Blending of COADS and UK Meterological Office Marine Data Sets

D.E. Parker, Hadley Centre, Meteorological Office, Bracknell, UK

Abstract

Following the confirmation by Woodruff (1990) that both the COADS and the Meteorological Office Marine Data Bank (MDB) contain unique observations, plans are being made for an observation-by-observation blend. A provisional blend of 5° latitude x 5° longitude monthly sea surface temperature (SST) anomalies yielded much improved coverage in the eastern Pacific in the late 19th and early 20th Centuries, affecting estimates of global temperature trends by up to 0.1°C. Representation of the major El Niño of 1877-8 was substantially improved. A blend of full observations will give further improvements because duplicates can be eliminated, values will be quality controlled, and parameters other than SST will be incorporated into the expanded data base. Coverage of all data needs to be maximized, to enable a more complete description of climatic variations, and to allow verification of numerical model simulations of decadal to century timescale atmospheric changes forced by large-scale SST changes observed since the late 19th Century. To this end, data which as yet are neither in COADS nor in the MDB also need to be added to the blend.

A brief resume is given of the logistics involved in the blend. Problems include the criteria for identifying duplicates; the combination of non-identical repeats into single, optimal observations; errors in location and time of observations; exchange formats including flags for quality control, units, and methods of measurement; and the need for regular updates.

1. Introduction

The collation of a quality-controlled data bank of marine observations, covering as much of the globe as possible since at least the mid 19th Century, is essential for the monitoring, analysis and understanding of climatic variations. The assessments of climatic changes made for the Intergovernmental Panel on Climate Change (IPCQ (Folland et al., 1990; 1992) depend heavily on estimates of changes of sea surface temperature (SST) as well as land surface air temperature. Other data sources, such as surface pressure and wind measured on board ships and buoys, are increasingly recognized as being important. The marine data banks used in the IPCC reports were the COADS (Woodruff et al., 1987) and the UK Meteorological Office Marine Data Bank (MDB) (Bottomley et al., 1990), and include quality-controls as described by these authors. Bottomley et al. (1990) also present corrections for systematic biases in SST and marine air temperature: the corrections to SST have since been refined (Folland, 1991). Systematic corrections have also been made to COADS SST (Jones et al., 1991; and Jones and Wigley, this volume).

Bottomley et al. (1990) showed that the COADS contained more observations than the MDB, though the difference in content varied through history and was greatest in the 1960s and 1970s, a deficiency since reduced by the addition of missing Pacific data to the MDB. Woodruff (1990) presented greater detail in a comparison of the two data banks for a selection of 10° latitude x longitude boxes. Although the greater overall content of the COADS was confirmed, a substantial number of observations were also found to be unique to the MDB (Fig. 1). Clearly, then, an observation-by-observation blend of the two data banks would yield a data base larger than either. The preliminary combination of results from COADS and the MDB in Folland et al. (1992) also underlined the potential value of a full blend of the data. Coverage of 5° latitude x longitude seasonal SST data between 1880 and 1910 increased by nearly 20% of the world's oceanic area when a grid- box analysis based on the MDB was supplemented by corresponding gridded values derived from COADS (Table 1). The improvement in coverage was particularly marked in the Eastern Pacific (Fig. 2a, b). Table 1 also indicates useful, though smaller, increases in coverage in more recent decades. However, the combination was inadequate in that only SST was involved, and in the restriction of the accompanying quality-control to extreme value checks and spatial and temporal near-neighbour checks (Table 2). An observation-by-observation blend would include all parameters, and would reduce the expected root-mean-square error of grid-box values by increasing the number of constituent observations (Trenberth and Hurrell, this volume).

2. Benefits for Research

a. Estimating global climatic trends

Bottomley et al. (1990) and Folland et al. (1990) include the results of frozen-grid experiments which show that estimates of global and hemispheric temperature trends are surprisingly insensitive to the major historical variations of coverage of data. Nevertheless, confidence in these estimates is reduced by the major permanent or almost permanent gaps in coverage in the analyses. See also Ropelewski et al. (this volume). The addition of COADS SSTs to those from the MDB in Folland et al. (1992) had a significant effect on estimates of global SST trends (Fig. 3), because SST anomalies between 1870 and 1910 in the extra areas covered by COADS, mainly the Eastern Pacific, were less cold than elsewhere in the world in that period. A full blend of the MDB and COADS will further improve the reliability of these estimates (see Introduction) and will allow corroboration using marine air temperatures. Substantial further benefits will accrue from the inclusion of additional US data, including Maury's, and Japanese data, which are yet to be digitized (Elms, this volume; Uwai and Komura, this volume); and from the blending in of a Russian data set (Yudin et al., this volume), as many millions of extra observations will thus be added from these sources.

b. Studies of regional climate

An expanded, quality-controlled data bank will facilitate research into regional climatic phenomena, such as sub Saharan rainfall (Folland et al., 1991), Australian rainfall (Nicholls, 1989; Drosdowsky, this volume) and the El Niño - Southern Oscillation (ENSO). Major early ENSO events such as that of 1877-8 (Allan et al., 1991) will be more clearly delineated, and an improved knowledge of the history of ENSO may help to elucidate its role in longer-term climatic fluctuations. Figures 4a, b illustrate a major pattern of variation evident in the preliminary combined COADS-MDB SST data set after filtering to remove shorter-than-decadal timescale fluctuations (Folland and Colman, personal communication). This filtering deliberately removes the variations on 3 to 7 year timescales normally associated with ENSO, but the field shown in Figure 4a has marked similarities to the classical ENSO SST pattern (Glantz et al., 1991). Thus, a longer term role of ENSO-like variations is implied (Fig. 4b). A full blend of the data banks is needed, if the air temperature, humidity, wind and pressure fields linked to ENSO and to Fig. 4 are to be properly analysed and understood.

c. Input to model simulations of climate

An improved understanding of recent (and possible future) climatic fluctuations is likely to be obtained by forcing general circulation models of the atmosphere with observed historical SST and sea-ice information, and comparing the response of the models with the observed behaviour of the atmosphere (e.g. Hense et al., 1990). To optimize this research, the forcing fields should be as complete and accurate as possible before they are interpolated to provide the globally complete fields required by the models. Complete atmospheric fields are also required for forcing ocean models (Jessel, this volume).

The verification of the atmospheric model experiments requires the best possible coverage of atmospheric data over the oceans as well as over land: wind, pressure, air temperature, dewpoint and possibly also cloudiness. Spatially complete SST is also needed to verify oceanic and combined atmosphere-ocean models. A full blend of COADS and MDB is an important step towards these objectives.

d. Cross-verification

An observation-by-observation blend involving all major parameters will enable research results using one parameter to be more readily corroborated using others. Specifically:

> i) Marine air temperature supports SST, because their anomalies on regional ocean-basin scales over seasons and longer are highly correlated (Bottomley et al., 1990);
> ii) Dewpoint may support SST and air temperature (Parker, 1988);
> iii) Wind and pressure data provide mutual support. Biases in wind information may be estimated and removed by using pressure gradients and other information (Diaz et al., this volume; Wu and Newell, this volume; Ward (1992))

The denser the data coverage, the more powerful can these cross-checks be.

3. Quality Control

a. Systematic biases

The blended data bank should include coded information on instrumentation and observing practices in all observations for which these are known, so that the user can adjust results accordingly.

Examples are:

> i) SST: was the observation taken in an engine intake, or using a canvas or wooden or rubber bucket? (Bottomley et al., 1990).

ii) Air temperature and dewpoint: what kind of screen, if any, was used?
iii) Wind: were speeds estimated visually and if so, what scale was used to convert from Beaufort forces to m/sec?; or measured by anemometer, and if so, at what height? (Isemer and Hasse, this volume; da Silva et al., this volume; Cardone et al. (1990)).
iv) All data: was the platform a ship, an anchored buoy, or a drifting buoy?

b. Precision

The units in which the original recordings were made, and any rounding or truncation of decimals in the process of conversion, should be clearly indicated.

c. Duplicate observations

Exact duplicates should, of course, be removed from the blend.

Near-duplicates should be combined into single, optimized records, or the best copy should be selected on the basis of internal consistency and agreement with previous and subsequent reports from, if possible, the same platform.

The retained observation should include an indicator of its source, e.g. COADS, MDB, or keyed USA or Japanese data, or combination of near-duplicates.

4. Logistics

The form at of the output must be determined by the foregoing considerations and by the need for WMO (Commission for Maritime Meteorology (CMM)) recognition. The enhanced COADS is expected to include observations in a modified Long Marine Report (LMR) format (see Slutz et al., 1985) but data output for international exchange should be in format TD-1120 which is being proposed to CMM as the recognized format for exchange of historical marine data (J. Elms, pers. comm.).

Annual updates of the blend are anticipated. These should include both new data, and any newly digitized old data as they become available.

5. Acknowledgements

J. Elms, B. Fullagar, M. Jackson and S. Woodruff participated in many discussions on the design of the COADS-MDB blend.

References

Allan, R.J., Nicholls, N., Jones, P.D., and Butterworth, I.J., 1991: A further extension of the Tahiti- Darwin SOI, early ENSO events and Darwin pressure. *Journal of Climate*, **4**, 743-749.

Bottomley, M., Folland, C.K., Hsiung, J., Newell, R.E., and Parker, D.E., 1990: *Global Ocean Surface Temperature Atlas (GOSTA)*. Joint Meteorological Office/Massachusetts Institute of Technology Project. Project supported by US Dept. of Energy, US National Science Foundation and US Office of Naval Research. Publication funded by UK Depts. of Energy and Environment. 20+iv pp and 313 Plates, HMSO London.

Cardone, V.J., Greenwood, J.G.m and Cane, M.A., 1990: On trends in historical marine wind data. *Journal of Climate*, **3**, 113-127.

Diaz, H.F., Fu C. and Quan X., 1992: A comparison of surface geostrophic winds with COADS ship wind observations. *In:* Diaz, H.F., Wolter, K., and Woodruff, S.D., (eds.), *Proceedings of the International COADS Workshop, Boulder, Colorado, 13-15 January 1992,* NOAA Environmental Research Laboratories, Boulder, Colorado (this volume).

Drosdowski, W., 1992: Relationships of global SST patterns to Australian rainfall variations. *In:* Diaz, H.F., Wolter, K., and Woodruff, S.D., (eds.), *Proceedings of the International COADS Workshop, Boulder, Colorado, 13-15 January 1992*, NOAA Environmental Research Laboratories, Boulder, Colorado (this volume).

Elms, J.D., 1992: Status of NCDC keying of historical marine data. *In:* Diaz, H.F., Wolter, K., and Woodruff, S.D., (eds.), *Proceedings of the International COADS Workshop, Boulder, Colorado, 13-15 January 1992*, NOAA Environmental Research Laboratories, Boulder, Colorado (this volume).

Folland, C.K., 1991: *Sea temperature bucket models used to correct historical SST data in the Meteorological Office*. Climate Research Technical Note CRTN 14, Available from Hadley Centre, Meteorological Office, London Road, Bracknell, Berks RG12 2SY, UK., 29 pp.

Folland, C.K., Karl, T.R., Nicholls, N., Nyenzi, N.D., and Vinnikov, K.Ya., 1992: *Observed climate variability and change*. Section C of the IPCC Science Report Supplement. IPCC, Cambridge University Press, in press.

Folland, C.K., Karl, T.R., and Vinnikov, K.Ya., 1990: Observed climate variations and change. *In:* Houghton, J.T., Jenkins, G.J. and Ephraums, J.J., (eds.), *Climate Change, the IPCC Scientific Assessment*, WMO/UNEP, IPCC, Cambridge University Press, pp. 195-238.

Folland, C.K., Owen, J., Ward, M.N., and Colman, A., 1991: Prediction of seasonal rainfall in the Sahel region using empirical and dynamical methods. *Journal of Forecasting*, **10**, 21-56.

Glantz, M.H., Katz, R.W., and Nichols, N., (eds.), 1991: *Teleconnections linking worldwide climate anomalies. Scientific basis and societal impact*. Cambridge University Press, 535 pp.

Hense, A., Glowienka-Hense, R., von Storch, H., and Stahler, U., 1990: Northern Hemisphere atmospheric response to changes of Atlantic Ocean SST on decadal time scales: a GCM experiment. *Climate Dynamics*, **4**, 157-174.

Isemer, H.-J. and Hasse, L., 1992: The scientific Beaufort equivalent scale: Effects on wind statistics and climatological air-sea flux estimates in the North Atlantic Ocean. *In:* Diaz, H.F., Wolter, K., and Woodruff, S.D., (eds.), *Proceedings of the International COADS Workshop, Boulder, Colorado, 13-15 January 1992*, NOAA Environmental Research Laboratories, Boulder, Colorado (this volume).

Jessel, I., 1992: Scientific work at MPI in Hamburg in context with COADS. *In:* Diaz, H.F., Wolter, K., and Woodruff, S.D., (eds.), *Proceedings of the International COADS Workshop, Boulder, Colorado, 13-15 January 1992*, NOAA Environmental Research Laboratories, Boulder, Colorado (this volume).

Jones, P.D. and Wigley, T.M.L., 1992: Corrections to pre-1941 SST measurements for studies of long- term changes in SSTs. *In:* Diaz, H.F., Wolter, K., and Woodruff, S.D., (eds.), *Proceedings of the International COADS Workshop, Boulder, Colorado, 13-15 January 1992*, NOAA Environmental Research Laboratories, Boulder, Colorado (this volume).

Jones, P.D., Wigley, T.M.L., and Farmer, G., 1991: Marine and land temperature data sets: A comparison and a look at recent trends. *In:* Schlessinger, M.E., (ed.), *Greenhouse-gas -induced climatic change: a critical appraisal of simulations and observations*, Developments in Atmospheric Science, 19, Elsevier, pp. 153-172.

Nicholls, N., 1989: Sea surface temperatures and Australian winter rainfall. *Journal of Climate,* **2**, 965-973.

Parker, D.E., 1988: A note on the use of voluntary observing fleet data to estimate air-sea fluxes. Long-Range Forecasting and Climate Research LRFC 17, 25 pp. Available from Hadley Centre, Meteorological Office, London Road, Bracknell, Berks, RG12 2SY, UK.

Ropeleweski, C.F., Marsico, D., Chelliah, M., and Smith, T., 1992: What can COADS tell us? Potential and Limitations. *In:* Diaz, H.F., Wolter, K., and Woodruff, S.D., (eds.), *Proceedings of the International COADS Workshop, Boulder, Colorado, 13-15 January 1992*, NOAA Environmental Research Laboratories, Boulder, Colorado (this volume).

da Silva, A.M., Young, C.C., and Levitus, S., 1992: Estimates of wind stress and heat fluxes with bias corrections. *In:* Diaz, H.F., Wolter, K., and Woodruff, S.D., (eds.), *Proceedings of the International COADS Workshop, Boulder, Colorado, 13-15 January 1992*, NOAA Environmental Research Laboratories, Boulder, Colorado (this volume).

Slutz, R.J., Lubker, S.J., Hiscox, J.D., Woodruff, S.D., Jenne, R.L., Joseph, D.H., Steurer, P.M. and Elms, J.D., 1985: *COADS. Comprehensive Ocean-Atmosphere Data Set Release 1*. CIRES/ ERL/NCAR/NCDC. Boulder, Colorado.

Trenberth, K.E. and J.W. Hurrell, 1992: Monitoring global monthly mean surface temperature. *In:* Diaz, H.F., Wolter, K., and Woodruff, S.D., (eds.), *Proceedings of the International COADS Workshop, Boulder, Colorado, 13-15 January 1992*, NOAA Environmental Research Laboratories, Boulder, Colorado (this volume).

Uwai, T. and Komuru, K., 1992: The collection of historical ships' data in Kobe Marine Observatory. *In:* Diaz, H.F., Wolter, K., and Woodruff, S.D., (eds.), *Proceedings of the International COADS Workshop, Boulder, Colorado, 13-15 January 1992*, NOAA Environmental Research Laboratories, Boulder, Colorado (this volume).

Ward, M.N., 1992: Provisionally corrected surface wind data, worldwide ocean-atmosphere surface fields and Sahel rainfall variability. *Journal of Climate*, in press.

Woodruff, S.D., Slutz, R.J., Jenne, R.L., and Steurer, P.M., 1987: A Comprehensive Ocean-Atmosphere Data Set. *Bulletin of the American Meteorological Society,* **68**, 1239-1250.

Woodruff, S.D., 1990: Preliminary comparison of COADS (US) and MDB (UK) ship reports. Chapter XXVII (36 pp) of *Observed Climate Variations and Change: Contributions in Support of Section 7 of the 1990 IPCC Scientific Assessment*, D.E. Parker, Ed., IPCC.

Wu, Z. and Newell, R.E., 1992: The wind problem in COADS and its influence on the water balance. *In:* Diaz, H.F., Wolter, K., and Woodruff, S.D., (eds.), *Proceedings of the International COADS Workshop, Boulder, Colorado, 13-15 January 1992*, NOAA Environmental Research Laboratories, Boulder, Colorado (this volume).

Yudin, K.B., Ylyanich, I.G., Popova, V.N., and Krakanovskaya, Ye.M., 1992: Russian Marine Meteorological Data Set. *In:* Diaz, H.F., Wolter, K., and Woodruff, S.D., (eds.), *Proceedings of the International COADS Workshop, Boulder, Colorado, 13-15 January 1992*, NOAA Environmental Research Laboratories, Boulder, Colorado (this volume).

Table 1. Percentage of oceanic coverage with seasonal SST data for past decades. [One month's data can constitute a season in a 5° lat x long box].

	NDB Alone	MDB Supplemented with COADS
1861 - 70	21	25
1871 - 80	28	41
1881 - 90	37	55
1891 - 1900	33	51
1901 - 10	40	61
1911 - 20	41	49
1921 - 30	57	61
1931 - 40	60	65
1941 - 50	54	58
1951 - 60	71	76
1961 - 70	78	82
1971 - 80	79	84
1981 - 90	88	89

Table 2. Quality controls on combined MDB-COADS 5° latitude x longitude monthly SST.

1.Convert all 5° lat x long box monthly values to anomalies from Bottomley et al. (1990) 1951-80 climatology.
2.Set anomalies outside +/-7°C to missing.
3.Calculate average anomaly for surrounding boxes if >4 have data: substitute this average in the box if its anomaly is missing or differs from it by more than 2.25°C.
4.Calculate average anomaly for previous and subsequent months if both are available: substitute this average in the box if its anomaly is missing or differs from it by more than 2.25°C.
5.Repeat (3) and (4) twice: ie do them 3 times all told.
6. For “MOHSST5”, boxes which had missing values initially are re- set to missing, thus retaining a realistic data coverage for analysis of regional and global trends.
7. For “MOHSST5A”, retain all interpolated data. MOHSST5A, after blending with sea ice, will be used in numerical model integrations, simulating the climate of the 20th Century.

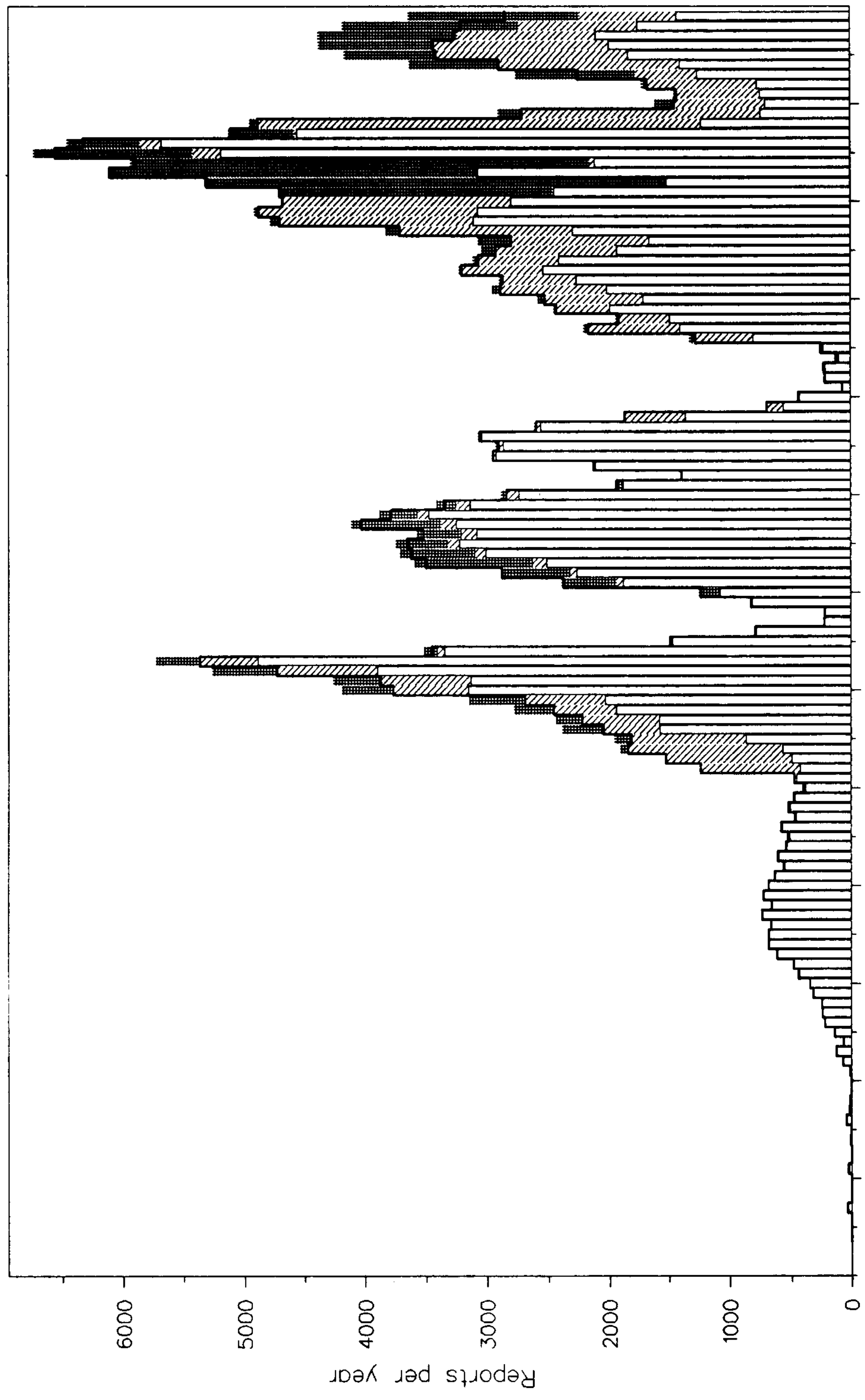

Figure 1. Annual numbers of observations in the area 10° - 20°N, 60° - 70°E. Hollow bars, COADS + MDB duplicates; light shading, unique to COADS; heavy shading, unique to MDB; heavy line, maximum of COADS and MDB if not blended. From Woodruff (1990).

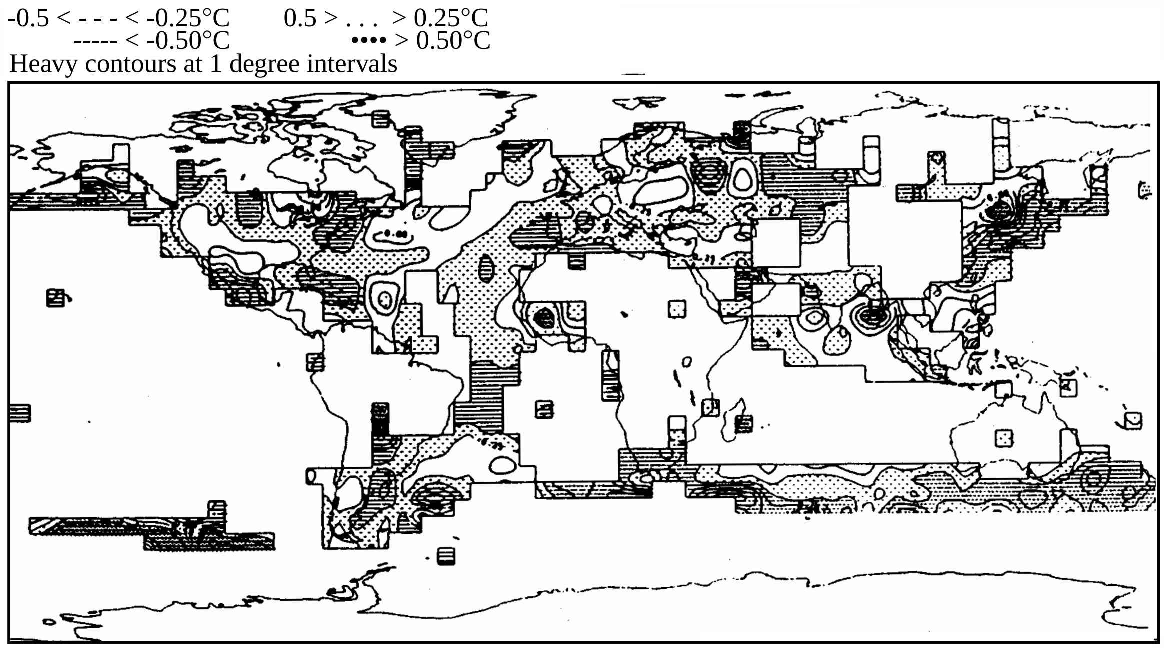

Figure 2a. Surface temperature anomalies, 1901-10 w.r.t. 1951-80. Oceanic data are SST based on MDB. Land air temperatures provided by P.D. Jones, Climatic Research Unit, University of E. Anglia, UK. The “refined” corrections described by Bottomley et al. (1990) were applied to all SST data.

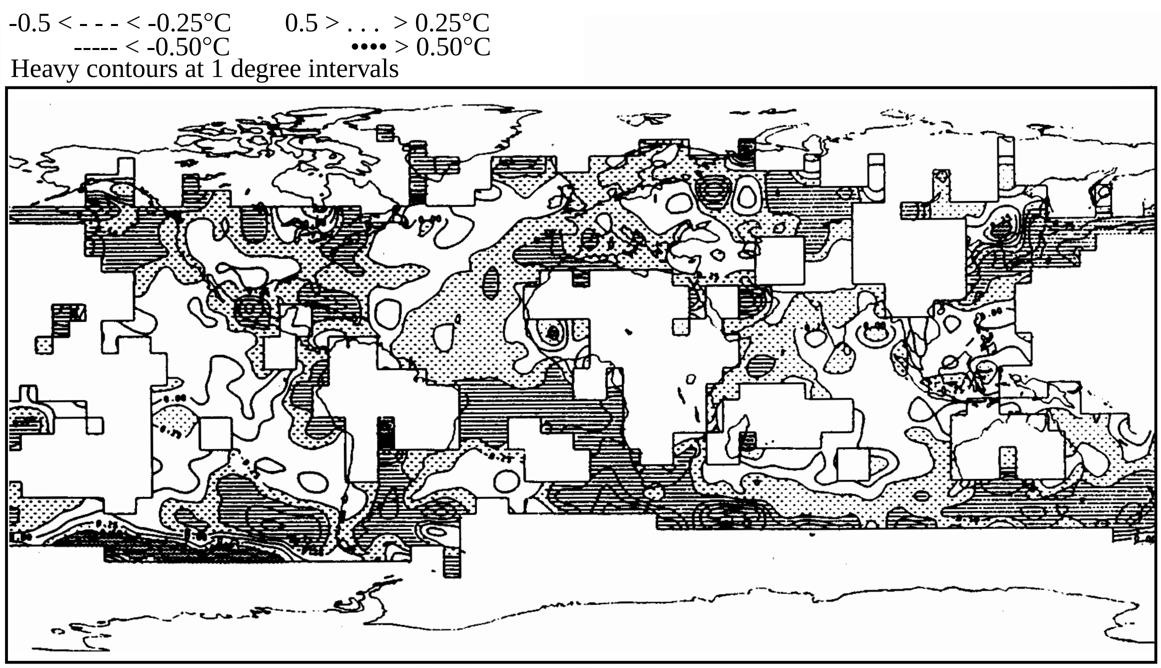

Figure 2b. As Figure 2a except that the MDB SST have been supplemented by COADS.

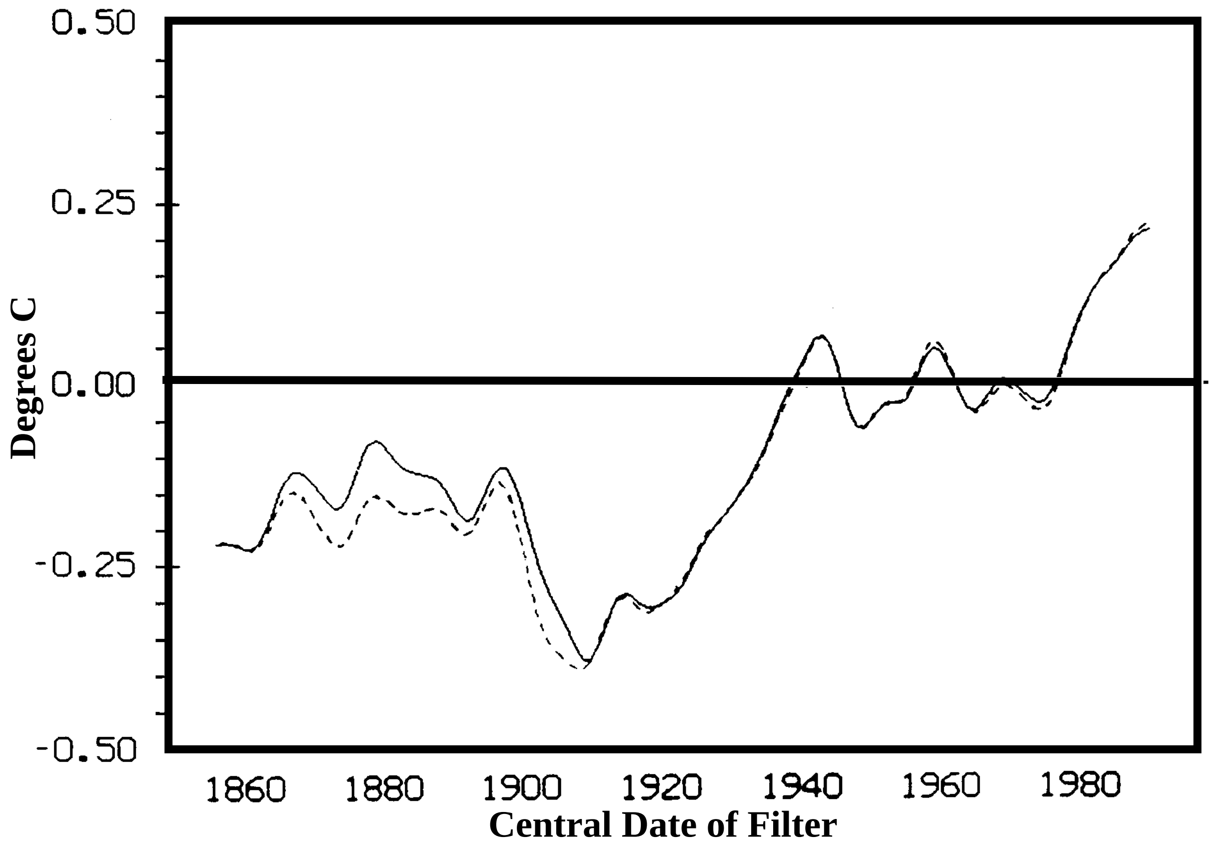

Figure 3. Smoothed time-series of globally-averaged SST anomalies w.r.t. 1951-80. The original annual series were extended backwards (forwards) by 7 terms each equal to the average of their first (final) 7 terms: the resulting series were smoothed using a 15-point filter created by normalising terms 4 through 18 of the 21-term binomial expansion. The “refined” corrections described by Bottomley et al. (1990) were applied to all data up to the end of 1941. Solid line: MDB supplemented with COADS. Dashed line: MDB alone.

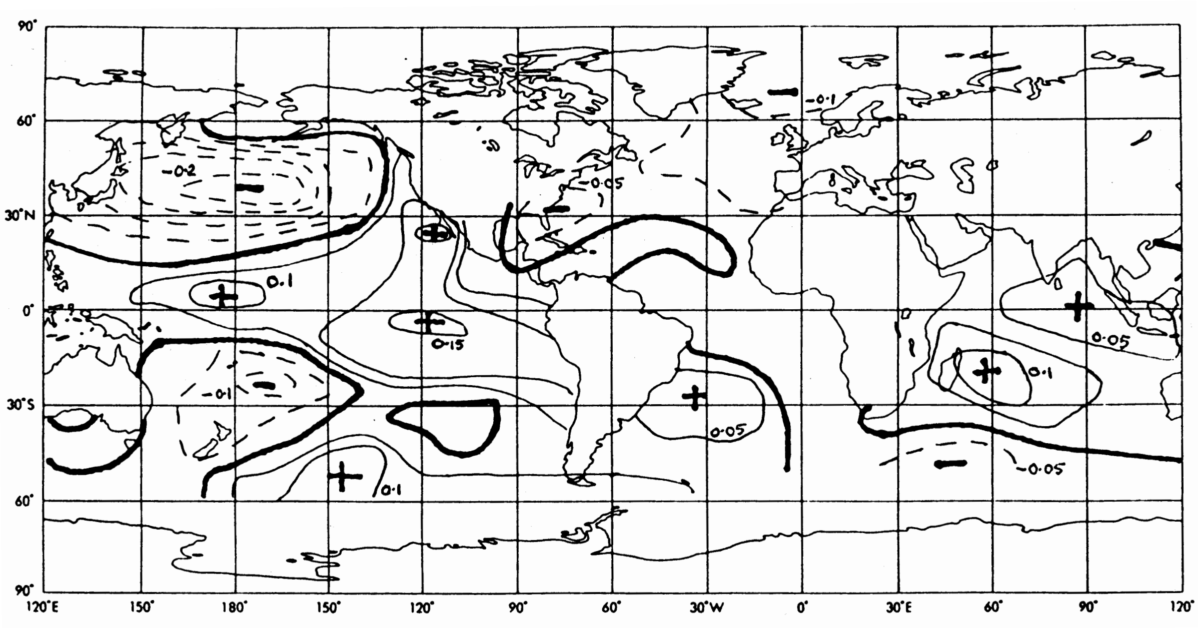

Figure 4a. Second global all-seasonal SST covariance eigenvector for 1901-1990, based on MDB supplemented with COADS and low-pass filtered to pass interdecadal variations. Contours at intervals of 0.05 nondimensional units, scaled such that values in grid boxes with SST data sum to 1, these grid boxes being a subset of 272 equal areas covering the entire globe, land and sea.

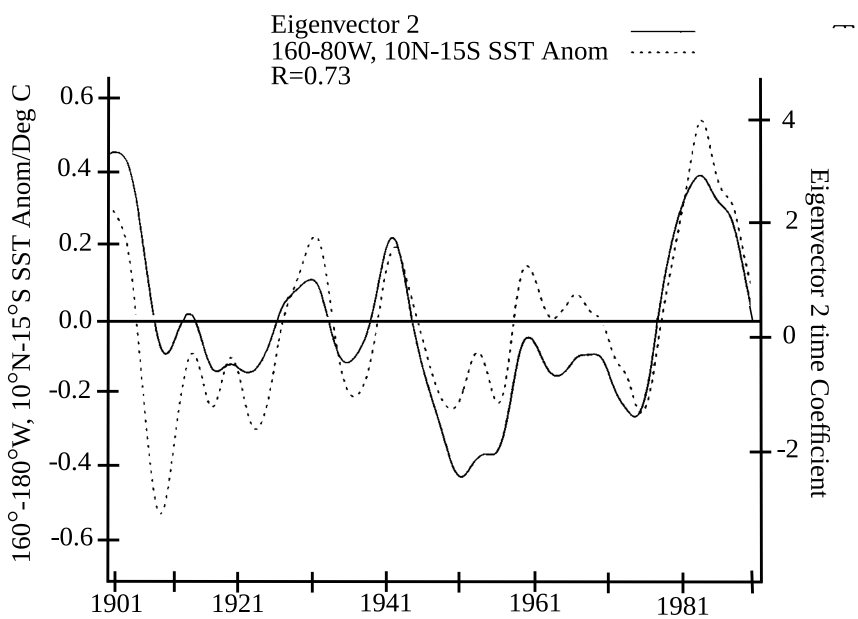

Figure 4b. Time sequence of the coefficient of the pattern shown in Figure 4a compared with SST anomalies for the area 10°N - 15°S, 80° - 160°W. The correlation between the two curves is 0.73.