An Objective Classification of Climatic Regions in the Pacific and Indian Oceans

Henry F. Diaz and Timothy J. Brown

Abstract: We have applied a number of objective statistical techniques to define homogeneous climatic regions for the Pacific Ocean, using COADS (Woodruff *et al* 1987) monthly sea surface temperature (SST) for 1950-1989 as the key variable. The basic data comprised all global 4°x4° latitude/longitude boxes with enough data available to yield reliable long-term means of monthly mean SST. An R-mode principal components analysis of these data, following a technique first used by Stidd (1967), yields information about harmonics of the annual cycle of SST. We used the spatial coefficients (one for each 4-degree box and eigenvector) as input to a K-means cluster analysis to classify the gridbox SST data into 34 global regions, in which 20 comprise the Pacific and Indian oceans. Seasonal time series were then produced for each of these regions. For comparison purposes, the variance spectrum of each regional anomaly time series was calculated. Most of the significant spectral peaks occur near the biennial (2.1-2.2 years) and ENSO (~3-6 years) time scales in the tropical regions. Decadal scale fluctuations are important in the mid-latitude ocean regions.

We present the results of our studies in applying a number of objective statistical techniques to define homogeneous climatic regions over the world oceans (see Diaz and Brown 1992). The variable used is COADS monthly sea surface temperature for 1950 through 1989. Reasons for exploring the nature of "climatically homogeneous" oceanic regions are twofold. First, from the point of view of climate monitoring and climate change detection efforts, it is useful to divide the ocean into smaller units. Second, it might be advantageous to study a number of air/sea interaction processes on regional scales. It is also of interest to compare SST changes in the different ocean basins. We will focus here only on the Indo-Pacific Ocean region.

Methodology

Using 4-degree latitude/longitude area boxes from the COADS data set (Woodruff *et al* 1987), an R-mode principal component analysis (PCA) was performed using the calendar month 1950-1989 long-term mean SSTs in each box as the variables, with a minimum requirement of five years of data within the 40-year period in each box. The use of the R-mode PCA technique in climate classification was first published by Stidd (1967) to classify the precipitation climate of Nevada, and also by Skaggs (1975) to study the 1930s drought in the United States.

K-means cluster analysis (Hartigan 1975) was then used to classify the gridbox SST data, using as input the spatial coefficients from the first two

---

In: K.T. Redmond and V.L. Tharp, Editors. 1993. Proceedings of the Ninth Annual Pacific Climate (PACLIM) Workshop, April 21-24, 1992. California Department of Water Resources, Interagency Ecological Studies Program, Technical Report 34.

eigenvectors of the monthly mean SST. Initially, a total of 1,939 4-degree boxes for the world oceans were used as input to the clustering algorithm, with an initial prescribed seed of 15 groups. Boxes in certain inland seas, and areas with relatively sparse data coverage were manually edited out of the final regional configuration, eliminating 123 boxes, for a final total of 1,816 4-degree boxes yielding 34 regional clusters over the world oceans (20 regions in the Indo-Pacific sector). In general, distinct regions were derived for each hemisphere and ocean basin. The resulting cluster patterns were visually examined to determine if the physical boundaries were climatologically consistent, and various tests were performed to ascertain the temporal coherence of the constituent boxes for each region. This involved, for instance, comparing the distribution of the departure values of the individual 4-degree boxes and calculating the mean inter-box correlation within each region.

Analysis Results

Figure 1 shows a map of the 20 cluster regions that were classified using the procedure described above, and Table 1 presents a few basic summary statistics for each region. In general, as expected, the largest regions are in the tropics. Seasonal values for the 4-degree boxes within each region were averaged together (with cosine of latitude weighting to account for the differences in area) to form regional time series. Figures 2-4 illustrate these results for three ocean areas: two in the tropics and one in mid-latitudes.

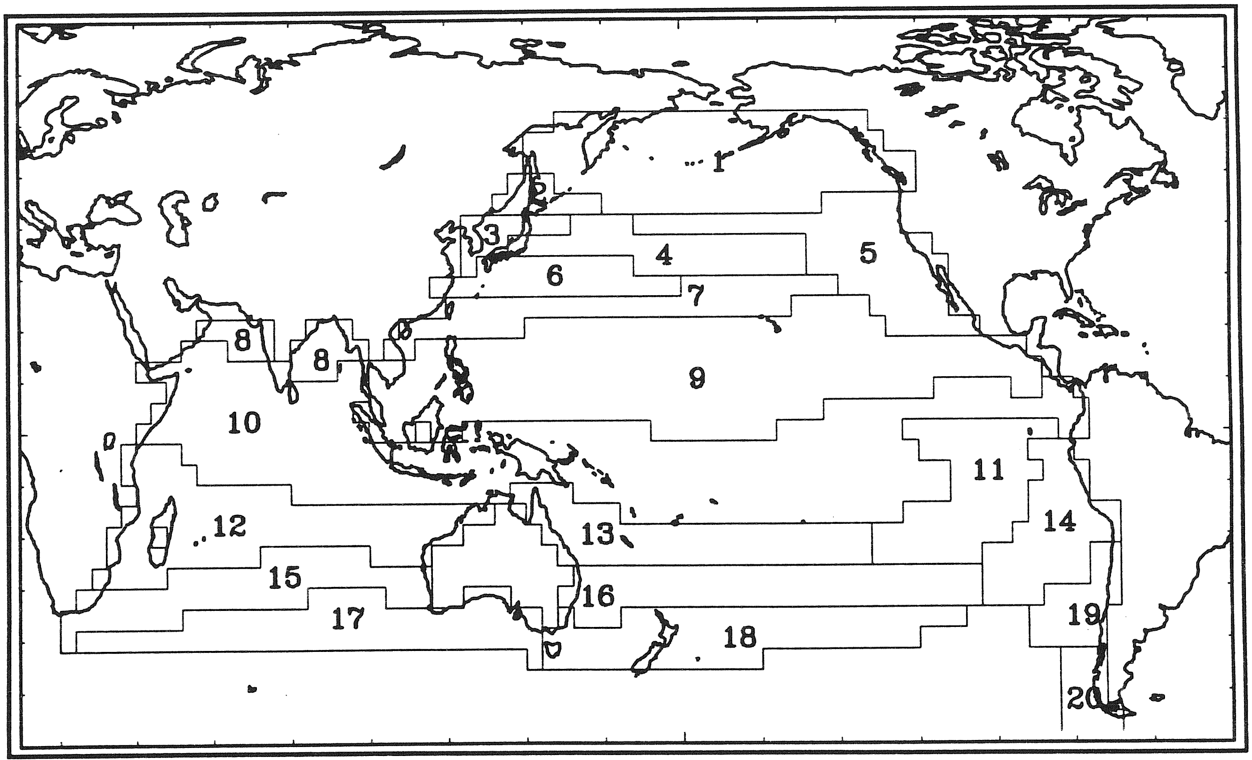

Figure 1. Map illustrating the regional SST boundaries.

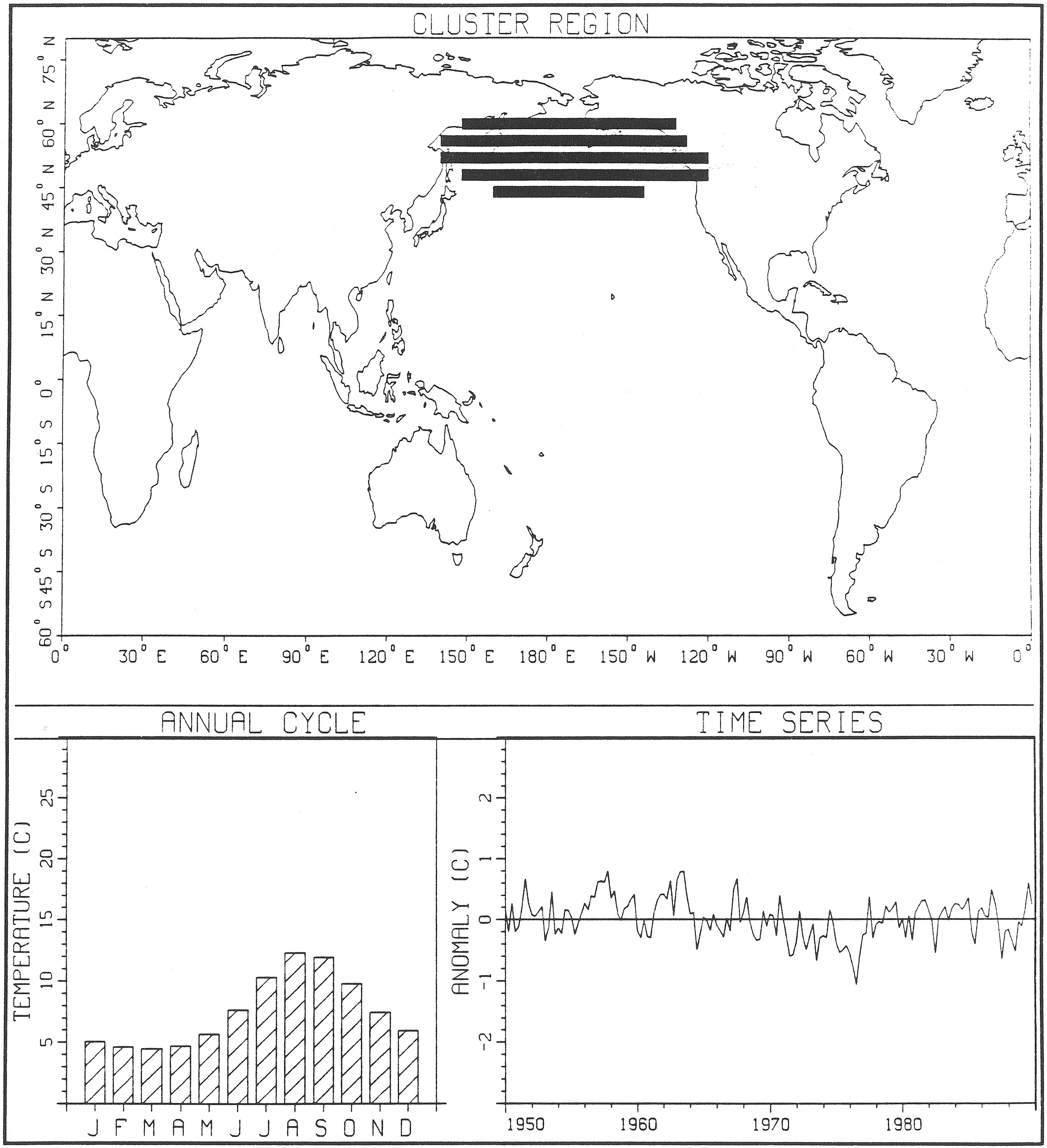

Figure 2. Time series of SST anomalies for the northern North Pacific region (region 1 in Figure 1) indicated by the shaded bars. Figure also shows the seasonal cycle of monthly mean SSTs averaged over the entire region.

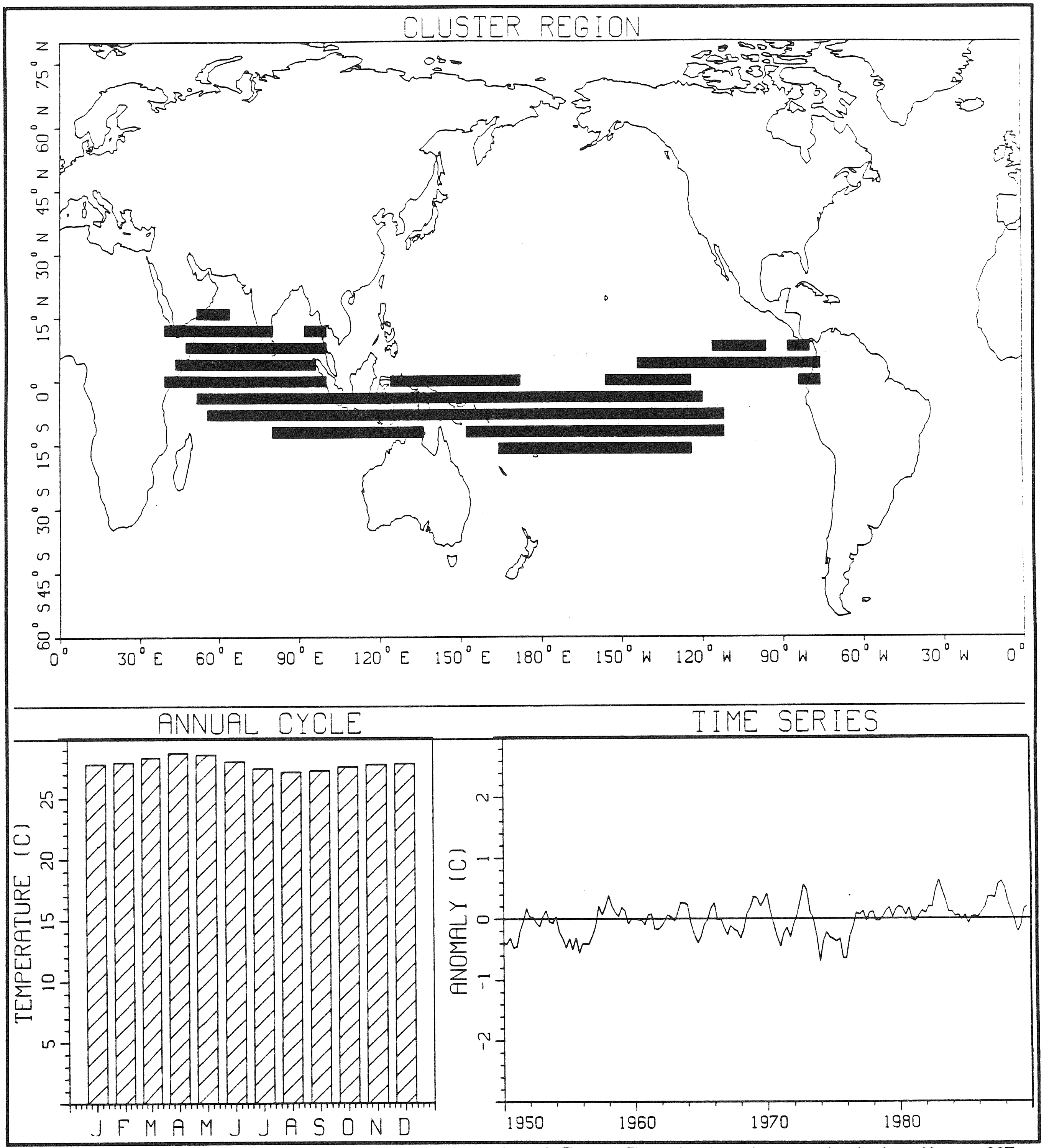

Figure 3. Time series of SST anomalies for the tropical warm pool region (region 10 in Figure 1). Figure also shows the seasonal cycle of monthly mean SSTs averaged over the entire region.

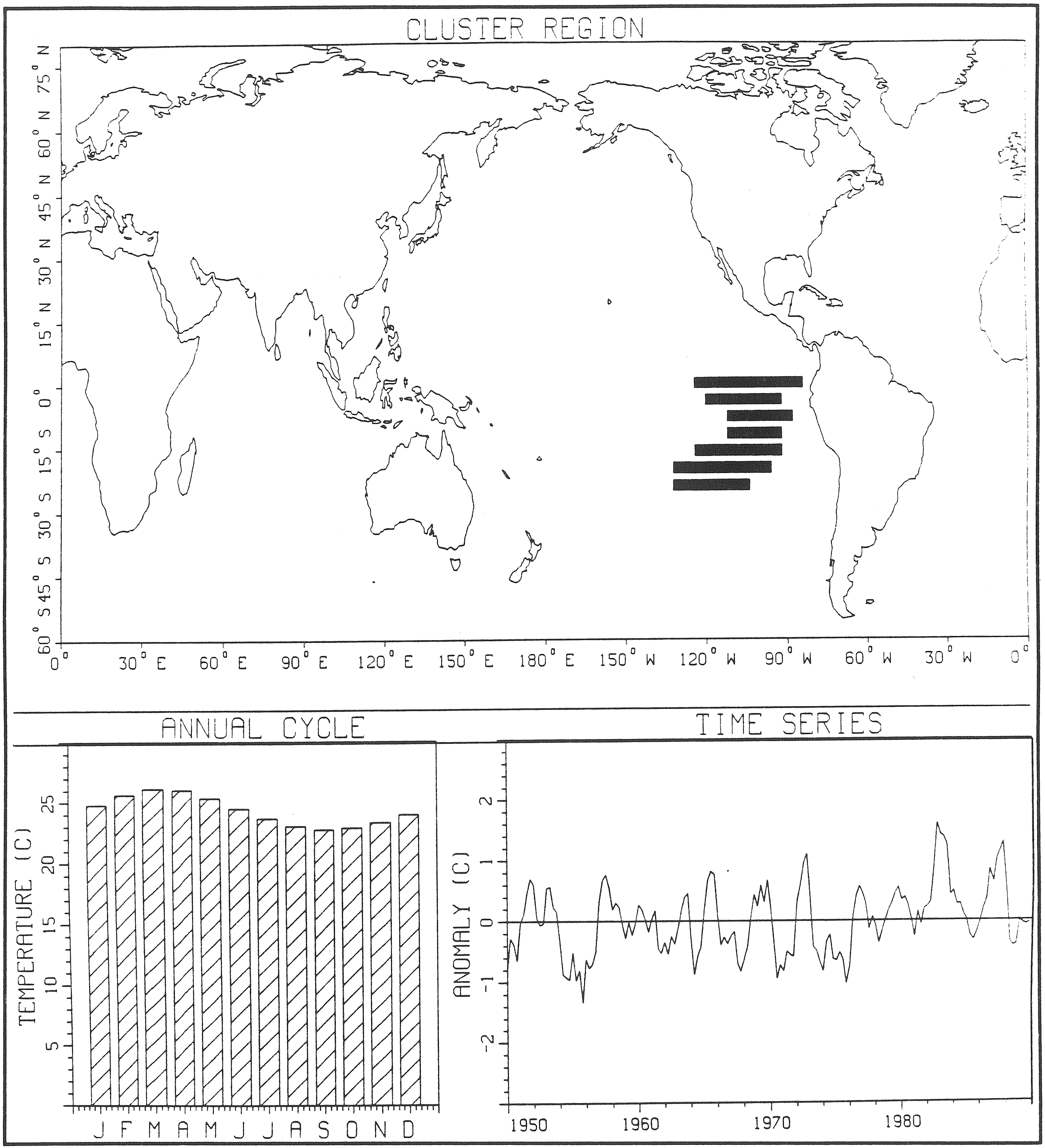

Figure 4. Time series of SST anomalies for an El Niño-sensitive area of the eastern tropical Pacific (region 11 in Figure 1). Figure also shows the seasonal cycle of monthly mean SSTs averaged over the entire region.

For the northern North Pacific region (see Table 1, region 1), seasonal temperature anomalies exhibit a cooling trend of about 0.4°C/decade, although SSTs in the 1980s were steady there, apart from typical interannual variability. Region 10, the large tropical region that encompasses parts of both the Indian and Pacific oceans, has the highest annual mean SST of any of the 34 global regions thus classified. It displays a small (though statistically significant) linear trend over the last four decades of about a tenth of a degree Celsius per decade. Most of the increase in SST is associated with a period of above-average SST beginning in the late 1970s. The other tropical region discussed here (region 11) is strongly affected by the El Niño/Southern Oscillation phenomenon. It has an upward SST trend of about the same magnitude as region 10, but because of its greater interannual variability, the change is not statistically significant.

Table 1
SELECTED SUMMARY STATISTICS BASED ON
ANNUAL SST MEANS (°C) FOR EACH CLASSIFIED REGION IN THE INDIAN AND PACIFIC OCEANS

See Figure 1.
The ordinary least squares trend (OLS) is given in °C/decade, and its t-value is shown as an indicative measure of statistical significance.

Region	Mean	S.D.	OLS Trend	t-Value
1	7.56	0.58	–0.37	–6.89
2	8.14	0.63	–0.36	-5.57
3	15.65	0.54	–0.27	–4.49
4	18.92	0.41	–0.23	-5.45
5	17.72	0.26	–0.01	–0.40
6	23.21	0.24	–0.08	-2.69
7	25.36	0.16	–0.02	–0.80
8	27.78	0.24	0.13	4.97
9	27.51	0.21	0.06	2.45
10	27.86	0.23	0.10	3.90
11	24.32	0.49	0.12	1.79
12	25.75	0.28	0.17	6.25
13	25.53	0.21	0.06	2.19
14	20.61	0.57	0.16	2.20
15	20.74	0.33	0.20	6.19
16	20.82	0.22	0.03	0.85
17	15.74	0.27	0.11	3.25
18	15.55	0.24	–0.05	-1.62
19	16.00	0.35	0.05	1.10
20	9.70	0.68	–0.16	-1.81

Principal component analysis using the 20 regional time series of 160 seasonal values (SST anomalies) as input indicates the first four (significant) eigenvector patterns account for 57 percent of the regional seasonal SST variance. We note that spatial patterns of the first two global regional eigenvectors (Diaz and Brown 1992) are similar to the second and third eigenvectors of SST anomaly calculated by Parker and Folland (1991) for the period since 1901 — their first eigenvector representing the long-term trend in the (adjusted) SST record. The temporal coefficients of our first global regional eigenvector closely follow the variations in global mean SST anomalies during this time ($r = 0.79$).

For comparison purposes, the variance spectrum of each regional anomaly time series was calculated. Most significant spectral peaks occur near the biennial time scale (2.1-2.2 years) and in the ENSO frequency band (~3-6 years). We have summarized these results in Figure 5, which illustrates the frequency bands in the SST variance spectrum containing significant power. For instance, in regions 10 and 11, primary peaks in the seasonal variance spectrum of SST anomalies (peaks under 2 years excluded) are in the biennial and ENSO frequency band. For the northern North Pacific region, both decadal scale and ENSO variability are found.

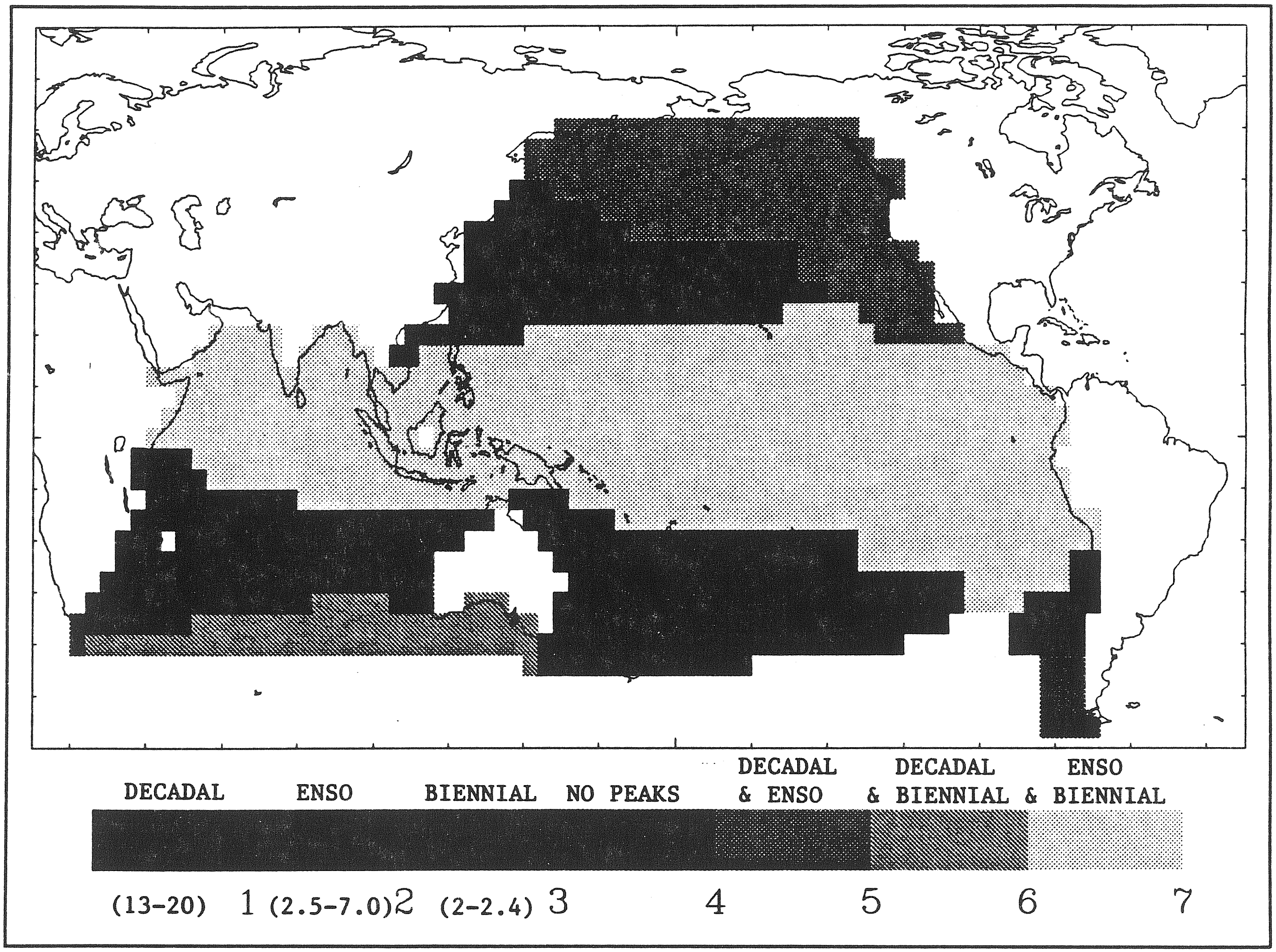

Figure 5. Map illustrating the characteristic time scale of temporal variability in the 20 Indo-Pacific regions.

Summary

The approach used here to objectively classify climate regions over the ocean departs from the usual methodology, which has relied on similarity of the interannual variability (as measured by the cross-correlation between area time series) to define spatially "homogeneous" regions. Here we have focused on similarity of the annual cycle of SST and applied a clustering algorithm to objectively classify oceanic SST. The results are useful for monitoring SST changes in upwelling areas along the West Coast of the Americas, the warm water pool in the equatorial Pacific and Indian oceans, *etc.*

Analysis of the variance spectrum of seasonal SST anomalies during the past 4 decades for our area of study illustrates regional differences in their characteristic time scale of variability. In general, the tropical ocean regions exhibit strong biennial and ENSO scale variability, whereas the extratropical regions contain greater variance at decadal time scales.

Acknowledgments

This work is part of an effort by several NOAA units, with funding from NOAA's Climate and Global Change Program, to develop a global climate perspectives system for climate diagnostics research and climate change detection efforts.

References

Diaz, HF, and TJ Brown. 1992. Application of statistical techniques for the regionalization of marine surface temperatures. Preprints, 12th AMS Conference on Prob. and Stat. in the Atmos. Sci., Toronto, Canada. pp. J20-J21.m

Hartigan, JA. 1975. *Clustering Algorithms.* J. Wiley & Sons, NY. 351 pp.

Parker, DE, and CK Folland. 1991. Worldwide surface temperature trends. In ME Schelesinger (ed.) *Greenhouse-Gas-Induced Climatic Change: A Critical Appraisal of Simulations and Observations.* Amsterdam, Elsevier. pp. 173-193.

Skaggs, RH. 1975. Drought in the United States, 1931-40. *Annals of the Assoc. of Amer. Geogr.* 65:391-402.

Stidd, CK. 1967. The use of eigenvectors for climatic estimates. *J. Appl. Meteor.* 6:255-264.

Woodruff, SD, RJ Slutz, RL Jenne, and PM Steurer. 1987. A comprehensive ocean-atmosphere data set. *Bull. Amer. Meteor. Soc.* 68:1239-1250.