This new and powerful interactive technology is likely to facilitate the rapid exploratory analysis of meteorological data fields.

Conclusions

There exist a variety of statistical/graphical displays based on nonparametric statistics which should be applied to problems in the intercomparison of meteorological fields. These techniques can serve as very important adjuncts to more conventional side-by-side contour plots. Recent advances in personal computer hardware/ software have resulted in novel, highly interactive. and extremely powerful statistical/ graphical tools which should greatly facilitate intercomparisons.

Comparisons of the seasonally and annually averaged surface u-winds from four historical datasets yield generally good global agreement. Using a variety of nonparametric representations, the four datasets examined divide up into two pairings: COADS, Oort and ECMWF, NMC. Perhaps, in retrospect, this is not surprising in that all of these datasets use many of the same observational data and the ECMWF and NMC analyses employ many similar data assimilation and modeling techniques.

Acknowledgement

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract No. W-7405-eng-48.

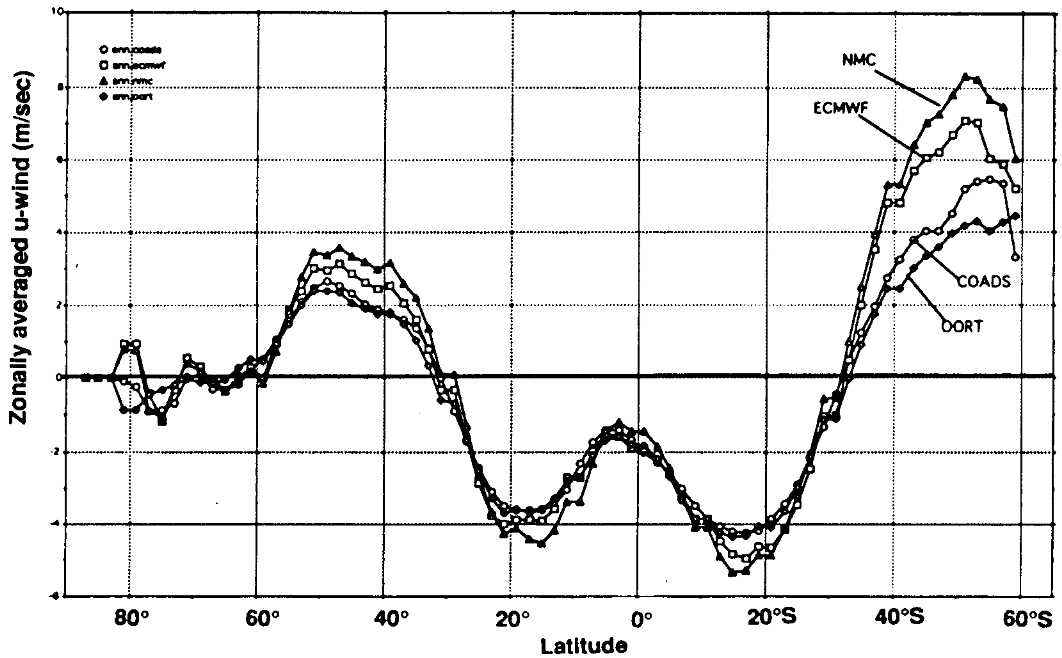

Figure 1. Zonally averaged annual zonal surface winds for four different climatologies: COADS, ECMWF, NMC, Oort. The general agreement between these datasets is good but two Pairings of data are evident: COADS, Oort and ECMWF, NMC.

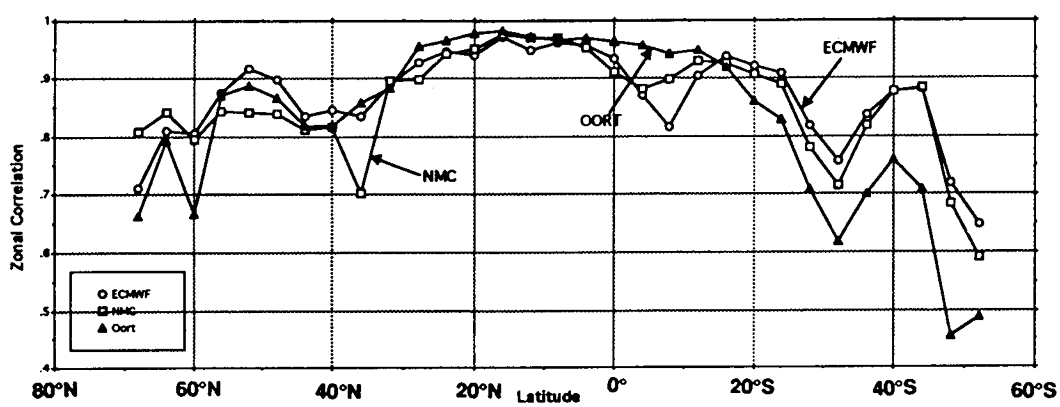

Figure 2. Zonal cross correlations of ECMWF NMC, Oort vs. COADS for annually averaged zonal surface wind (u). At each latitude, the cross correlation between the fields of all grid points longitudinally are plotted as a function of latitude.

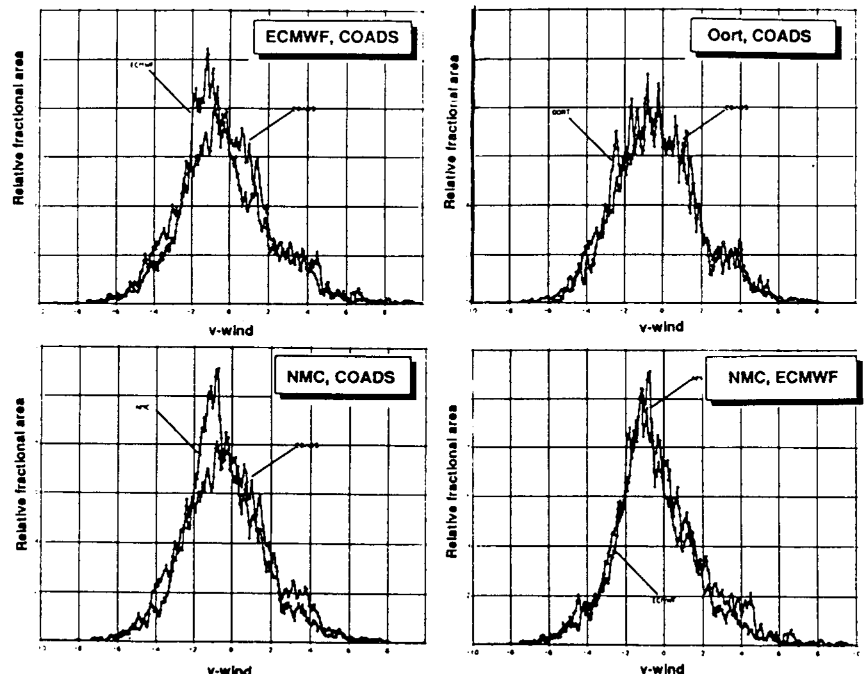

Figure 3. Area weighted histograms of the annually averaged meridional surface wind (v) for different datasets. For each histogram, the vertical scale represent the fractional global area containing a given range of v-wind (horizontal axis). The two left panels clearly show differences among the datasets whereas the two right panels indicate good agreement.

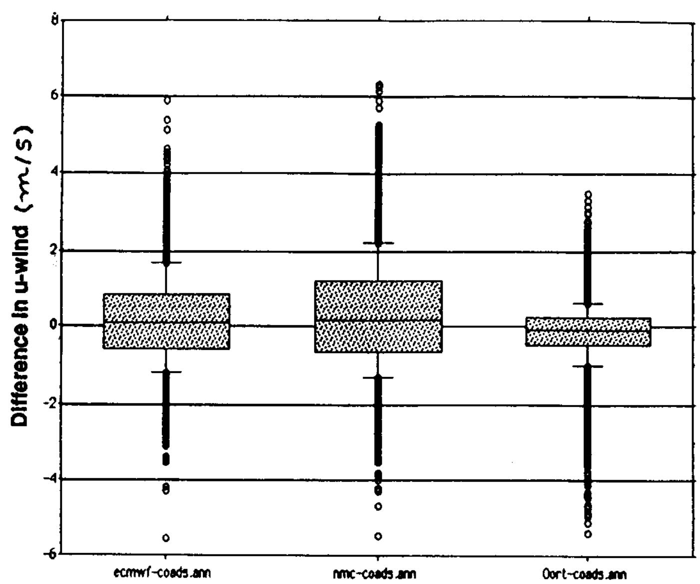

Figure 4. Boxplots showing the distribution of pointwise differences: dataset - COADS for three datasets for annually averaged zonal surface wind (u) at all COADS grid points. In these icons the shaded box contains the central 50% of the grid points (75th percentile to 25th percentile) and the 10% tails at each end of the distribution are represented as small circles.

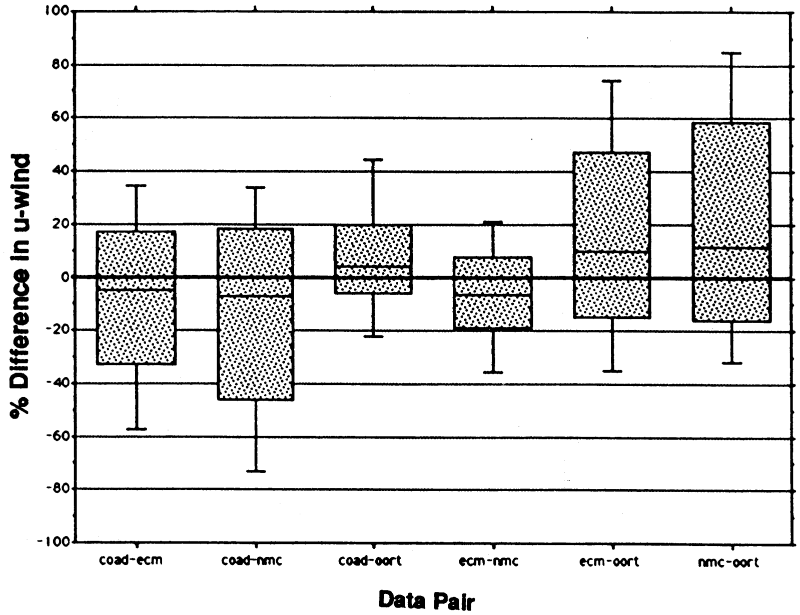

Figure 5. Distributions of all pairwise comparisons of the percentage difference in annually averaged zonal surface wind (u) are displayed as boxplots. Here, the extrema in each icon indicate the 10th and 90th percentiles.

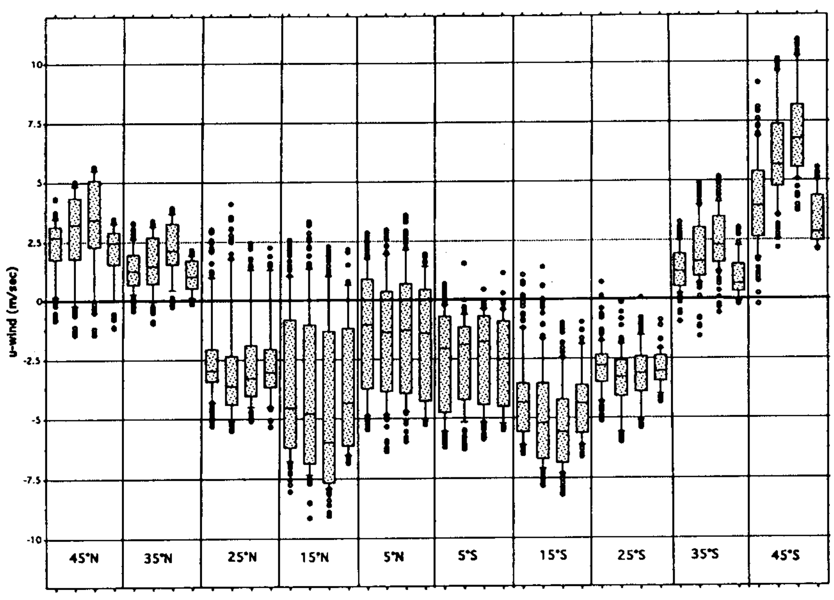

Figure 6. Side-by-side distributions of annually averaged zonal surface wind (u) for four different datasets at 10 latitudinal cuts. In each subpanel the left to right order is COADS, ECMWF, NMC, Oort. The small circles represent those grid points in the extreme 10% tiles.

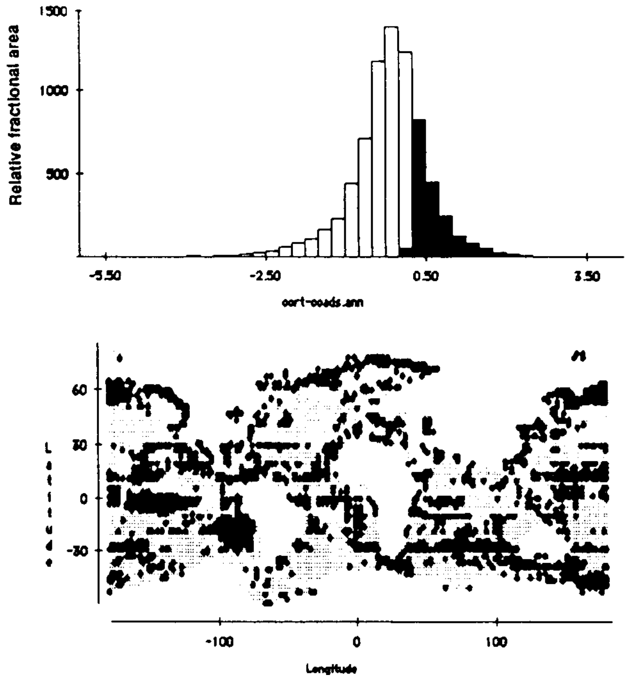

Figure 7. Top panel is the number of histogram of the distribution of differences: COADS - Oort for annually averaged zonal surface wind (u). The lower panel interactively shows the actual spatial locations of these maximal highlighted differences.

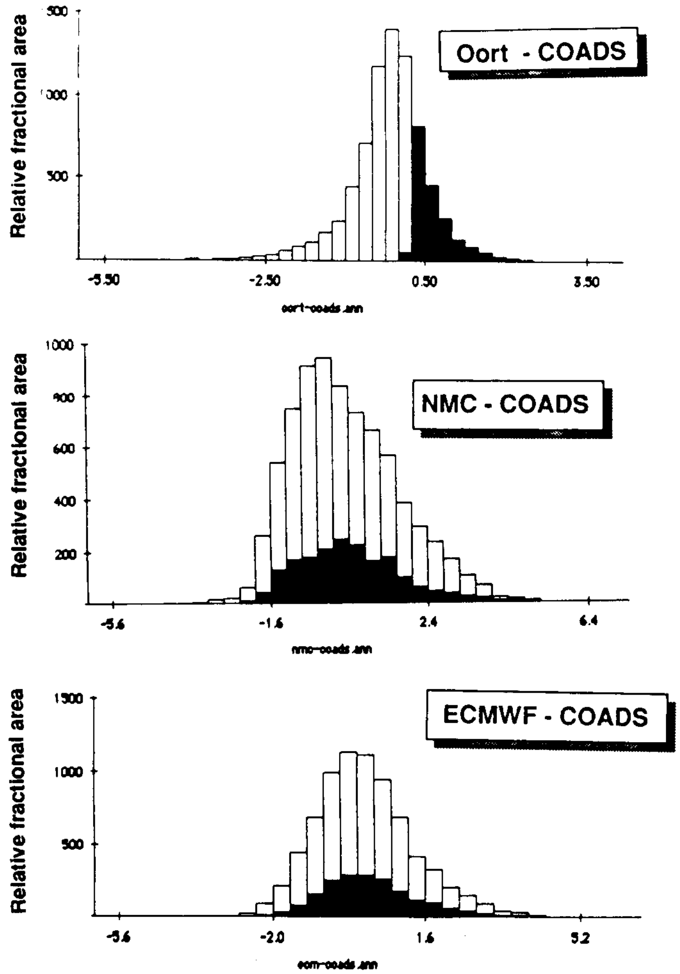

Figure 8. Histograms of the pointwise differences in annually averaged u-wind for: Oort - COADS, NMC - COADS, ECMWF - COADS. The maximal difference points selected in the shaded Oort - COADS distribution are <u>not</u> the same maximal points for the other pairwise differences.