JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 90, NO. C6, PAGES 11,642–11,650, NOVEMBER 20, 1985

Production of Global Sea Surface Temperature Fields for the Jet Propulsion Laboratory Workshop Comparisons

JEFFREY E. HILLAND

Jet Propulsion Laboratory, California Institute of Technology, Pasadena

DUDLEY B. CHELTON

College of Oceanography, Oregon State University, Corvallis

ENI G. NJOKU[1]

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

Sea surface temperature (SST) is measured from space by the advanced very high resolution radiometer (AVHRR), scanning multichannel microwave radiometer (SMMR), high resolution infrared sounder (HIRS) and VISSR atmospheric sounder (VAS). Typical accuracies have been reported from 0.5°C regionally to 2.0°C on a global basis. To evaluate the accuracy of the satellite-derived sea surface temperatures, a series of three workshops was organized to provide uniform data reduction and analysis. The analytical techniques used to intercompare satellite and in situ measurements are described in detail. Selected results showed the overall average rms errors were in the range 0.5°–1.0°C.

1. INTRODUCTION

To provide uniform data processing for the sea surface temperature (SST) workshop series, principal investigators agreed to send SST retrievals produced by their best algorithms to the JPL/Pilot Ocean Data System. Uniform data reduction was assured by providing a single processing system and a set of analytical procedures. The algorithms were designed to intercompare sensor performances on global and ocean basin length scales over four month-long periods: November 1979, Decemer 1981, March 1982, and July 1982. Thus the sea surface temperature evaluation scheme reflected measurement consistency from instrument to instrument as well as temporal changes in the ocean.

The impetus for making an SST accuracy assessment came from the perception that surface temperatures were routinely being determined from space with an accuracy of better than 1°C. Since December 1981, the advanced very high resolution radiometer (AVHRR) has made SST measurements in the infrared portion of the spectrum by using a multichannel technique developed by *McClain et al.* [1983]. In order to determine SST more accurately under cloudy conditions, infrared soundings from the high resolution infrared sounder (HIRS) and microwave soundings from the microwave sounding unit (MSU) have been combined in a scheme described by *Susskind et al.* [1982]. Another instrument, the visible-infrared spin scan radiometer atmospheric sounder (VAS), has provided daytime SST retrievals from geostationary orbit [*Smith and Woolf*, 1981]. The ocean surface has been been viewed in the microwave portion of the spectrum by the scanning multichannel microwave radiometer aboard Nimbus 7. *Wilheit et al.* [1983] have described the SMMR SST retrieval techniques in detail.

In contrast to satellite methods, in situ data collected from ships, expendable bathythermographs (XBT's), and drifting buoys have provided independent bulk measurements of SST. These platforms have long served as oceanographers' primary tools. Hence a large body of knowledge has been compiled with the result that biases on the order of +0.4°C and rms accuracies in the range 0.2°–1.0°C have been reported by *Barnett* [1984], *Tabata* [1982], and *Saur* [1963].

Against this background of spacecraft and in situ measurements, NASA organized a series of workshops to determine sensor SST accuracies, to evaluate overall system performance, and to make recommendations for future sensor development. Workshop planning and initial results are discussed in *NASA/JPL* [1983]. Workshop 2 and 3 results were more comprehensive as a result of refinements in the analysis procedures. In addition, more data were available for analysis with the acquisition of five-channel AVHRR, HIRS/MSU, and VAS data sets described in *NASA/JPL* [1984, 1985].

We will describe the data set characteristics in detail. Specifically, resolution, sampling characteristics, and data quality will be reviewed. These features, to a large extent, determined the methodology used for data reduction and analysis. The algorithms used to derive SST from each instrument are presented in the accompanying papers and will not be discussed here. However, the rationale behind development of analysis techniques used in the workshop environment as well as the procedures themselves will be delineated. Finally, to provide a complete overview of the processing system, examples of data products and images discussed in the companion papers are presented in the appendix.

2. DATA SET CHARACTERISTICS

Because each sensor collected data in a unique manner, as a result of resolution and scanning methods but also as result of sensor duty cycles (the percentage of time the sensor operated satisfactorily measured relative to the total time assigned for

[1]Also at Department of Engineering, Harvey Mudd College, Claremont, California.

Paper number 5C0643.
0148-0227/85/005C-0643$05.00

TABLE 1. Sensors Evaluated During the SST Workshops and Resolution of Derived SST's

Platform	Sensor	Duty Cycle, %	Spatial Resolution of Derived SST, km	Coverage
TIROS N, NOAA 6, 7	AVHRR	100	25	global
TIROS N, NOAA 6, 7	HIRS/MSU	100	125	global
Nimbus 7	SMMR	15	150	global
GOES	VAS	54	60	NW Atlantic, SE Pacific,
Ships	thermometer			global
XBT	thermistor			North Pacific
FGGE buoys	thermistor			southern hemisphere

operation) and processing, the data distribution varied greatly. (Details of these sensors are provided in other papers in this issue.) Table 1 summarizes pertinent sampling parameters for each sensor. The comparison shows that the high-resolution, nearly continuous duty cycle of the IR-VIS sensors yields an enormous number of discrete radiances. Spacecraft-measured radiances were spatially averaged as part of the conversion process to geophysically meaningful temperatures before delivery to JPL, thus reducing the data volume. Other data set features pertinent to the processing and analysis are elaborated on below, sensor by sensor.

AVHRR

Global day/night coverage across a 2700-km-wide swath at 4-km resolution characterizes the fundamental sampling of AVHRR instruments aboard the TIROS-N and NOAA satellites. Prior to mid-November 1981 the Global Operational Sea Surface Temperature Computation (GOSSTCOMP) provided 50-km resolution SST retrievals by using a single window (centered at 11 μm) algorithm. After this date the improved five-channel instrument was used to derive SST from the 3.7-, 11- and 12-μm windows by utilizing the triple-window technique known as the multichannel sea surface temperature (MCSST) algorithm. In this improved technique, SST's were derived at spatial intervals of 25 km. The standard NOAA GOSSTCOMP and MCSST products provided geolocated SST's and supporting parameters such as platform source, data quality, and day/night status.

HIRS/MSU

The HIRS and MSU instruments flown on TIROS-N, NOAA 6, and NOAA 7 served as sources for derived SST's. The large number of IR and microwave channels are combined in a physical algorithm to produce surface temperatures under clear or cloudy conditions. Arrays of HIRS soundings (instantaneous field of view 17.4 km at nadir) are averaged across the 2300-km-wide swath to form SST's at a spatial resolution of 125 km. Because the retrievals are spatially averaged, points within 60 km of land tend to be contaminated by warmer land temperatures. For workshop purposes, space/time location and quality parameters were provided so that information could be segregated for various study months, regions, and sampling conditions, such as day/night and clear/cloudy.

SMMR

Dual polarized microwave radiance measurements at 6.6, 10.7, 18, and 21 GHz were the fundamental input to SMMR SST algorithms. The 37-GHz channels were not used to derive SST. The SMMR samples along a nadir-centered 780-km-wide swath with a spatial resolution of 150 km at 6.6 GHz. Various quality control criteria were applied to the data by the Nimbus 7 algorithm development team, but the most obvious and influential, with regard to workshop processing, was a land proximity mask. All data within 600 km of land, including large islands, were eliminated as a result of antenna sidelobe contamination. Furthermore, in order to distinguish the highest quality values, data sent to JPL (Jet Propulsion Laboratory) were flagged for day/twilight/night status and cell (1–5) location in the swath. Cell locations were established during the antenna pattern correction processing performed by the Nimbus 7 team. Briefly, the irregularly spaced swath points were averaged onto a regular grid that divided the swath width into five cells numbered sequentially 1–5, beginning with the leftmost cell. The end cells in the swath were deemed unreliable because of polarization correction errors. Furthermore, the instrument was turned on and off every other day because of spacecraft power limitations. Finally, because night retrievals were considered most reliable, the overall duty cycle was reduced to 15%.

VAS

The Geostationary Operational Environmental Satellite (GOES) carrying VAS provided a stable platform for scanning the full earth disc. During regular operations, one image is collected per hour for 18 hours during weekdays, resulting in a 54% duty cycle. However, daytime only IR and VIS data were collected for the workshop series. These data were screened for cloud-free areas as part of the SST derivation scheme. Three of the twelve thermal bands sensed by VAS were used to derive SST's at a spatial resolution of 50 km. Finally, retrievals from the eastern tropical Pacific and northwestern Atlantic were provided for evaluation.

In Situ

The primary surface data set consisted of ship intake temperature measurements collected from radio reports by the Fleet Numerical Oceanography Center (FNOC). Typically, intake temperatures are accurate to the nearest 1°C. Additionally, biases on the order of tenths of a degree Celsius have been reported [*NASA/JPL*, 1983]. However, these data are the sole source of global in situ measurements and at best provide spotty spatial coverage in the southern hemisphere. Temporal resolution of most reports is 6 hours. Marine reports were closely scrutinized for pathological errors related to erroneous ship locations and extreme temperatures.

Complementary data sets consisting of XBT drops across the northern and tropical Pacific and measurements from drifting buoys in the southern hemisphere provided additional validation information. XBT drops between North America, Japan, Hawaii, Tahiti, and the Panama Canal were made about every 200 km, yielding 400–1000 observations during any month. Stated accuracies [*NASA/JPL*, 1984] were 0.1°C with biases of about +0.16 to +0.38°C relative to salinity-temperature-depth (STD) instruments.

Drifting buoys launched during the First GARP Global Experiment (FGGE) supplemented ship observations in the southern hemisphere from 20°S to 65°S. More than 100 buoys reported SST's at roughly 6-hour intervals during November 1979. In an examination of the buoy program [*Garrett*, 1981], comparisons made with ship measurements within 1 hour and 100 km of the buoy observation yielded a worst case standard

TABLE 2. Analysis and Graphical Display Techniques Applied to Satellite-Derived SST's

Techniques	Data Type Analyzed		
	Raw Absolute	Raw Anomaly	Binned Anomaly
Compute raw anomalies by interpolating climatology to satellite point and subtracting	X		
Form 2° latitude × 2° longitude monthly averages (binning)		X	
Calculate absolute SST by adding binned anomalies to 2° binned climatological SST	X		X
Histogram of SST and summary statistics: mean, standard deviation, number of observations		X	
Contour map of binned absolute SST	X		X
Thematic map of binned anomalies			X
Thematic map of anomaly differences			X
Thematic map of number density within a 2° cell			X
Scatter diagram of binned anomaly SST versus Reynolds climatological SST Summary statistics: bias, standard deviation about bias, correlation, number of observations	X		X
Cross-correlation tables. Statistics: correlation, bias, standard deviation about bias, number of 2° cells			X
Error partitioning tables. Statistics: overall rms error contributed by each sensor and rms error for each sensor combined with two other sensors.			X
Scatter diagram of sensor versus sensor. Statistics: correlation, bias, standard deviation about bias, number of 2° cells		X	X
Scatter diagram of binned anomaly differences for a sensor pair versus number of observations in 2° cells for either sensor			X

deviation of 1.48°C and an average bias of +0.28°C relative to the ships. Allowing for buoy temperature sensor stabilization reduced the standard deviation and bias to 1.15°C and +0.75°C when compared to intake temperatures and 0.56°C and +0.18°C, respectively, when compared to bucket measurements. It should be noted that these statistics were determined for buoy SST's taken 24 hours after the ship recording.

3. Analysis Techniques

Vastly different spatial and temporal sampling characteristics as well as a low signal-to-noise ratio characterized the data. Hence analysis procedures were designed to reduce the noise by forming monthly 2° latitude by 2° longitude average SST anomaly fields. Noise levels were determined from point-to-point or "spot" comparisons of SST anomalies. An anomaly is defined as a departure of absolute SST from climatology. Anomalies were computed by linearly interpolating in space and time a 1° × 1° monthly climatology generated by *Reynolds* [1982] to the irregularly spaced satellite or surface point and subtracting the climatology from the measured temperature T. The resultant value, hereafter referred to as a "raw" anomaly ΔT, was used as the fundamental signal rather than absolute SST. Thus a picture of ocean variability was depicted by each sensor.

A variety of statistical and display routines, summarized in Table 2, were used to portray raw anomaly quantitative results and spatial features. First and second moments were computed in the usual Gaussian sense

$$\Delta\bar{T} = \frac{1}{n}\sum_{i=1}^{n} \Delta T_i \tag{1}$$

and

$$\Delta T_{\text{rmsd}} = \left[\frac{1}{n}\sum_{i=1}^{n}(\Delta T_i - \Delta\bar{T})^2\right]^{1/2} \tag{2}$$

where n is the total number of points, ΔT_i is the raw anomaly located at latitude y and longitude x at time t, and ΔT_{rmsd} is the root mean square deviation about the mean.

Monthly average fields were formed from raw anomalies by averaging all points that fell within a 2° by 2° cell centered on odd latitudes and longitudes. The average temperature or "binned" anomaly for latitude j and longitude i is simply expressed as

$$\bar{T}_{ji} = \frac{1}{m}\sum_{k=1}^{m} T_{yx} \tag{3}$$

where m is the total number of points in cell j, i. T_{yx} is the kth raw anomaly located at latitude y and longitude x within the cell centered at j, i.

Statistics of the binning process were retained for the purpose of comparing sampling characteristics, temperature extremes, and data dispersion within each cell. Raw anomalies exceeding ±5.75°C were eliminated before binned anomaly fields were formed because the natural variability of the ocean is typically much less than this magnitude. It follows that any signal of this intensity is likely the result of poor sensor performance or algorithm deficiencies, except perhaps in the case of a strong El Niño. No further quantitative editing was performed on SST anomalies. However, the data were stratified into latitude/longitude bands and were segregated on the basis of a quantitative interpretation of the status flag associated with each observation. Table 2 summarizes the screening pro-

TABLE 3. Data Analysis Products Spatial and Temporal Length Scales

Analysis Product	Global*§	North Pacific†	South Pacific‡	North Atlantic¶	Monthly Average
Anomaly histogram	X	X	X	X	X
Absolute SST contour map	X	X	X	X	X
Thematic map of binned anomalies	X	X	X	X	X
Thematic map of anomaly differences	X				X
Thematic map of number density	X				X
Scatter diagram of anomaly SST vs climatological SST	X	X	X	X	X
Cross-correlation table	X	X	X	X	X
Error partition table	X	X	X	X	X
Scatter diagram sensor vs. sensor		X	X	X	Spot merge ±6, ±12 hr; 20, 100 km
Scatter diagram anomaly diff. vs num. observ. in 2° cell		X	X	X	X

*Global study area: 60°S to 60°N, 0°–360°E.
†North Pacific study area; 0°–60°N, 100°–290°E.
‡South Pacific study area; 60°S to 0°, 100°–290°E.
¶North Atlantic study area; 0°–60°N, 290°–360°E.
§Global and regional thematic maps and scatter diagrams within 20° latitude bands extended to 60° latitude. All other products terminated at 55° latitude to eliminate spurious points due to sea ice.

cedures applied to each data set. Field data statistics were calculated as in (1) and (2) by substituting the mean anomalies for the raw values.

Sensor and algorithm performance were measured relative to climatology and each sensor. Statistics of the relationships quantified the bias, standard deviation, and correlation. The correlation between any two sensors (climatology was treated as a sensor) is given by

$$R_{12} = \frac{\sum_{n=1}^{N} \bar{T}_{1n}\bar{T}_{2n} - \frac{1}{N}\sum_{n=1}^{N}\bar{T}_{1n}\sum_{m=1}^{N}\bar{T}_{2m}}{\left[\sum_{n=1}^{N}\left(\bar{T}_{1n} - \frac{1}{N}\sum_{m=1}^{N}\bar{T}_{1m}\right)^2\right]^{1/2}\left[\sum_{n=1}^{N}\left(\bar{T}_{2n} - \frac{1}{N}\sum_{m=1}^{N}\bar{T}_{2m}\right)^2\right]^{1/2}} \quad (4)$$

where subscripts 1 and 2 refer to sensor pairs and subscripts n and m denote the cells common to both sensors. It is important to note that the mean and standard (rms) deviation of bivariate plots of SST binned anomalies against climatology were computed about the mean of the sensor data as in (1) and (2). Statistics displayed on correlation tables and scatter diagrams of raw anomalies were calculated about the mean bias, defined as the mean difference between common cells or, in the case of raw anomalies, matched points.

Another information extraction technique, known as "error partitioning," was used to ascribe measurement error to each sensor. This method employs sensor triplets in a set of three simultaneous equations that can be solved for each sensor's contribution to the total error in an rms sense. The mean square difference between binned anomalies from two sensors is expressed as

$$D_{12} = \langle(\bar{T}_1 - \bar{T}_2)^2\rangle = \frac{1}{N}\sum_{n=1}^{N}(\bar{T}_{1n} - \bar{T}_{2n})^2 \quad (5)$$

[see *NASA/JPL*, 1983, Appendix G] where $\langle\rangle$ indicates the sample mean computed over all binned cells n that are common to the sensor triplet, and $\bar{T}_1$ and $\bar{T}_2$ are the binned SST anomalies. Suppose the error in the measured anomaly is

$$\varepsilon_k = \bar{T}_k - \bar{T}_t \quad (6)$$

where $\bar{T}_k$ is the measurement from sensor k and $\bar{T}_t$ is the true anomaly. Then it is possible to express (5) as

$$D_{12} = \langle(\varepsilon_1 - \varepsilon_2)^2\rangle = \langle\varepsilon_1{}^2\rangle + \langle\varepsilon_2{}^2\rangle \quad (7)$$

It has been assumed that sensor errors are uncorrelated so that the mean cross product of the errors is zero. Similar expressions may be derived for D_{13} and D_{23}, thus forming a set of three equations that can be solved for the rms sensor errors: $[\langle\varepsilon_k{}^2\rangle]^{1/2}$, $k = 1, 2, 3$. In this manner an estimate of the error attributed to each sensor in a triplet with any other two sensors can be made. Thus for M sensors there will be $(M^2 - 3M + 2)/2$ error estimates for a given sensor in combination with pairs of all other sensors. By averaging the errors, an estimate of the overall error can be partitioned to each sensor.

The analytical techniques were applied to binned anomaly fields and raw anomalies. Information was derived from field data on global and regional scales and raw data in the form of "spot" comparisons on spatial and temporal scales commensurate with sensor sampling and geophysical variability. The space-time scales used for preparing analysis products are presented in Table 3. This comparison shows the large-scale summaries prepared from sensors providing global coverage. In contrast, sensors that operated in limited regions were merged by using a number of different space-time scales or "windows"

TABLE 4. Time-Space Windows Used for Merging

Sensor	Time, ±hours	Radial Distance, km
AVHRR	12	100
Ships	6	100
XBT	12	20, 100
FGGE buoys	12	20

Scatter diagrams were prepared for sensors matched to the table entry.

for the purpose of examining the errors in spot comparisons. This merging process matched pairs of points that were within a given time tolerance (plus or minus hours) and radial distance (kilometers). Table 4 provides the specifications used to merge sensors with the table entries.

4. Data Products

The data products prepared for the workshop series are reviewed in *NASA/JPL* [1983, 1984, 1985]. A complete collection of materials spanning four data months and eight sensors is too large to present here. However, for discussion purposes a representative set of products has been compiled in the appendix. Data reduction results, discussed in detail in the companion papers, are also presented in the appendix.

For the purpose of examining the distribution of raw anomalies, histograms similar to Figure 1 were generated for all sensors. In this example the AVHRR provided an unbiased estimate of SST with few outliers, indicating that surface temperatures for December 1981 were not noisy. It should be noted that the 200,000 points represented on this plot are from the first week in December. Successive weeks have similar statistics, with the result that binned fields were smooth. Histograms of the raw data provided a quantitative measure of extreme values because points exceeding ±5.75°C were eliminated before binning.

Color maps displaying the number of raw anomaly observations falling within a 2° cell are illustrated in Plate 1. The blue shades indicate fewer AVHRR points were retrieved in historically cloudy zones [*Hughes*, 1984] along the equator, southwestern Pacific, southeastern Atlantic, and the Pacific southward of 40° latitude. In the less cloudy regions the high spatial sampling frequency of AVHRR typically yielded more than 300 samples per cell for any month, whereas the lower-resolution satellite and in situ sensors accumulated on the order of 10 samples per cell. In some cases there was a total absence of retrievals during the month as a result of clouds, data editing or, for in situ platforms, no coverage. Areas that appear black represent cells for which there were no SST retrievals. The color legend has been assigned a different range of magnitudes for each sensor to make the data density variability apparent.

Monthly mean SST anomalies were portrayed in color maps for each study month. In Plates 2–5 anomalies colder than climatology appear as blue shades and warmer anomalies as gray to red shades. This color convention applied to all maps of SST anomalies and anomaly differences. Difference maps shown in Plates 6–9 were formed by subtracting ship anomalies from satellite anomalies for all cells common to both ship and satellite sensors. These anomaly difference maps, formed relative to ships, delineate the satellite–in situ sensor biases but also inherently retain ship biases. Thus to decouple ship errors, satellite-satellite biases were evaluated with the differencing technique by using the sensors previously differenced with ships. Color maps displaying the sensor biases have been assembled for the four study months in Plates 10–13.

Two more types of color maps were produced to illustrate features specific to particular sensors. Since the VAS coverage

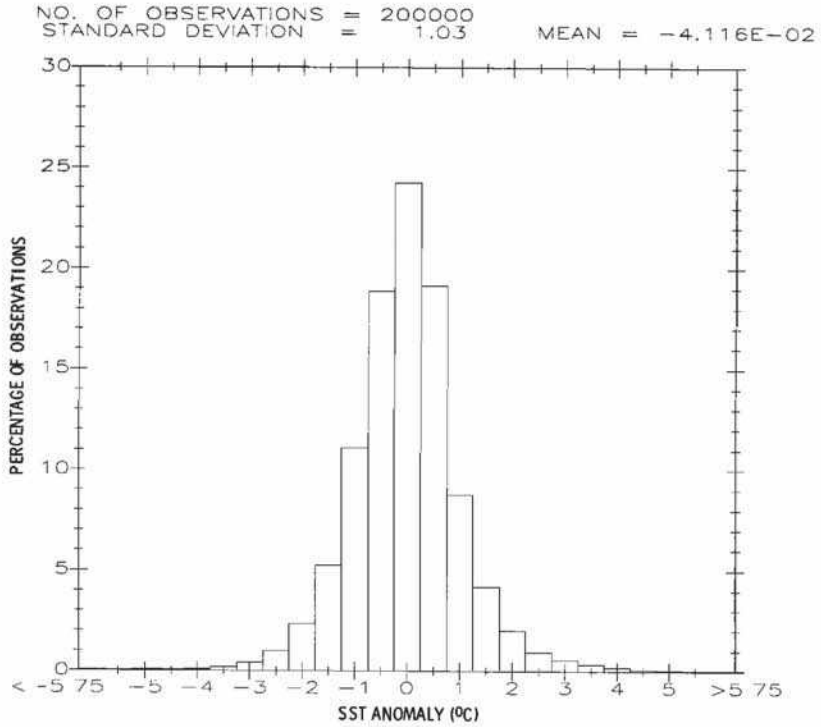

Fig. 1. Histogram of raw SST anomalies (SST) for NOAA AVHRR, December 1981, 55°S to 55°N, 0°–360°E.

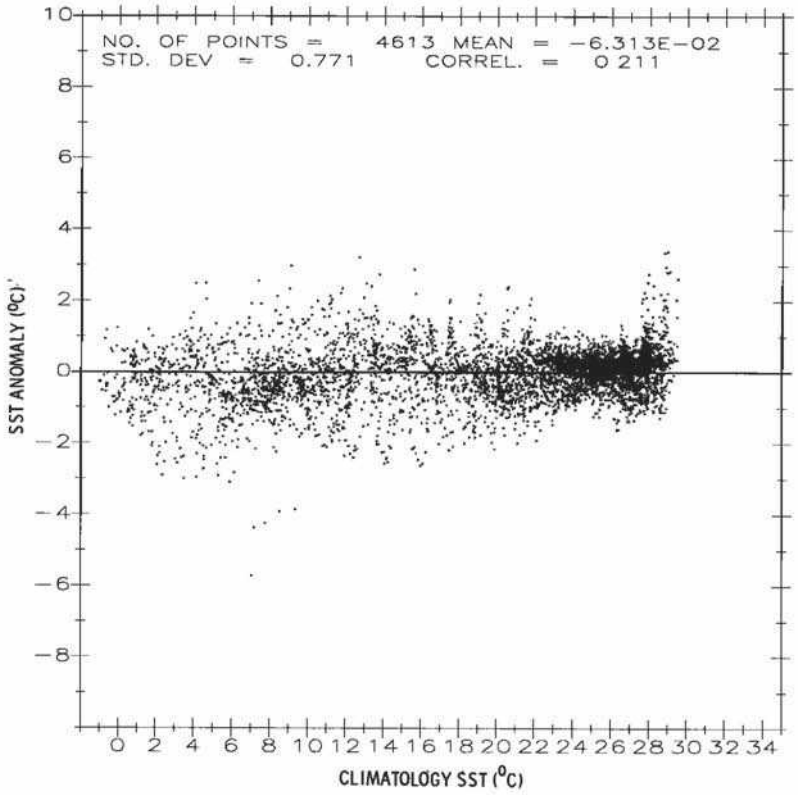

Fig. 2. Scatterplot of binned anomalies versus binned climatology for NOAA AVHRR, December 1981, 55°S to 55°N, 0°–360°E.

was limited to the tropics near longitude 90°W, special cutouts were made to compare all sensors in this region. Plate 14 shows the comparisons for March and July 1982. In the case of the HIRS, data were provided to the workshop with specified data quality weights at each SST retrieval point. Maps could then be made with or without these weights applied in the binning process. Plate 15 shows weighted and unweighted maps supplied by J. Susskind for March 1982 [see *Susskind and Reuter*, this issue] and a weighted map for July 1982 produced by the workshop (cf. Plate 5, unweighted).

Global instrument biases were quantified on bivariate plots of binned SST anomalies versus binned climatology and illustrated in Figures 2 and 3. The low (0.21) correlation with climatology is an indication that the AVHRR retrievals represented the change in ocean temperature independent of climatology, i.e., the bias and standard deviation of the measured anomalies do not seem to be highly dependent on the ocean temperature. This can be examined more closely using the plots of Figure 3. This diagram displays the data in Figure 2 stratified into 20° latitude bands with the statistics computed about the zonal mean. The AVHRR statistics shown suggest that a small temperature-dependent bias exists. However, similar results from HIRS and ships show good agreement with the AVHRR. The maximum positive anomalies were reported in the 0°–20°N band by all three sensors.

In contrast to the large-scale field summaries, raw (not averaged) SST anomalies were compared on scales less than 2° × 2° and 1 month. Points were sampled nearly coincidently in space and time in order to measure the relative accuracy and correlation between sensors. Figure 4 shows raw anomalies derived from AVHRR matched to within ±12 hours and 100-km radius of XBT points collected in the North Pacific. Here the AVHRR exhibits a small negative bias but shows good agreement with XBT's and a scatter of 1°C about the bias, consistent with the results obtained by *McClain* [1983]. Similar regional comparisons were made for the South Pacific and North Atlantic. Additionally, bivariate comparisons of binned anomalies were made on monthly regional scales. This was necessary because of the vastly different resolution sizes which varied from spots for ships and XBT's to 50 km for AVHRR and 150 km for SMMR.

Bivariate comparisons proved useful for examining relative sensor performance on a pair-by-pair basis. However, to evaluate the set of all sensors, cross-correlation calculations were made by using equation (4). Global correlation analysis results for December 1981 are presented in Table 5. The results generally show weak correlations typical of low signal to noise ratios, suggesting that the sea surface temperatures differed little from the climatology. Exceptions are the moderate correlations between ships and AVHRR, and XBT's and AVHRR, of 0.61 and 0.70, respectively. These results are more indicative of the northern latitudes because the majority of ship and XBT measurements were confined to the North Atlantic and North Pacific. Furthermore, the AVHRR accuracy, as measured by the standard deviation, is approximately ±0.7°C. The table entries were computed over the number of points common to both sensors. A significant result of workshop 1 was the realization that, even after screening, ship observations tended to be noisy because many cells were sampled only a few times per month. Consequently, error caused

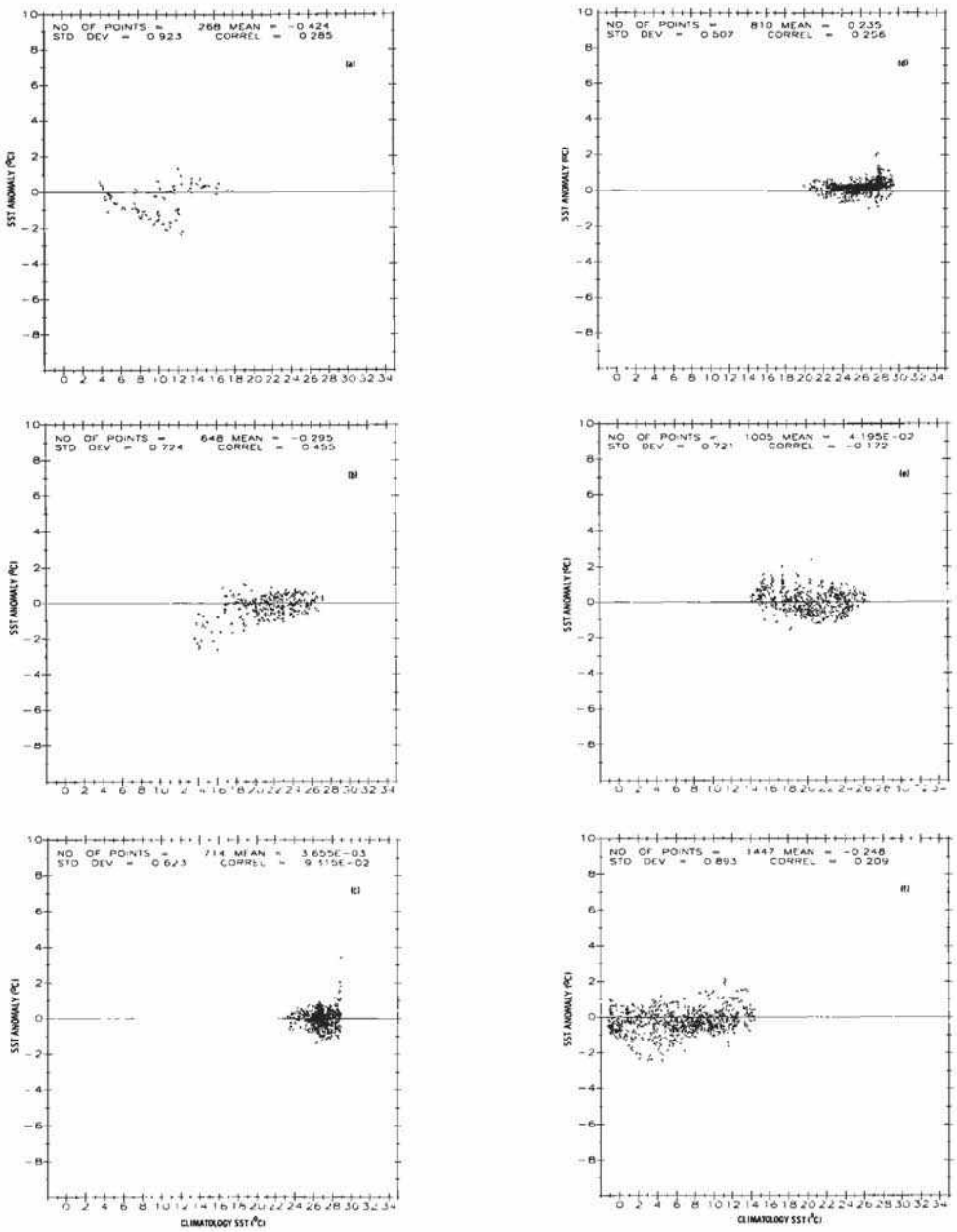

Fig 3. Scatterplot of binned anomalies versus binned climatology stratified by latitude zone: (*a*) 60°N to 40°N, (*b*) 40°N to 20°N, (*c*) 20°N to 0°, (*d*) 0° to 20°S, (*e*) 20°S to 40°S, (*f*) 40°S to 60°S. Data shown are for NOAA AVHRR, December 1981.

by undersampling was reduced in the correlation analysis by using only those ship cells with more than five observations.

Estimates of the rms sensor errors were made with the error partitioning technique developed in equations (5)–(7). It is important to recall that this method assumes that the errors are uncorrelated, which may not be strictly true in some cases. For example, cloud elimination techniques may cause errors to be correlated, or alternatively, effects resulting from instrument temperature changes may not have been eliminated from the algorithms. The latter could be true for two sensors located on the same satellite that experience similar heating-cooling cycles over an orbit. Selected results for December 1981 are presented in Table 6. The partitioned results in this example are between AVHRR, SMMR, HIRS, and climatology, where climatology consists of zeros. An anomaly data set consisting of zeros represents a sensor that measured SST equivalent to climatology.

In Table 6 the upper right triangle contains the rms error for the triad and the lower left triangle contains the number of common cells used in the computation. Thus an rms error of 0.60°C can be allocated to AVHRR in combination with SMMR and HIRS. A table was generated for each sensor, and the overall average rms error was determined for each sensor. For December 1981 the overall rms errors were 0.60, 0.98, 0.68, and 0.43°C for AVHRR, SMMR, HIRS, and climatology, respectively. A significant result is that the AVHRR and ships were very close to the true global rms SST anomaly (quantified by the climatology "error"). Although SMMR error is of

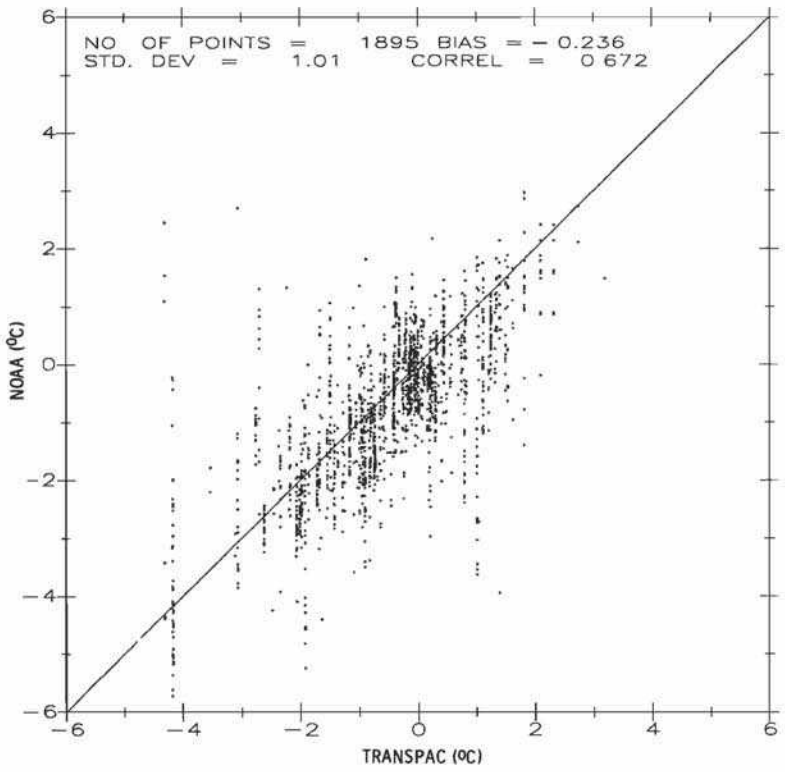

Fig. 4. NOAA AVHRR SST raw anomalies versus Transpac XBT raw anomalies. Matched within a ±12-hour, 100-km window for December 1981, 0°–55°N, 100°–290°E.

the order of 1.00°C, this should be sufficient accuracy to measure 1°–2°C regional anomalies and may be the only sensor that can determine SST precisely in cloudy areas.

5. Summary and Conclusions

The SST workshop series provided the first opportunity to compare the performance of several satellite and surface sensors. To evaluate sea surface temperature retrievals during a variety of ocean conditions, the months of November 1979, December 1981, March 1982, and July 1982 were selected for study. Workshop 1 was devoted to the analysis of November 1979 data and a review of the analysis techniques. December 1981 data were examined in workshop 2, and March and July 1982 SST's were the subject of workshop 3. This paper presented the analytical techiques used to reduce the data and quantify the accuracy of the sensors. The methodology used to evaluate the data was developed in response to early analysis in workshop 1. Hence the same set of procedures was applied to all data months. Because the geophysical signal was small compared to the absolute ocean temperature, climatology was first subtracted from the raw SST retrievals. It addition the sampling scales of each sensor were very different, so it was necessary to develop a scheme that converted the data to uniform space/time scales. Moreover, to eliminate data noise without excessively reducing the natural ocean signature, the raw anomalies were averaged over a period of 1 month and a 2° latitude by 2° longitude box.

Maps displaying the data density distribution, sea surface temperature anomalies, and anomaly differences proved extremely useful for determining the qualitative agreement of the sensors. Cloudy regions and areas where sensors systematically performed poorly because of clouds or calibration problems were clearly visible in the color maps.

Quantitative comparison of the field or so-called binned data showed that the accuracies varied from sensor to sensor, depending on the ocean region and data collection period. Typically, the standard deviation and bias of satellite measure-

TABLE 5. Two-Degree Average SST Cross Correlations for the Global Oceans; December 1981, 55°S to 55°N, 0°E to 360°E

		HIRS /MSU	SMMR Night	Ships, >5/cell	TRANSPAC XBT	Climatology
NOAA AVHRR	*C*	0.32	0.20	0.61	0.70	0.00
	B	0.00	0.09	0.25	0.27	0.00
	S	1.09	1.16	0.70	0.75	0.83
	N	7379	3864	1681	366	7410
HIRS /MSU	*C*		0.19	0.19	0.20	0.00
	B		0.24	−0.43	−0.16	0.00
	S		1.17	1.16	1.17	1.06
	N		3873	1705	366	7604
SMMR night	*C*			0.21	0.09	0.00
	B			−0.72	−0.80	−0.09
	S			1.17	1.41	1.06
	N			690	284	3874
Ships, >5/cell	*C*				0.57	0.00
	B				−0.03	−0.03
	S				0.91	0.71
	N				212	1727
TRANSPAC XBT	*C*					0.00
	B					0.23
	S					0.98
	N					366

C, cross correlation; *B*, bias (average over column temperatures minus row temperatures); *S*, standard deviation; *N*, number of common observations.

TABLE 6. Error Partitioning Results for AVHRR, December 1981

	AVHRR	SMMR	HIRS	Climatology
AVHRR	—	—	—	—
SMMR	—	—	0.60	0.62
HIRS	—	3780	—	0.59
Climatology	—	3781	4612	—

Upper right triangle contains rms error (°C) for triad, lower left triangle contains the number of points common to the sensors in the triad. Overall mean rms error, 0.60°C.

ments compared to ships and XBT's varied between 0.5–1.2°C and 0.2°–0.8°C, respectively. The standard deviation between two sensors includes errors from both sensors. It was therefore necessary to use a partitioning approach (see section 3) to separate errors. The overall average rms errors for AVHRR, SMMR, and HIRS for all four workshop months were about 0.5°C, 1.0°C, and 0.7°C, respectively.

Point-to-point comparisons proved valuable for comparing sensors that sample with nearly the same spot size (AVHRR, XBT, FGGE buoys). Ship data were shown to be too noisy on a spot basis, even after stringent screening. Therefore regional comparisons were primarily limited to the North Pacific.

The first-order analysis of SST data has been summarized in this paper. The workshop series provided a much needed assessment of the current state of the art of temperature retrieval but also provoked critical thought about the manner in which the comparisons were performed. Specifically, it is of utmost importance to understand the geographical dependence of the accuracy of satellite-derived temperatures. Moreover, the error dependence on the number of samples collected in a given area should be closely examined, as should the intracell temperature variance.

Acknowledgments. Support and implementation of the SST workshop processing were due to the efforts of the staff of the NASA/JPL Pilot Ocean Data System. Carol Miller, Jim Brown, Dan Bonbright, and Rich Norman developed the applications software. Brent Houghton and Frank Kuykendall were responsible for developing the image processing routines. Catherine Kerr, Ruby Lassanyi, Frank Salamone, John Burke, and Steve Carpenter helped create many of the data products. Funding support was provided by the Information Systems office of NASA's Office of Space Science and Applications.

REFERENCES

Barnett, T. P., Long-term trends in surface temperature over the oceans, *Mon. Weather Rev., 112*, 303–312, 1984.

Garrett, J., The performance of the FGGE drifting buoy system, *Adv. Space Res., 1*, 87–94, 1981.

Hughes, N. A., Global cloud climatologies: A historical review, *J. Clim. Appl. Meteorol., 23*, 724–751, 1984.

McClain, E. P., W. G. Pichel, C. C. Walton; Z. Ahmad, and J. Sutton, Multi-channel improvements to satellite derived global sea surface temperatures, *Adv. Space Res., 2*, 43–47, 1983.

NASA/JPL, Satellite-derived sea surface temperature: Workshop 1, *Publ. 84-5*, Jet Propul. Lab., Pasadena, Calif., 1984.

NASA/JPL, Satellite-derived sea surface temperature: Workshop 2, *Publ. 84-5*, Jet Propul. Lab., Pasadena, Calif., 1984

NASA/JPL, Satellite-derived sea surface temperature: Workshop 3, *Publ. 85-63*, Jet Propul. Lab., Pasadena, Calif., 1985.

Reynolds, R., A monthly averaged climatology of sea surface temperature, *Tech. Rep. NWS 31*, Nat. Oceanic Atmos. Admin., Silver Spring, Md., 1982.

Smith, W. L., and H. M. Woolf, Algorithms used to retrieve surface skin temperature and vertical temperature and moisture profiles from VISSR Atmospheric Sounder (VAS) radiance observations, paper presented at Fourth Conference on Atmospheric Radiation, Am. Meteorol. Soc., Toronto, June 16–18, 1981.

Saur, J. F. T., A study of the quality of sea water temperatures reported in logs of ships' weather observations, *J. Appl. Meteorol., 2*, 237–247, 1983.

Susskind, J., and D. Reuter, Retrieval of sea surface temperatures from HIRS2/MSU, *J. Geophys. Res.*, this issue.

Susskind, J., J. Rosenfield, D. Reuter, and M. T. Chahine, The GLAS physical inversion method for analysis of HIRS/MSU sounding data, *NASA Tech. Memo. 84936*, 1982.

Tabata, S., An evaluation of the quality of sea surface temperatures and salinities measured at station P and line P in the Northeast Pacific Ocean, *J. Geophys. Res., 87*, 374–385, 1982.

Wilheit, T. T., J. Greaves, J. Gatlin, D. Han, B. M. Krupp, A. S. Milman, and E. Chang, Retrieval of ocean surface parameters from the Scanning Multichannel Microwave Radiometer (SMMR) on the Nimbus-7 satellite, *IEEE Trans. Geosci. Remote Sensing, 15*, 225–244, 1983.

D. B. Chelton, College of Oceanography, Oregon State University, Corvallis, OR 97331.

J. E. Hilland, Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109.

E. Njoku, Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109, and Department of Engineering, Harvey Mudd College, Claremont, CA 91711.

(Received January 31, 1985;
accepted July 31, 1985.)

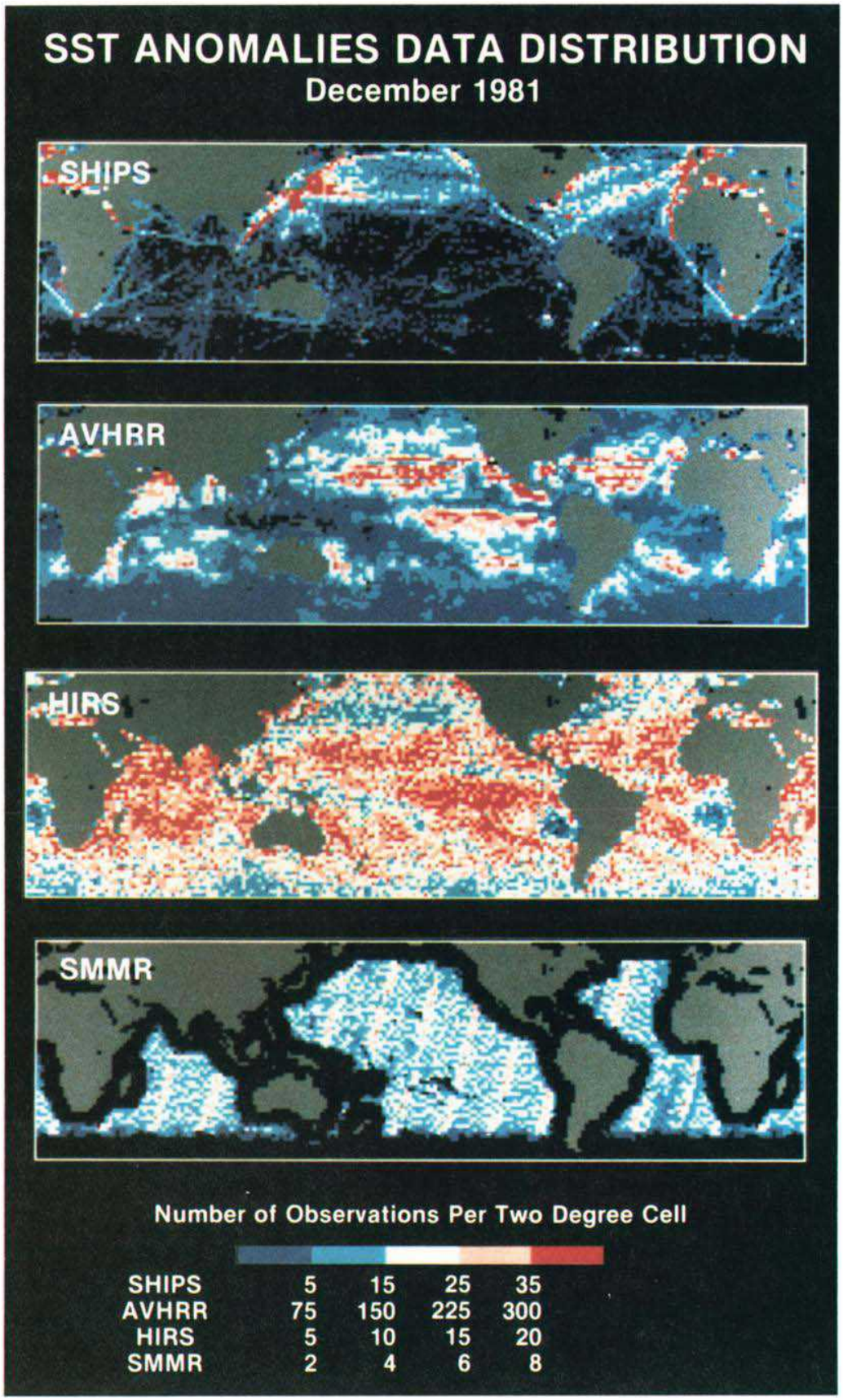

Plate 1. Thematic map of the number of raw SST measurements within each 2° latitude by 2° longitude cell. Mapped area: 60°S to 60°N, 0°–360°E (global), December 1981.

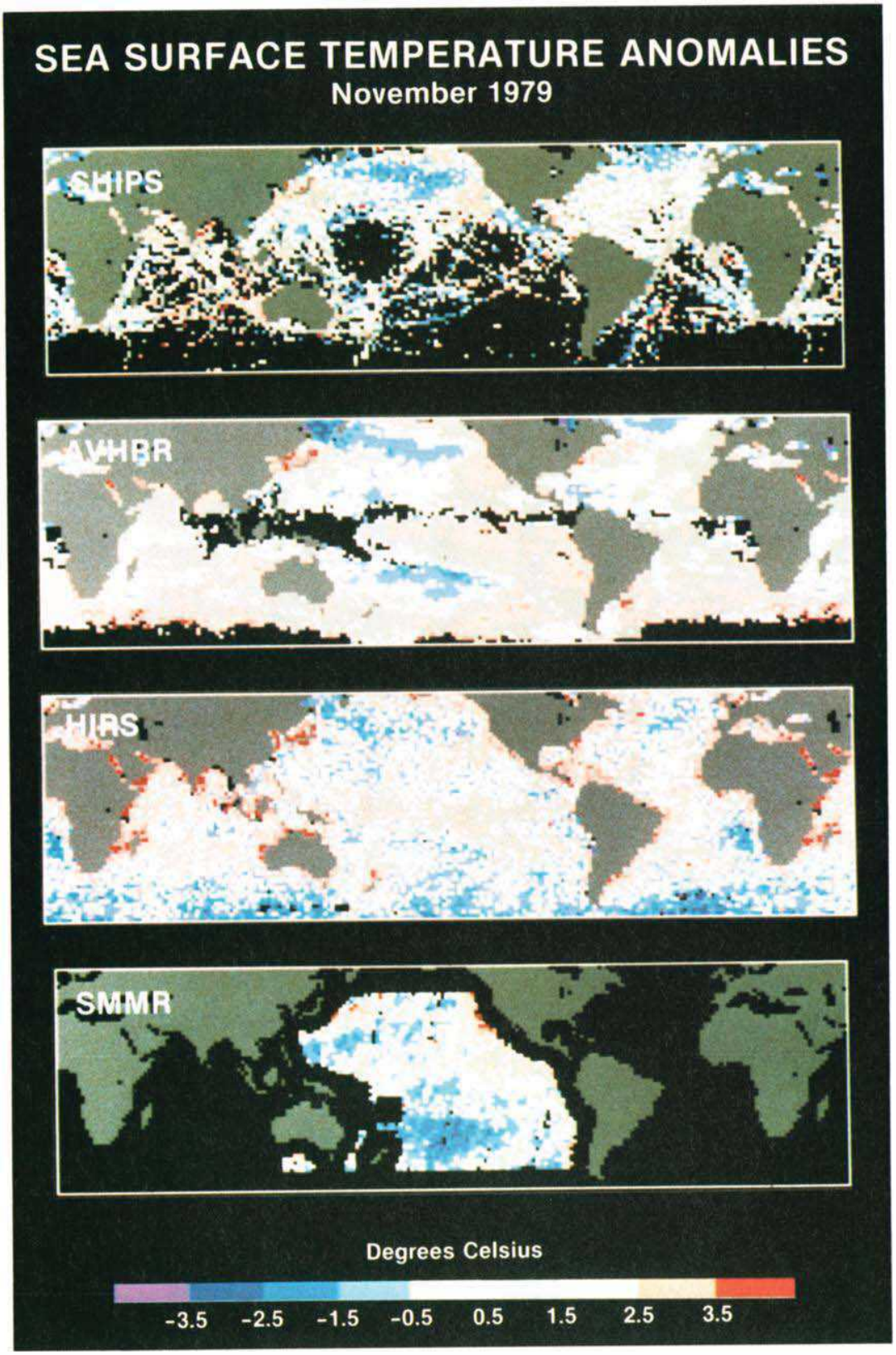

Plate 2. Thematic map showing binned SST anomalies for the global ocean, November 1979.

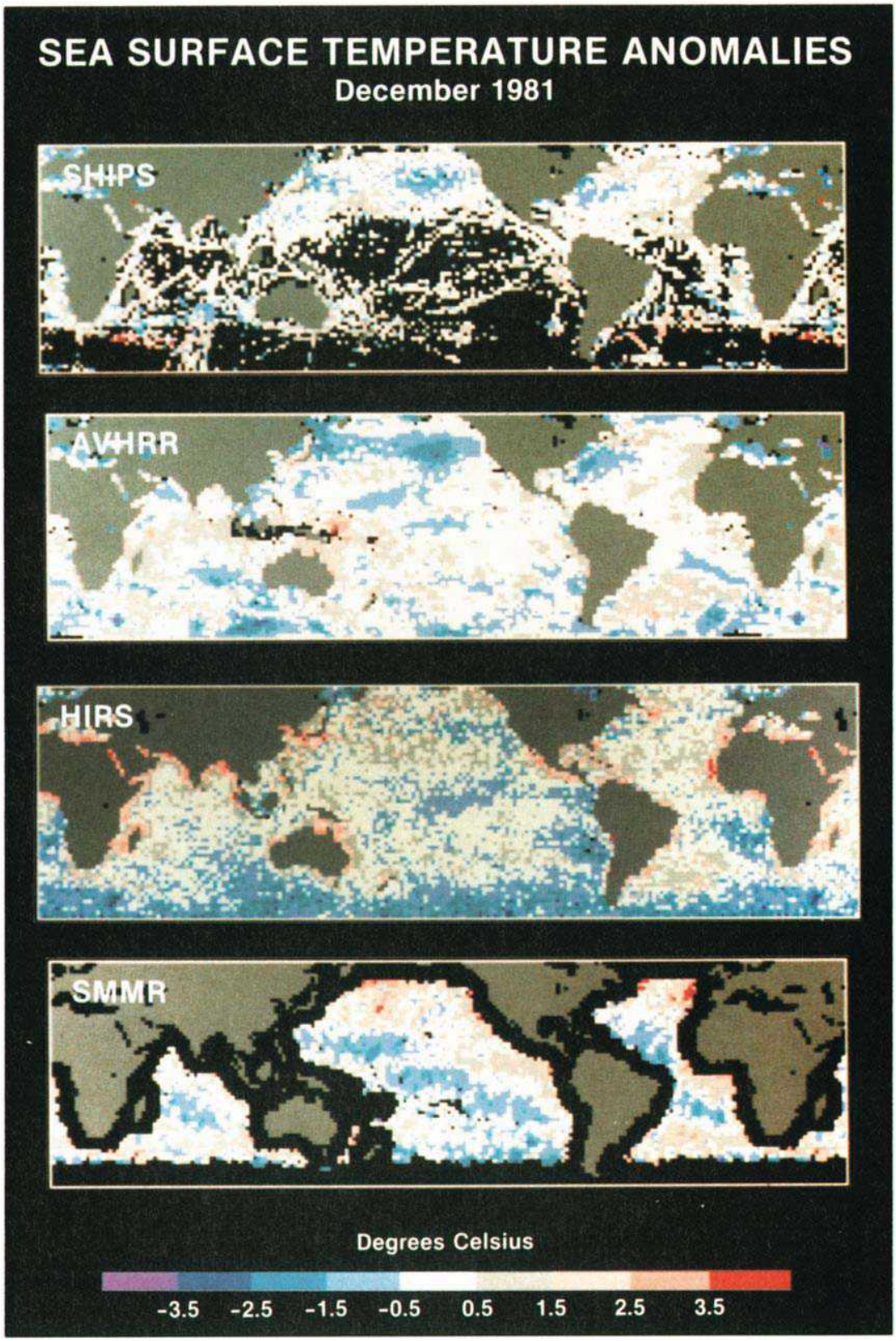

Plate 3. Thematic map showing binned SST anomalies for the global oceans, December 1981.

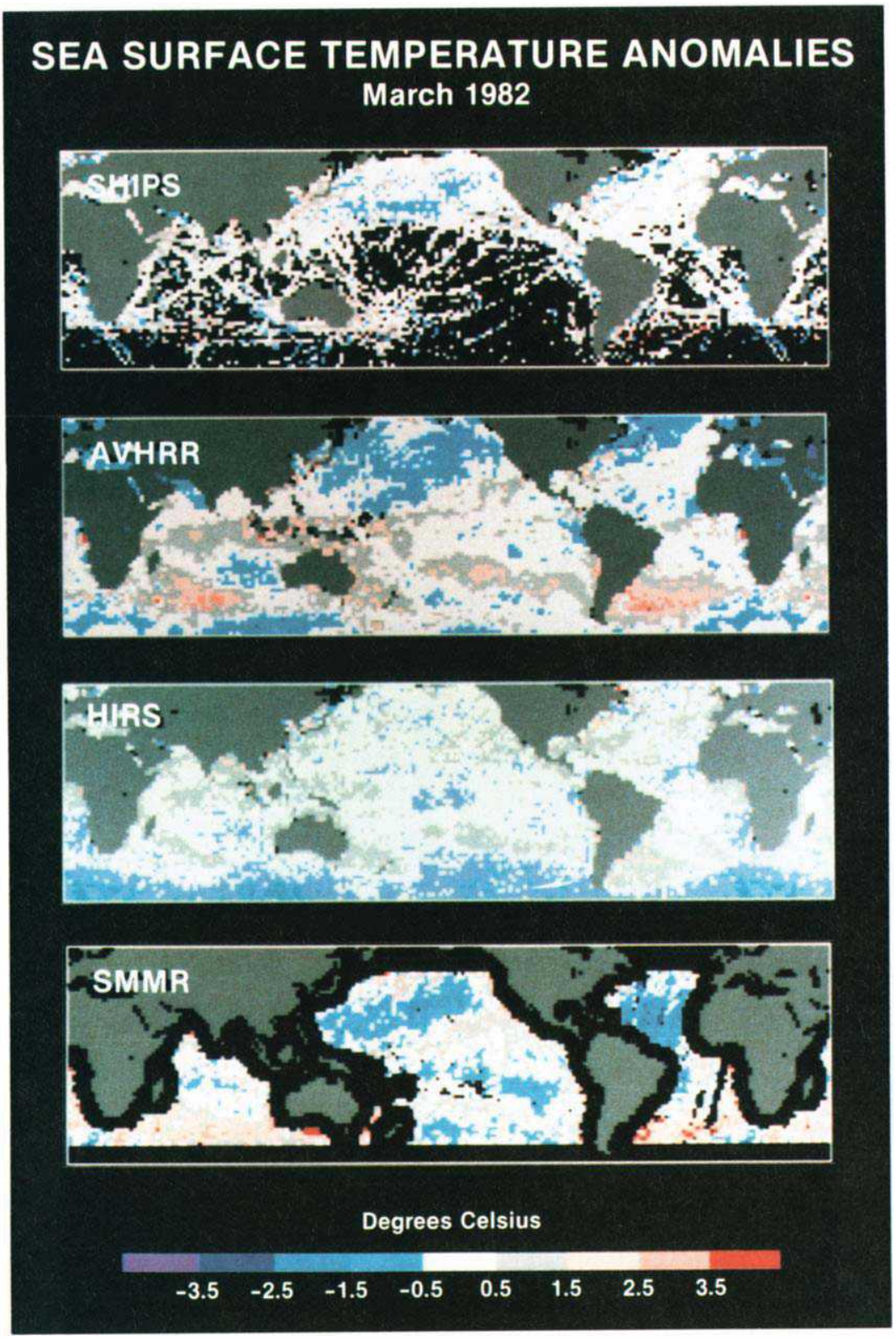

Plate 4. Thematic map showing binned SST anomalies for the global oceans, March 1982.

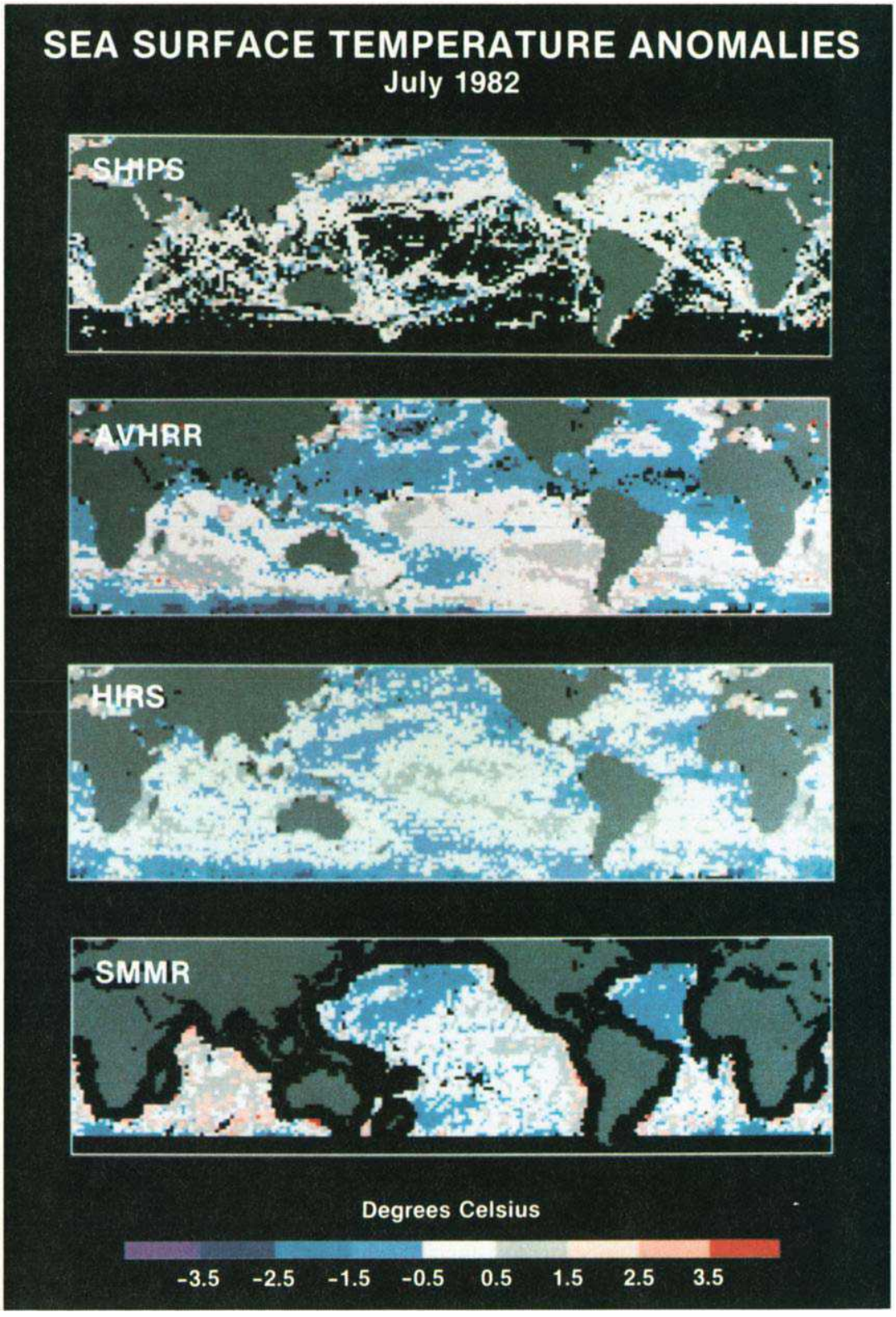

Plate 5. Thematic map showing binned SST anomalies for the global oceans, July 1982.

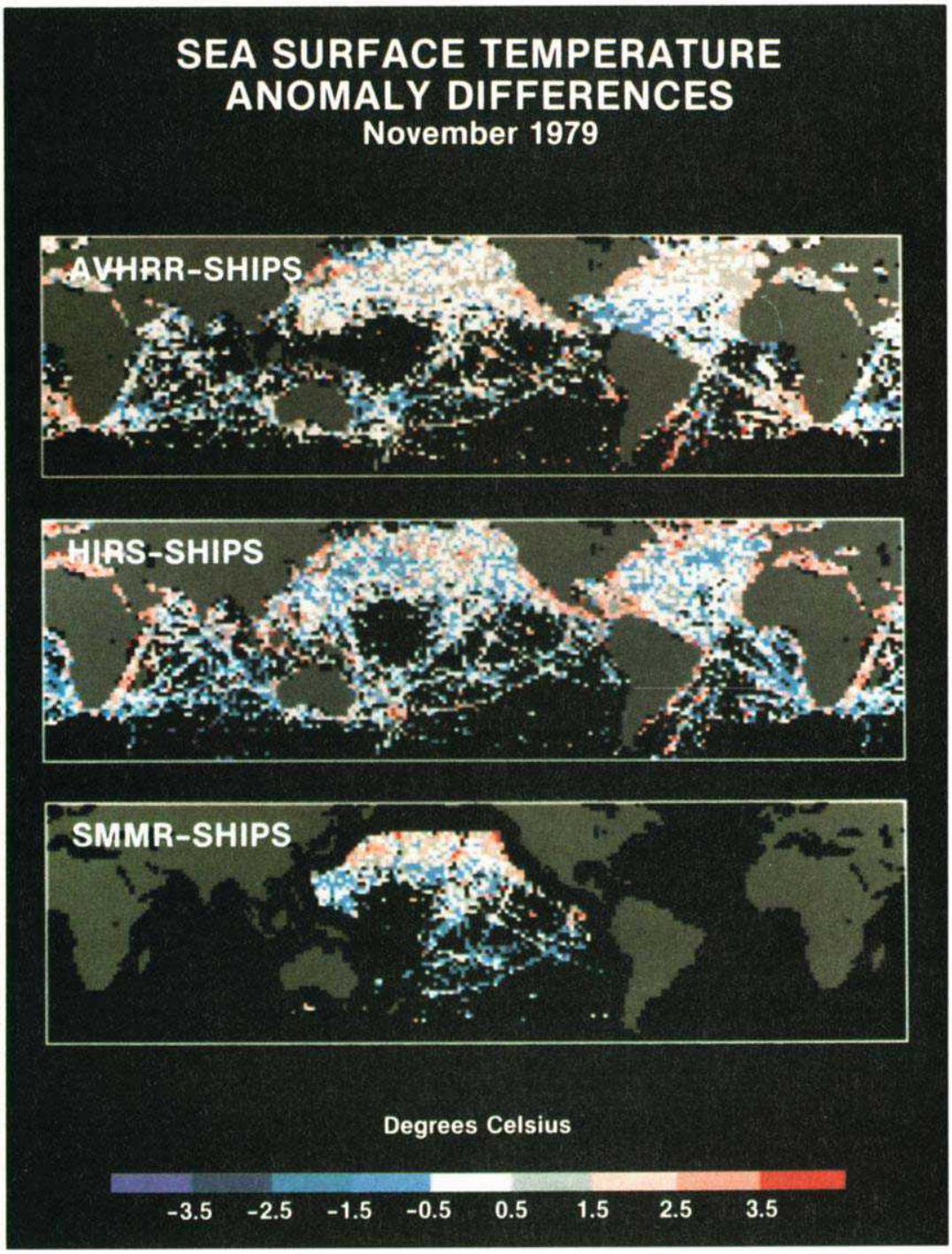

Plate 6. Thematic map of binned SST anomaly differences (sensor-ships), November 1979.

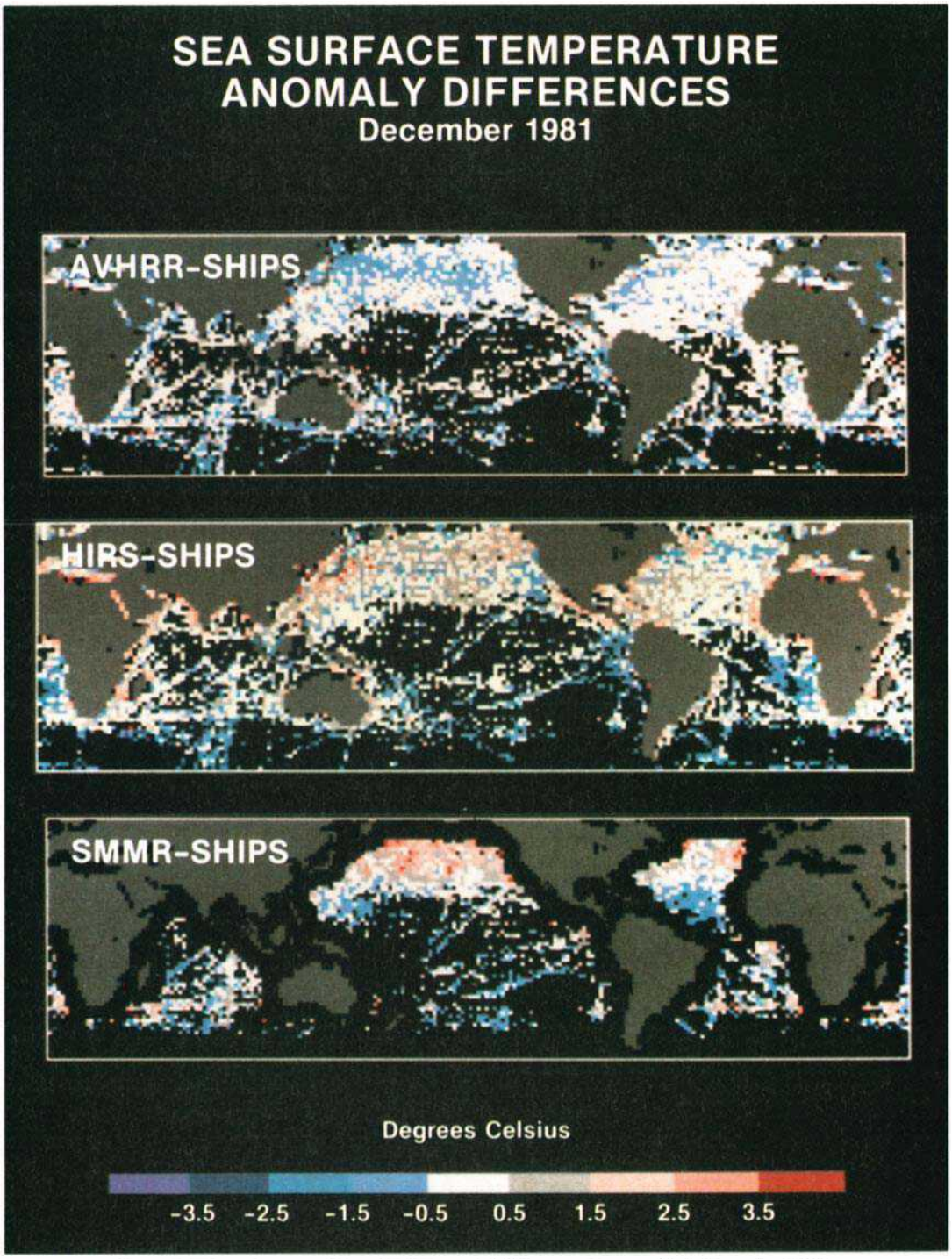

Plate 7. Thematic map of binned SST anomaly differences (sensor-ships), December 1981.

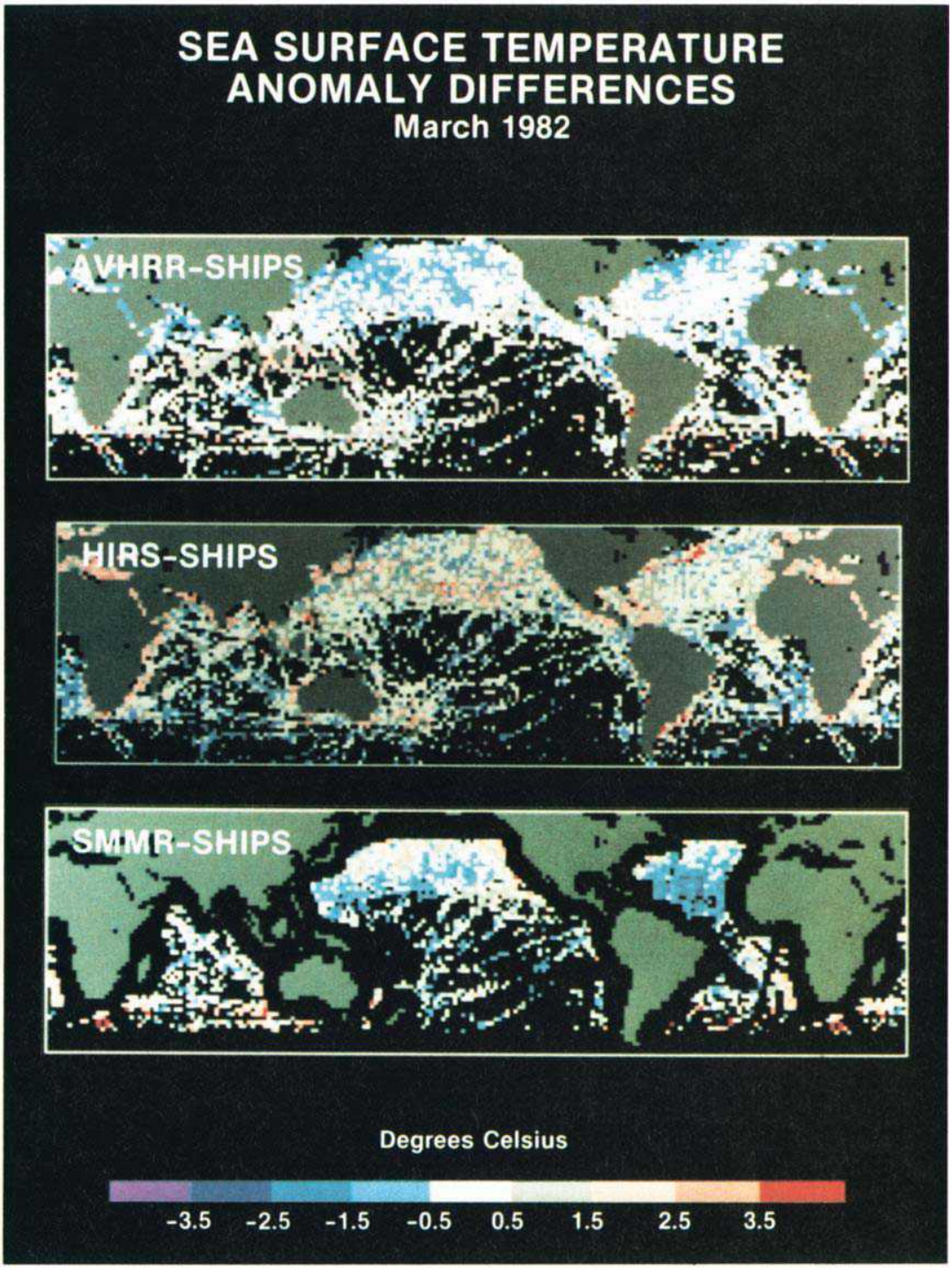

Plate 8. Thematic map of binned SST anomaly differences (sensor-ships), March 1982.

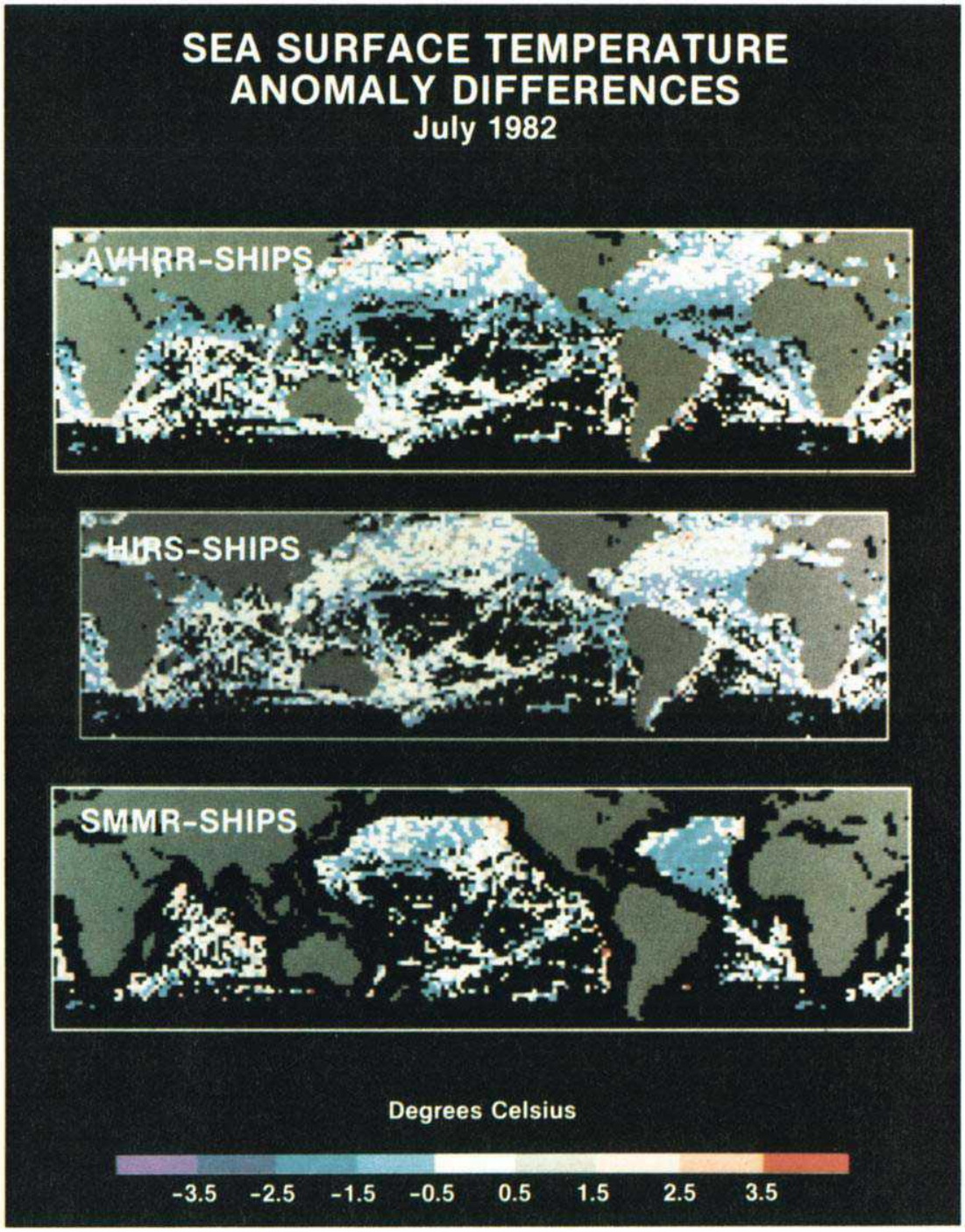

Plate 9. Thematic map of binned SST anomaly differences (sensor-ships), July 1982.

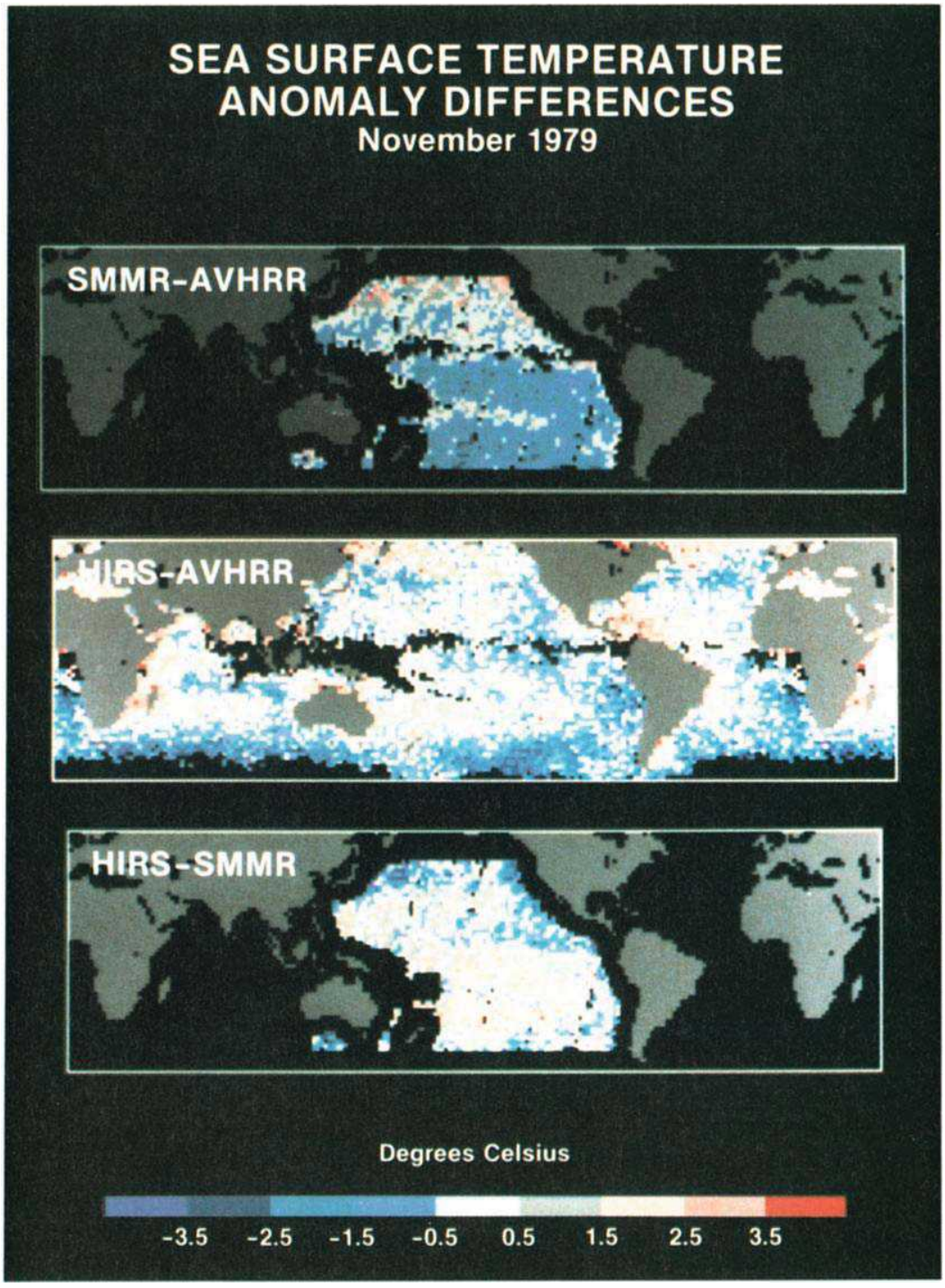

Plate 10. Thematic map of binned SST anomaly differences (sensor 1, sensor 2), November 1979.

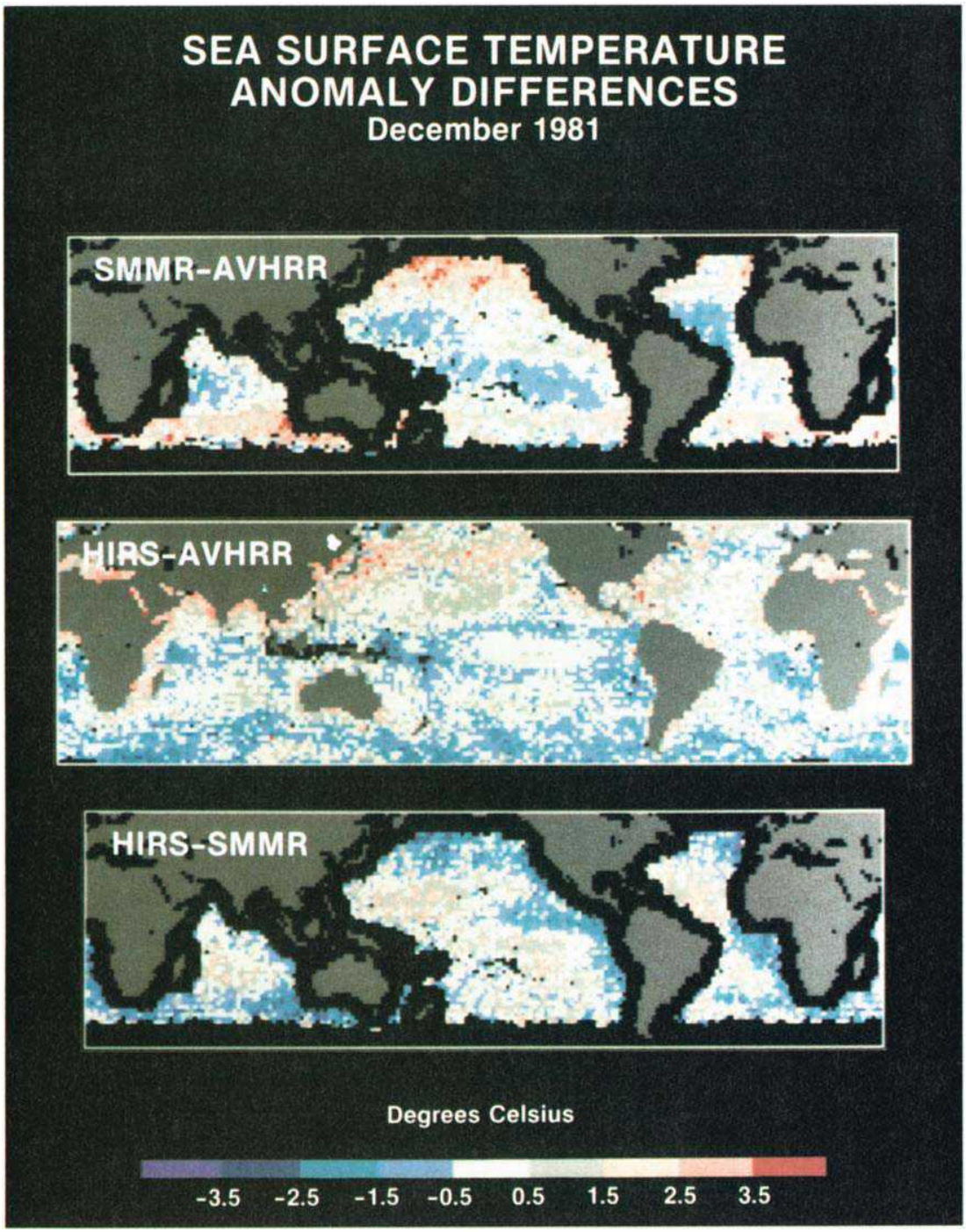

Plate 11. Thematic map of binned SST anomaly differences (sensor 1, sensor 2), December 1981.

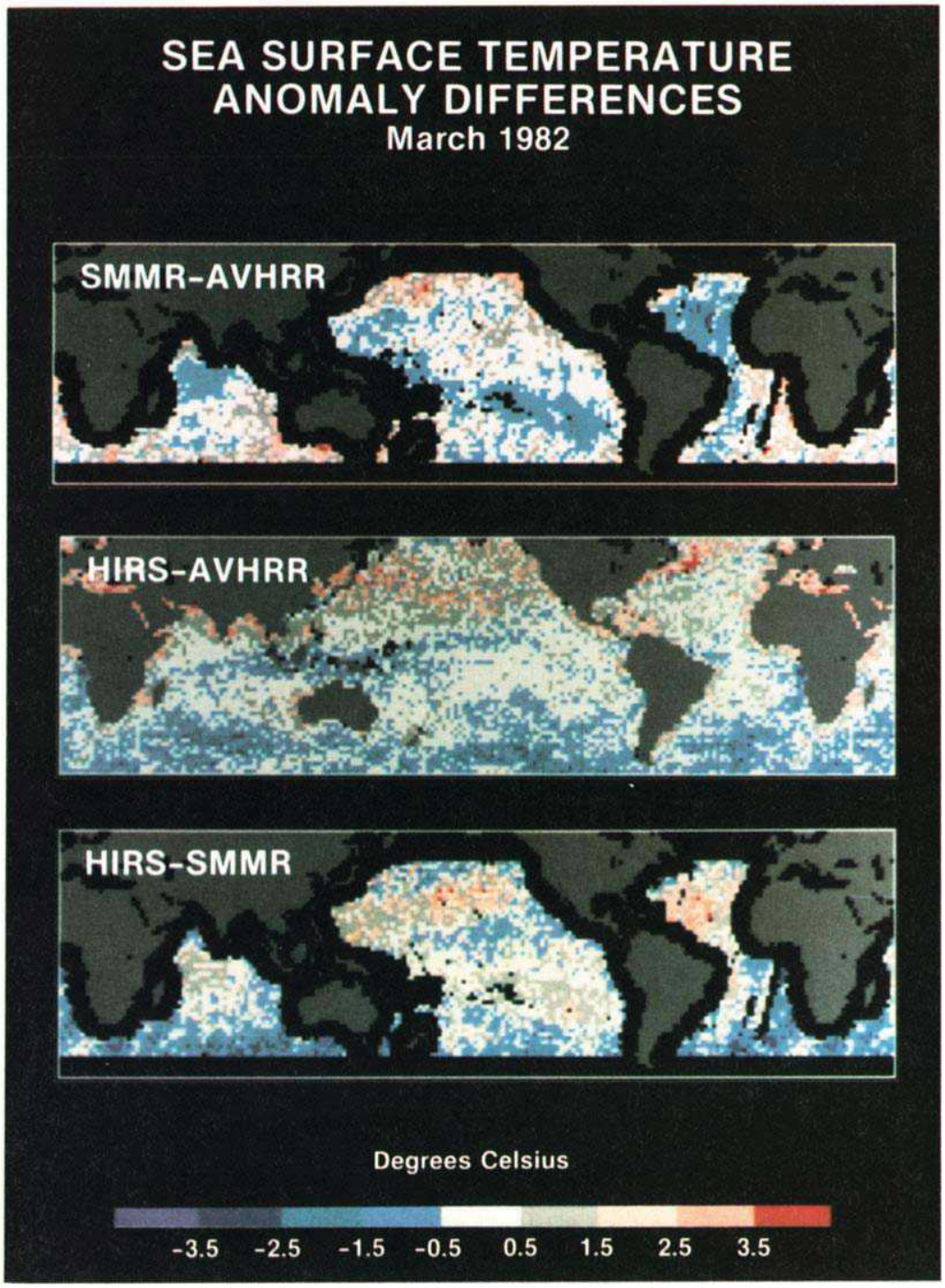

Plate 12. Thematic map of binned SST anomaly differences (sensor 1, sensor 2), March 1982.

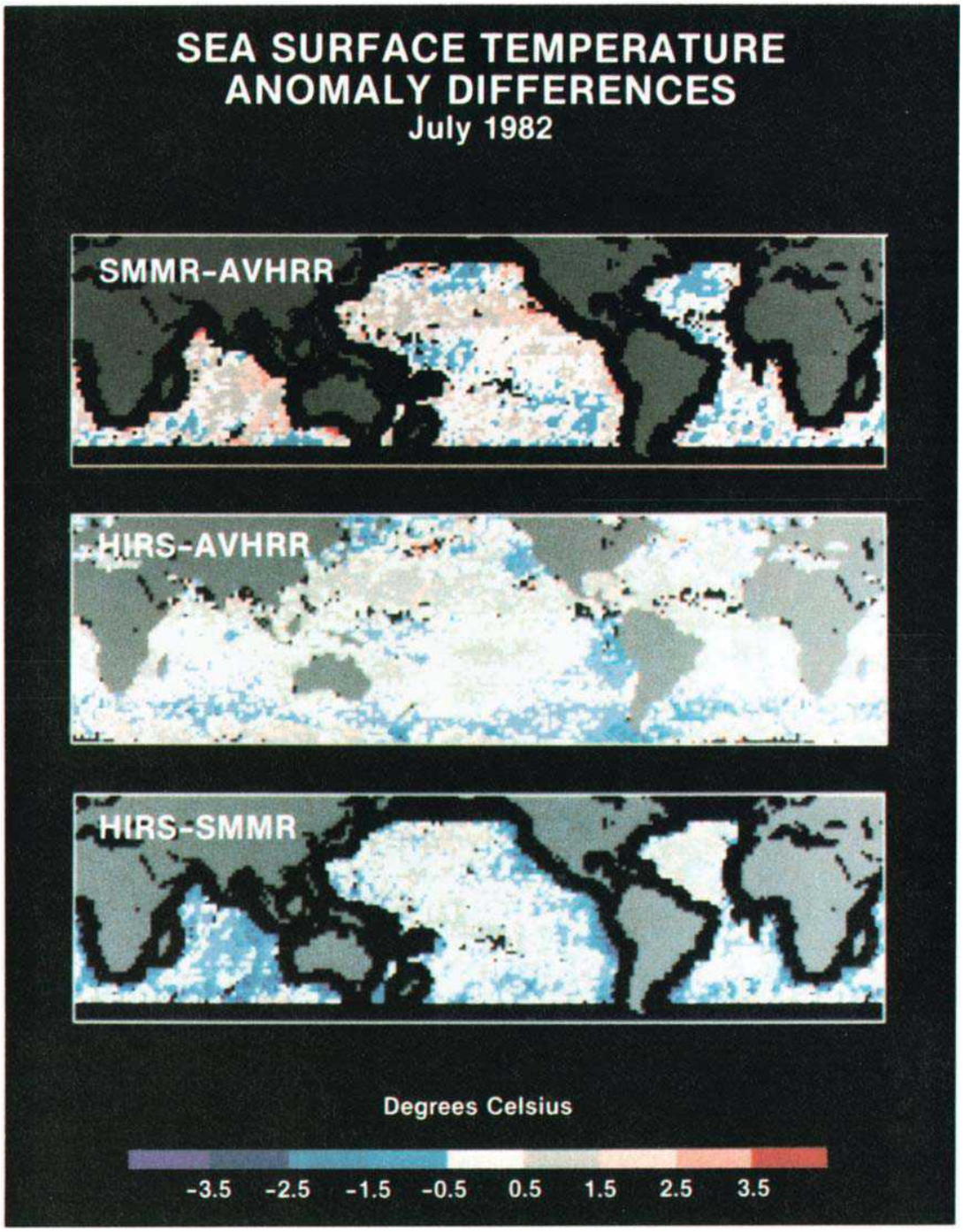

Plate 13. Thematic map of binned SST anomaly differences (sensor 1, sensor 2), July 1982.

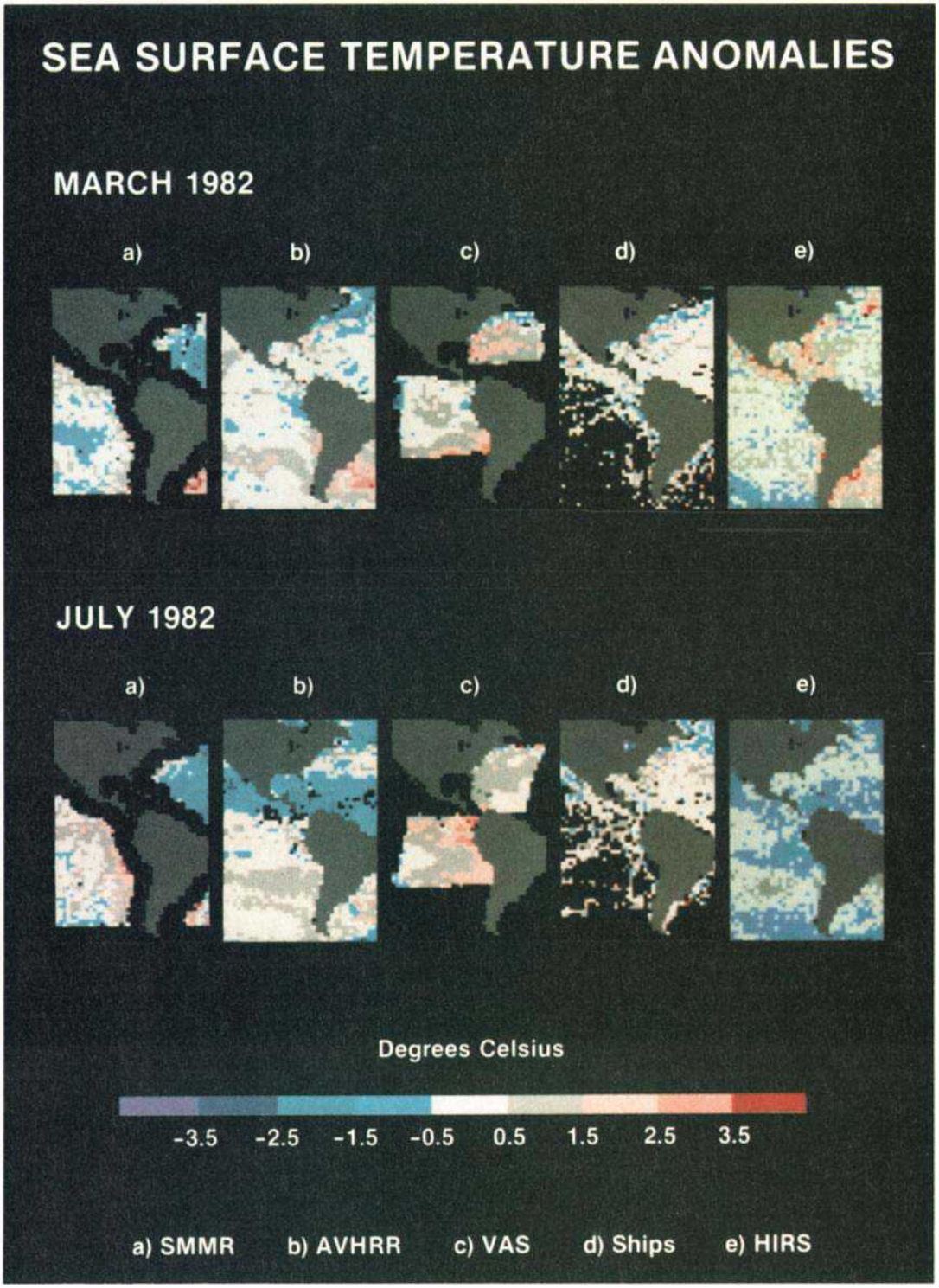

Plate 14. Sea surface temperature binned anomalies for the region of VAS coverage during March and July 1982: (*a*) SMMR, (*b*) AVHRR, (*c*) VAS, (*d*) ships, (*e*) HIRS.

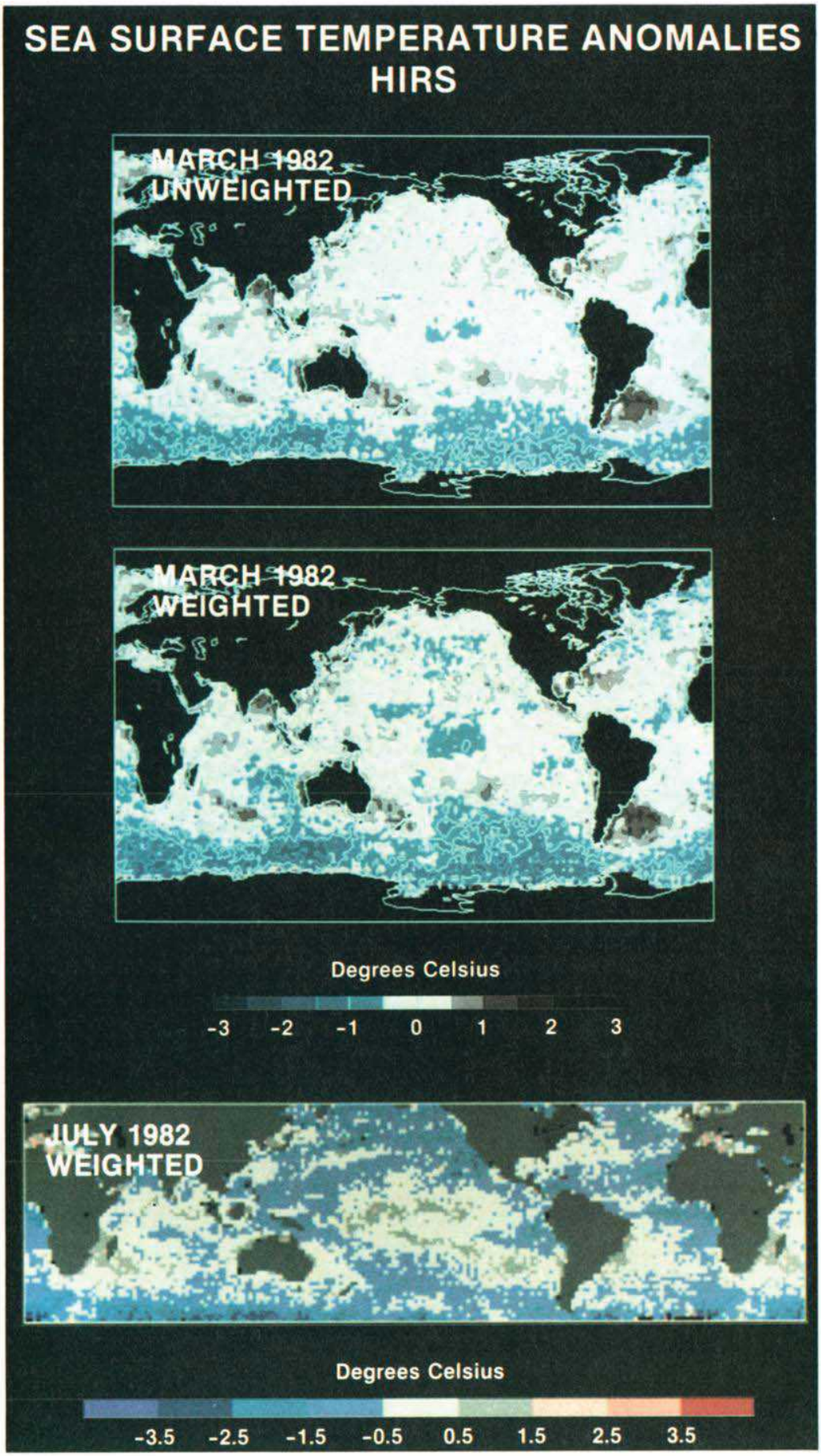

Plate 15. Sea surface temperature anomalies for HIRS: (top and middle) March 1982; (bottom) July 1982.