



## SEA-SURFACE TEMPERATURE ESTIMATION

Use of regression models for time/space interpolation of sea-surface temperature observations

E. R. Anderson

Research and Development Report

9 March 1967

U. S. NAVY ELECTRONICS LABORATORY, SAN DIEGO, CALIFORNIA 92152

TK  
7855  
5591  
105  
No. 1427

---

OF THIS DOCUMENT IS UNLIMITED

## THE PROBLEM

Develop statistical, physical, and computer techniques and methods for interpreting, summarizing, and extrapolating environmental data to support Navy requirements in research, developmental, and operational aspects of underwater detection, location, communications, and navigation. Specifically, study the use of regression models for time/space interpolation of sea-surface temperature observations.

MBL/WHOI



0 0301 0040519 7

## RESULTS

1. A regression model considering latitude, longitude, and day-of-year as the independent variables, together with empirically determined interaction terms, was found capable of estimating the seasonal variation of sea-surface temperature off the west coast of the United States, in water depths greater than 100 fathoms, to a standard deviation of less than  $1^{\circ}\text{F}$ .
2. The analysis suggests that more information than previously suspected can be obtained from a given number of observations provided realistic regression models can be developed. This suggestion has important implications with regard to sampling. A sampling interval based on the model can be used in place of the fixed time interval employed in the classical manner with an area grid. The oceanographic problem becomes one of searching for adequate models. It is indicated adequate models can be derived for many ocean areas from the present archive of oceanic temperature data.
3. On the assumption that the regression model is reasonably valid, the regression technique has the potential of being an effective method for identifying and editing raw temperature data for erroneous observations and for detecting and isolating temperature anomalies.
4. This study suggests that regression techniques may provide the basis for a new approach to summarizing archived sea-surface temperature data.

## RECOMMENDATIONS

1. Extend regression models to include depth as an independent variable.
2. Apply regression modeling techniques to describing the distribution of other oceanic parameters, such as salinity.

## ADMINISTRATIVE INFORMATION

Work was performed under SR 104 03 01, Task 0586 (NEL L40571). This report covers work from about January 1960 to June 1964 and was approved for publication 9 March 1967.

The author expresses appreciation to R. M. Lesser and Mrs. Jean Baker for doing the early computer programming, to H. W. Frye for assuming the computer programming responsibilities when Mr. Lesser and Mrs. Baker left the Laboratory, and to Mrs. G. L. Jones for assistance in preparing and handling the voluminous amounts of data utilized.



# CONTENTS

AUTHOR'S NOTE...	<i>page 7</i>
Oceanometrics defined...	8
INTRODUCTION...	11
THE REGRESSION MODEL...	12
Seasonal variation...	13
Latitudinal variation...	21
Longitudinal variation...	28
Interactions...	35
Summary...	36
RESULTS OF REGRESSION ANALYSIS...	37
DISCUSSION...	67
Time/space consistency...	67
Time/space distribution...	68
Data screening...	70
Anomaly detection...	73
Summarization of historical data...	75
CONCLUSIONS...	79
RECOMMENDATIONS...	81
REFERENCES...	83

## ILLUSTRATIONS

- 1 Example of day-to-day variation of sea-surface temperature for five locations in the eastern Pacific Ocean and the Atlantic Ocean... *page 14, 15*
- 2,3 Latitudinal and longitudinal variation of sea-surface temperature... *22-25, 30-33*
- 4 Area locator chart... *38*
- 5 Space/time distribution of sea-surface temperature observations in Areas B, C, and E... *40-41*
- 6 Statistical results for regression analyses on sea-surface temperature data taken in Areas A, B, C, and E... *42, 43*
- 7 Location of data used for large-area regression model... *44*
- 8 Time/space distribution of data used for sea-surface temperature comparisons... *46*
- 9 Comparison between observed and computed sea-surface temperatures... *47*
- 10 Differences between observed and computed sea-surface temperatures for water depths greater than and less than 100 fathoms... *49*
- 11 Differences between observed and computed sea-surface temperatures as a function of percent of observations made in water depths less than 100 fathoms... *51*
- 12-14 Differences between observed and computed sea-surface temperature as a function of day-of-year, latitude, and longitude... *52-57*
- 15-17 Regression models for three latitudinal strips... *60-65*
- 18 An example of the use of a regression equation to edit raw sea-surface temperature data... *72*
- 19 Differences between observed and computed sea-surface temperatures illustrating the use of a regression model for anomaly detection... *74*
- 20 Sea-surface temperature contours for 8 November 1949 computed from a regression model... *77*



## AUTHOR'S NOTE

This study was undertaken as a result of the following recommendation made in an NEL report published in 1960:<sup>1,\*</sup>

"Explore the utility of multiple regression (response surfaces), or more complex analyses of variance, in summarizing the entire body of collected data, since the manner in which the observations are collected limits the amount of the available data which can be used for comparisons at different places and times."

This recommendation resulted from a review of the manuscript of NEL Report 965 by Dr. George W. Snedecor, then Consultant in Statistics at NEL. The initial attempts to apply regression models to summarizing sea-surface temperature data were undertaken jointly by Dr. Snedecor and the author. The original intent was to publish these results as a coauthored report. With the retirement of Dr. Snedecor a number of years ago this is not now possible. However, the author wishes to especially acknowledge Dr. Snedecor's enthusiastic motivation, interest, and contribution to the results presented in this report.

---

\* See references at end of report.

## Oceanometrics Defined

The work described in this paper is often referred to as the "climatology of the oceans." Since the dictionary definition of "climatology" refers to the atmosphere only, Dr. Snedecor proposed that this aspect of oceanography be called "oceanometrics." The use of such a word has precedents in the fields of biology and economics, where "biometrics" and "econometrics" are well-defined words.

Since Dr. Snedecor first proposed the word some five years ago, its definition has been undergoing gradual evolution. Originally it was felt that the word implied a relationship to oceanography similar to the relationship of climatology to meteorology. In the minds of many people climatology is associated with the statistical summarization of measurements of atmospheric parameters, such as temperature, wind speed, and the like, with no implication concerning its ultimate application.

As a result a second definition evolved suggesting that oceanometrics occupied a position between the extremes of "pure dynamical oceanography" and "climatology of the ocean." In this sense pure dynamical oceanography is thought of as attempting to

construct simplified models and, from these models, to derive laws that describe what is happening in the ocean; and climatology of the oceans is thought of as collecting data on oceanographic parameters and presenting statistical summaries of these data with, in the extreme case, no thought to physical theory.

Recently a third definition has been suggested based on the assumption that most sciences develop in three stages: description, prediction, and control. At the present time the science of oceanography is phasing from description to prediction. The first stage, usually referred to as "descriptive oceanography," is primarily concerned with the reporting of data collected during exploratory data-collection cruises. The second stage, prediction, is primarily concerned with the quantitative analysis of oceanic data. This stage involves the symbolic expression of the physics of ocean behavior, or mathematical modeling. It is generally referred to as "dynamic oceanography" if it is limited to deterministic variables—that is, variables which are unencumbered by random variability. It is proposed that the second stage be referred to as "oceanometrics" if it is concerned with stochastic variables—that is, variables which include random variability.



## INTRODUCTION

This study is the second in a series concerned with the analysis of sea-surface temperature observations. The first study<sup>2</sup> dealt with the effect of missing data in long time-series sea-surface temperature measurements on certain regression and autocorrelation analyses.

For many decades observations of sea-surface temperature have been taken and recorded by merchant and naval vessels. Subsequently these observations have been catalogued and archived by many agencies. In the United States these agencies include the U. S. Naval Hydrographic Office (now the U. S. Naval Oceanographic Office), the National Weather Records Center, and the National Oceanographic Data Center.

As the volume of data accumulated, it became the basis for many generalized summaries.<sup>3-12</sup> All these summaries use arbitrary temporal and spatial averaging. Krümmel averaged the data over all years and all months. His areas were 5 degrees of latitude in the north-to-south dimension and extended east to west across an entire ocean. Böhnecke, the U. S. Weather Bureau, and the U. S. Naval Hydrographic Office used areas of 1-, 2-, or 5-degree squares of latitude and longitude and averaged all years together by month or season.

In recent years a requirement by researchers in the fields of fisheries oceanography, military oceanography, and meteorological oceanography for more detailed descriptions of sea-surface temperature distributions has developed. In response to this requirement the Bureau of Commercial Fisheries and the American Geographical Society have begun the preparation of more detailed charts. The Bureau of Commercial Fisheries is preparing detailed month-by-month charts of sea-surface

temperature in the North Pacific and the American Geographical Society is preparing similar charts for the Atlantic in the area of the Gulf Stream. The technique used to summarize the data is the same as that used in the earlier studies. Both studies summarize data for a particular year by monthly time intervals. The Bureau of Commercial Fisheries uses 2-degree-square areas and the Geographical Society uses 30-minute-square areas.

This study examines the potential of multiple-regression analysis as an approach to analyzing sea-surface temperature observations. Although multiple-regression techniques were developed by the statisticians many decades ago, they have rarely been used by oceanographers, because of the complexity inherent in developing realistic models and the magnitude of the arithmetic task required to evaluate the necessary constants. With the rapid progress in developing high-speed digital computers, the arithmetic computational difficulties have been overcome to the extent that it is now practical to consider relatively complex models.

## THE REGRESSION MODEL

In the use of regression analysis it is necessary to know, or to assume, (1) the major independent variables, or main effects; and (2) a functional relationship between these variables, or a regression model. In any given situation the desired functional relationship is generally determined from analytical or theoretical considerations or from a study of scatter diagrams prepared from the data being analyzed. In this study the latter approach is used.

The assumed independent variables are latitude, longitude, and day-of-year. A "point," or "cell," in the model is a 10-minute-by-10-minute area for a 1-day time period. A 10-minute-square area was selected, since the location of the data points is probably not known exactly and the initial interest is in the seasonal, or day-to-day, change in temperature.

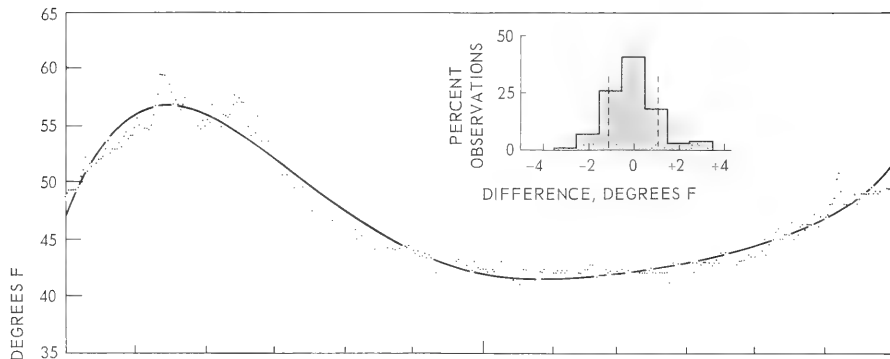
## Seasonal Variation

Sea-surface temperature records acquired over approximately 96 years were used to establish a functional relationship descriptive of the seasonal variation. The data were for the following locations:

50°N	145°W	Pacific Ocean	Weather Ship PAPA
51°50'N	131°W	Pacific Ocean	St. James Island
54°10'N	133°W	Pacific Ocean	Langara Island
32°50'N	117°15'W	Pacific Ocean	Scripps Pier, La Jolla, Calif.
35°N	48°W	Atlantic Ocean	Weather Ship ECHO

One year of measurement for each location is presented in figure 1 to give the reader some feel for how the individual temperature measurements vary throughout the year. A subjective examination of the data shows a more or less regular sinusoidal variation with season. In addition there are variations of a few days' to a few weeks' duration at irregular intervals. At the open-ocean locations, PAPA and ECHO, the shorter-period variations occur less frequently and their magnitude is smaller than at the

PAPA 1957



LANGARA 1954

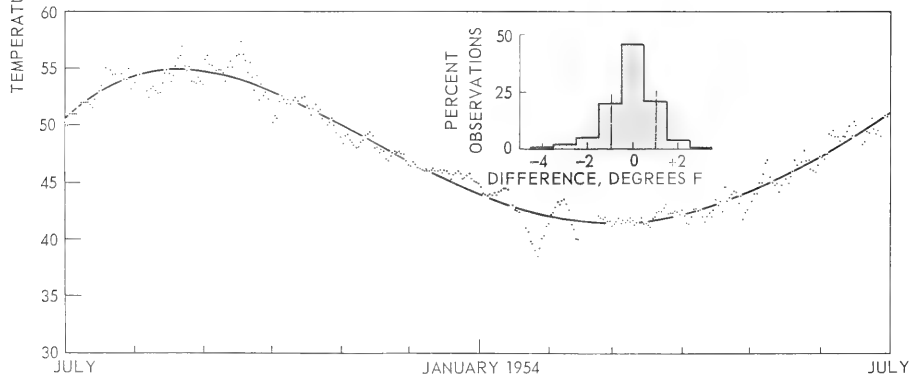
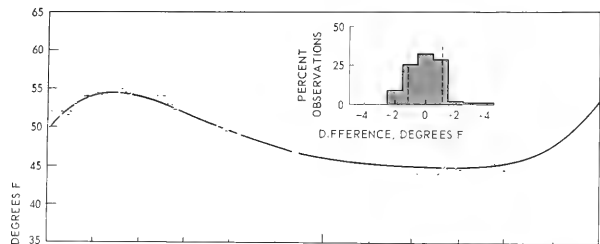
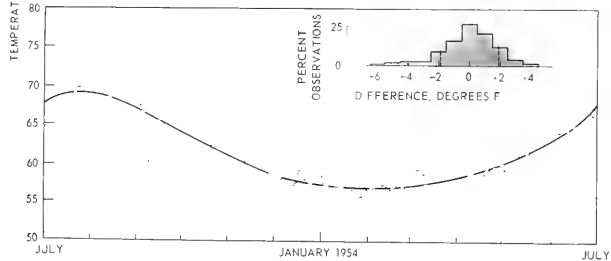


Figure 1. An example of the day-to-day variation of sea-surface temperature for five locations in the eastern Pacific Ocean and the Atlantic Ocean.

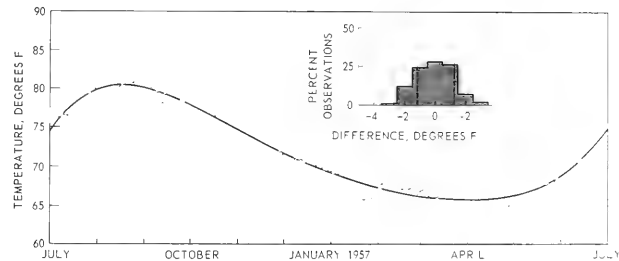




ST. JAMES 1954



SCRIPPS 1954



ECHO 1954

REVERSE SIDE BLANK

Figure 1 (Continued)



coastal location, Scripps Pier. It is also recognized that there is a diurnal variability present in these data since the daily observations were taken at random times during the day. The latter short-period variations will not be included in the model, with the result that their effect will contribute to the unexplained variance.

Since the seasonal variation in sea-surface temperature is not symmetrical about an origin, the use of the following fifth-degree polynomial is suggested by the scatter diagrams:

$$T' = a_0 + a_1D + a_2D^2 + a_3D^3 + a_4D^4 + a_5D^5 \quad (1)$$

where  $D$  is measured in days from some arbitrary origin,  $T'$  is the least-squares fitted value, or estimate, of surface temperature, and the sub scripted  $a$ 's are regression coefficients to be estimated.

Equation (1) was fitted to 5 years of data taken at each of the five locations listed above to demonstrate the adequacy of the fifth-degree polynomial as an estimator of the seasonal sea-surface temperature variation. The origin of time was taken as July 1 and the years referred to are fiscal years. The notation "1954" refers to a fiscal year 1954 starting 1 July 1953, and ending 30 June 1954.

The following related quantities are used as measures of the "goodness of fit" of equation (1) to the observed data:  $R$ , multiple correlation coefficient;  $100R^2$ , percent variance explained by regression; and  $\sigma$ , standard deviation in degrees Fahrenheit of the observations about the regression curve.

<u>Fiscal year</u>	<u>1957</u>	<u>1958</u>	<u>1959</u>	<u>1960</u>	<u>1961</u>
<u>Weather ship PAPA</u>					
Observations	254	199	210	245	247
$R$	0.98	0.99	0.99	0.99	0.99
$100R^2$	95.9	97.2	97.7	97.6	97.5
$\sigma$	1.1	0.9	0.8	0.7	0.9
<u>Weather ship ECHO</u>					
Observations	229	221	277	264	100
$R$	0.96	0.98	0.98	0.98	0.98
$100R^2$	92.3	96.7	96.7	95.4	96.4
$\sigma$	1.2	1.0	1.0	1.2	1.0
<u>Cape St. James</u>					
Observations	325	296	313	163	245
$R$	0.95	0.94	0.93	0.95	0.92
$100R^2$	90.5	88.8	85.6	90.4	83.9
$\sigma$	1.0	1.5	1.2	1.1	1.0

<u>Fiscal year</u>	<u>1957</u>	<u>1958</u>	<u>1959</u>	<u>1960</u>	<u>1961</u>
<u>Langara Island</u>					
Observations	345	354	343	343	306
$R$	0.96	0.95	0.91	0.98	0.96
$100R^2$	92.4	90.8	83.1	95.9	91.3
$\sigma$	1.3	1.3	1.7	1.0	0.9
<u>Scripps Pier</u>					
Observations	362	357	364	359	358
$R$	0.96	0.93	0.91	0.91	0.94
$100R^2$	91.6	86.2	83.0	83.4	87.9
$\sigma$	1.3	1.6	1.8	1.9	1.8

An examination of these statistics supports the conclusion that a fifth-degree polynomial is an acceptable estimator of the seasonal variation. Relatively, the fit is best for open-ocean data (PAPA and ECHO), with 92 to 98 percent of the observed variability explained; next best at island locations (St. James Island and Langara Island), with 83 to 96 percent of the variability explained; and poorest at Scripps Pier, located about 1000 feet from shore, with 83 to 92 percent of the variability explained. Equation (1) has been fitted to many other years of data with similar results.

On figure 1 the solid line is a plot of equation (1) using the regression coefficients for the indicated location and year. The histograms show the distribution of differences between the observed and estimated sea-surface temperatures and the vertical dotted lines indicate one standard deviation of these differences.

It is of interest to note that from September 1 to January 1 the rate of cooling varies from  $2.0^{\circ}\text{F}$  per 30-day period at St. James Island to  $3.4^{\circ}\text{F}$  per 30-day period at PAPA. The rate of warming from May 1 to July 1 varies from  $3.2^{\circ}\text{F}$  per 30-day period at Langara Island to  $4.0^{\circ}\text{F}$  per 30-day period at ECHO and St. James Island.

## Latitudinal Variation

The observations used to examine the latitudinal variation in sea-surface temperature were taken from Punched Card Deck 116, U. S. Merchant Marine and Other Ship Observations, 1949 --- , of the National Weather Records Center. These are marine weather observations which include, among other parameters measured, a sea-water temperature. These observations are usually taken by a mercury-in-glass thermometer installed in the ship's sea-water-intake system and are reported to the nearest whole degree Fahrenheit.<sup>13</sup> Listings of all weather observations made in the North Pacific north of  $20^{\circ}\text{N}$  for the years 1956 and 1957 were obtained from the Records Center.

From the punched-card deck the sea-surface temperatures taken north of  $20^{\circ}\text{N}$  and along a given longitude,  $\pm 0.2^{\circ}$  of longitude, were selected to examine the variation of surface temperature as a function of latitude. The longitude strips selected were:  $126^{\circ}$ ,  $129^{\circ}$ ,  $132^{\circ}$ ,  $135^{\circ}$ ,  $138^{\circ}$ , and  $141^{\circ}\text{W}$ . The data from three of these strips for March and September 1956 and 1957 are plotted in figure 2. March and September were used to minimize the effect of the seasonal change in temperature.

138 DEGREES W LONGITUDE

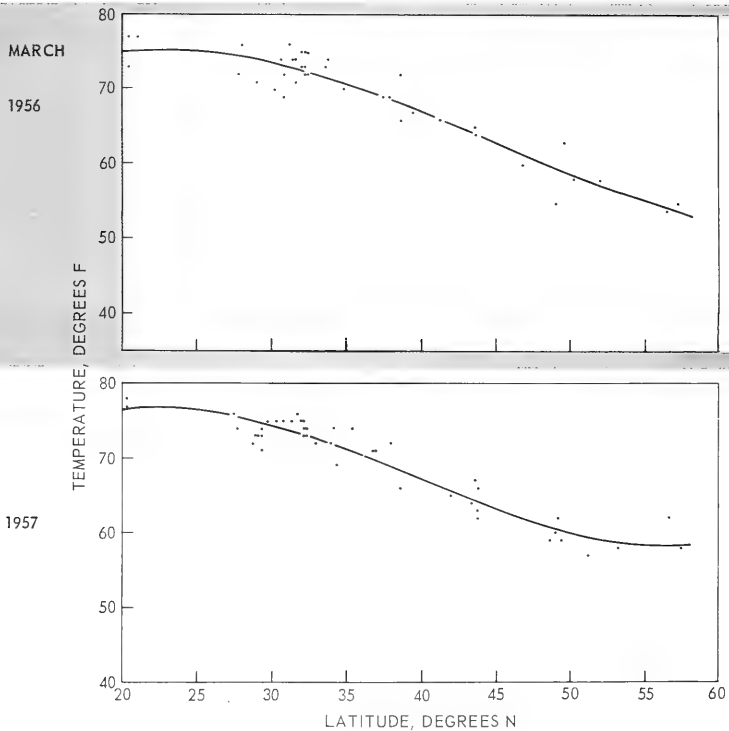


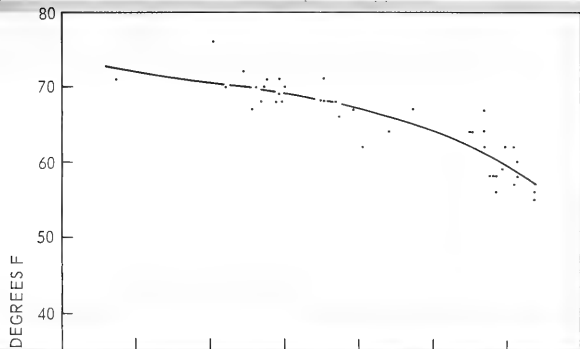
Figure 2. Latitudinal variation of sea-surface temperature.



132 DEGREES W LONGITUDE

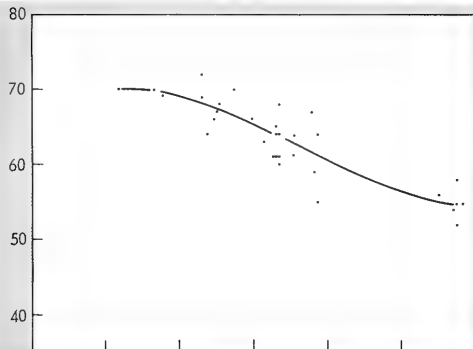
MARCH

1956

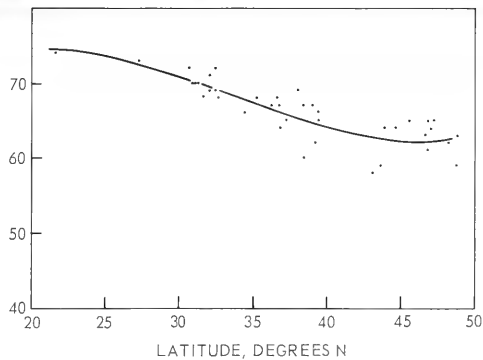
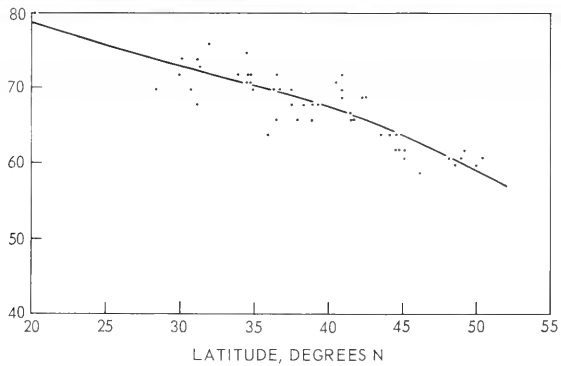


126 DEGREES W LONGITUDE

TEMPERATURE, DEGREES F



1957



138 DEGREES W LONGITUDE

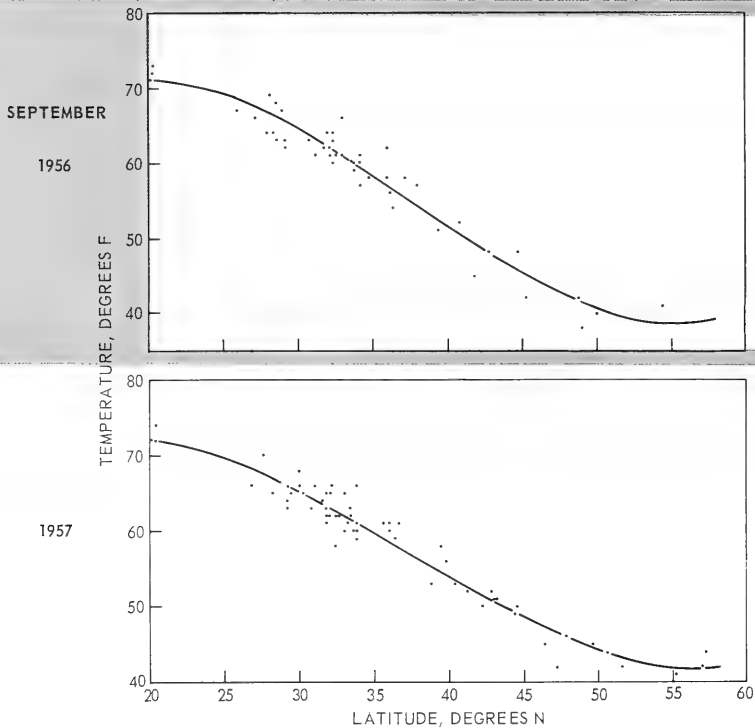
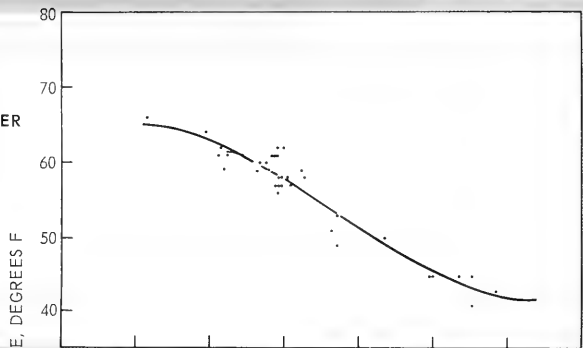
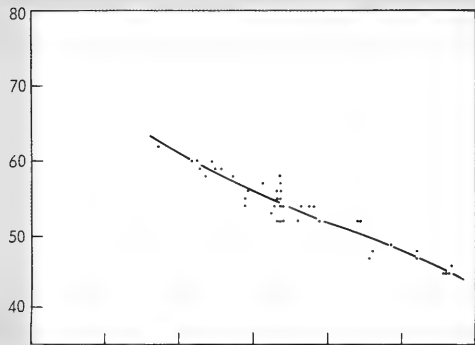


Figure 2 (Continued)

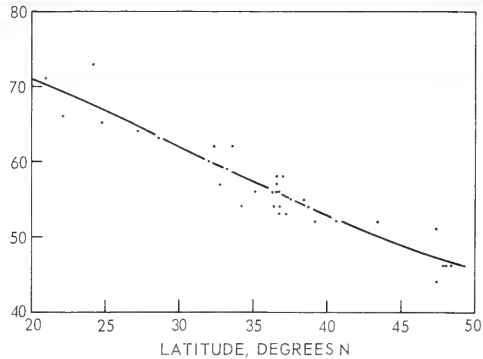
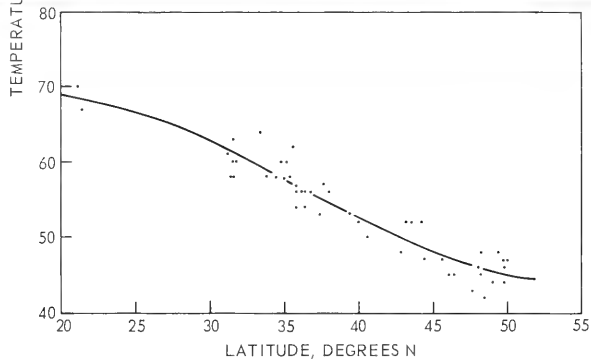
132 DEGREES W LONGITUDE



126 DEGREES W LONGITUDE



1957



A subjective study of these scatter diagrams suggests a polynomial of the following form:

$$T' = a_0' + a_6 L - a_7 L^2 - a_8 L^3 \quad (2)$$

where  $L$  is the latitude,  $T'$  is the estimated sea-surface temperature, and the subscripted  $a$ 's are the regression coefficients to be estimated.

Equation (2) was fitted to the above 24 sets of data. The following statistics for the data shown in figure 2 were obtained:

Year	Longitude	Number of Observations	$R$	$100R^2$ , percent	$\sigma$ , degrees F
<u>MARCH</u>					
1956	126°W	44	0.95	89.6	1.5
	132°	42	0.96	91.8	1.8
	138°	50	0.97	93.7	2.1
1957	126°W	38	0.95	90.0	2.1
	132°	105	0.94	89.0	2.3
	138°	65	0.97	94.4	1.9

Year	Longitude	Number of Observations	R	100 R <sup>2</sup> , percent	$\sigma$ , degrees F
<u>SEPTEMBER</u>					
1956	126° W	37	0.88	76.7	2.6
	132°	44	0.90	81.9	2.3
	138°	41	0.95	89.7	2.1
1957	126° W	41	0.84	71.1	2.1
	132°	59	0.88	78.4	2.3
	138°	52	0.96	92.0	1.7

Relatively the fit is best for data taken the greatest distance from shore, the percent variance explained by regression varying systematically from 94 to 71 percent. The reader is reminded that the original data were reported to the nearest whole degree Fahrenheit and that many of the temperatures are "injection" temperatures taken at some depth below the surface. The standard deviation would be expected to be greater for these data than for the data used to establish the seasonal variation.

On figure 2 the solid line is a plot of equation (2) using the regression coefficients for the proper latitude, year, and month. A third-degree polynomial appears to exhibit the flexibility necessary to obtain a reasonable estimation of the latitudinal variation.

It is of interest to note that from 30° to 40° N the temperature decreases about 1.2° F per 1-degree change in latitude, about 60 nautical miles.

## Longitudinal Variation

The observations used to examine the longitudinal variation were also obtained from Punched Card Deck 116.

From this deck the sea-surface temperatures taken along a given latitude,  $\pm 0.2^\circ$  of latitude, were selected. Latitude strips, at 3-degree intervals, were selected starting at  $30^\circ$  N. The data for four of these strips— $30^\circ$ ,  $36^\circ$ ,  $42^\circ$ , and  $48^\circ$  N—for March and September 1956 and 1957 are plotted on figure 3.

The data suggest that a third-degree polynomial can be used as a model where

$$T' = a_0'' + a_9G + a_{10}G^2 + a_{11}G^3 \quad (3)$$

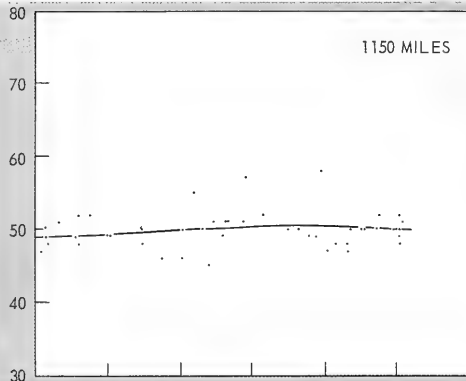
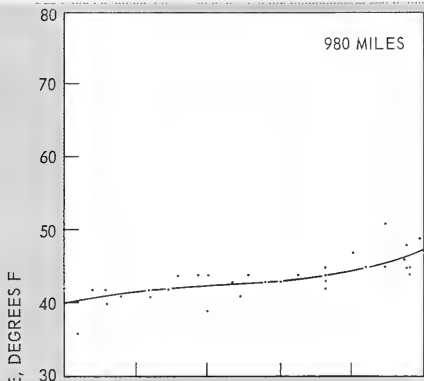
and  $G$  is the longitude.

Equation (3) was fitted to the above sets of data. The following statistics for the data shown on figure 3 were obtained:

Year	Latitude	Number of Observations	R	$100R^2$ , percent	$\sigma$ , degrees F
<u>MARCH</u>					
1956	30° N	95	0.82	68.0	2.0
	36°	102	0.65	41.7	2.3
	42°	54	0.17	2.8	2.3
	48°	66	0.71	50.1	1.6
1957	30° N	110	0.82	66.8	1.6
	36°	107	0.70	49.3	2.0
	42°	74	0.60	35.7	2.1
	48°	65	0.15	2.3	2.0
<u>SEPTEMBER</u>					
1956	30° N	67	0.88	77.8	1.7
	36°	96	0.94	88.2	2.2
	42°	52	0.86	74.2	2.3
	48°	66	0.69	48.2	2.8
1957	30° N	91	0.75	55.6	2.0
	36°	89	0.69	47.4	2.5
	42°	72	0.66	43.6	3.0
	48°	63	0.62	38.2	1.4

48 DEGREES N LATITUDE

42 DEGREES N LATITUDE

MARCH  
1956

1957

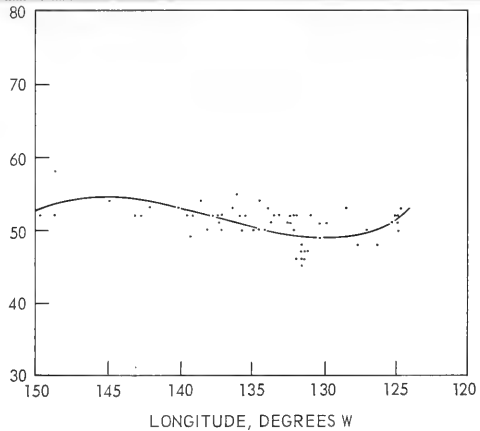
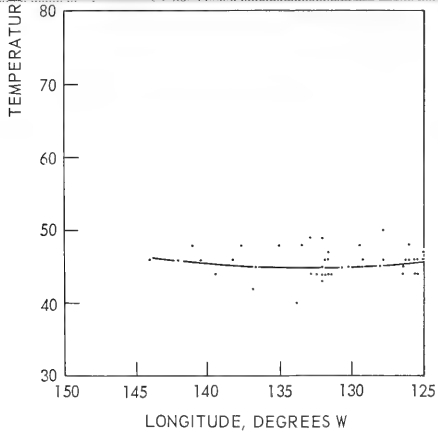
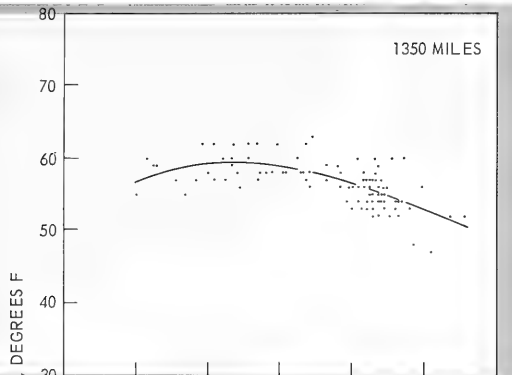


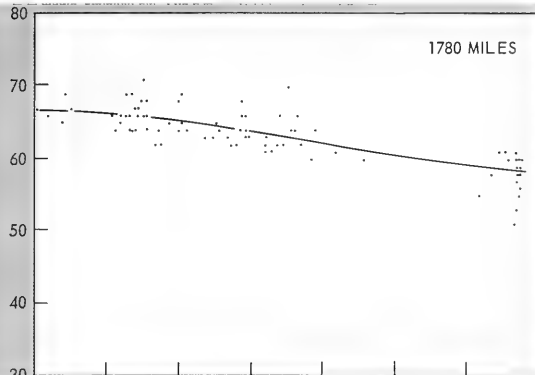
Figure 3. Longitudinal variation of sea-surface temperature.



## 36 DEGREES N LATITUDE

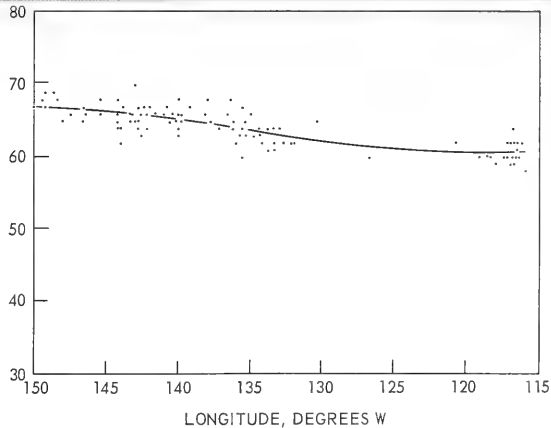
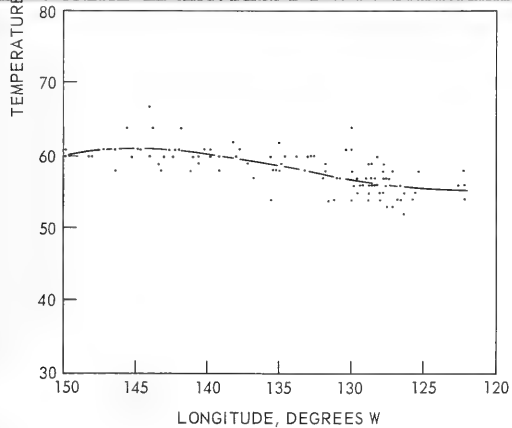


## 30 DEGREES N LATITUDE



MARCH

1956

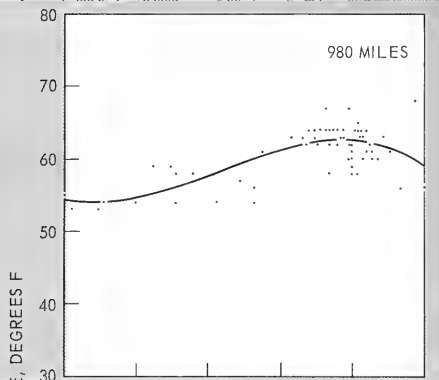


1957

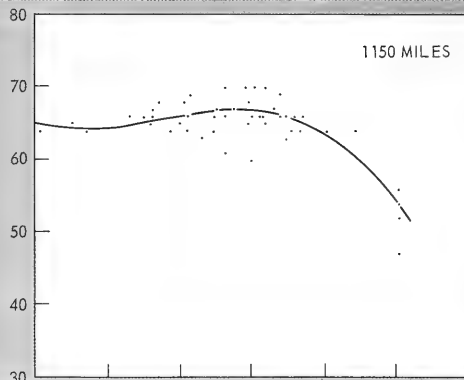
SEPTEMBER

1956

48 DEGREES N LATITUDE



42 DEGREES N LATITUDE



1957

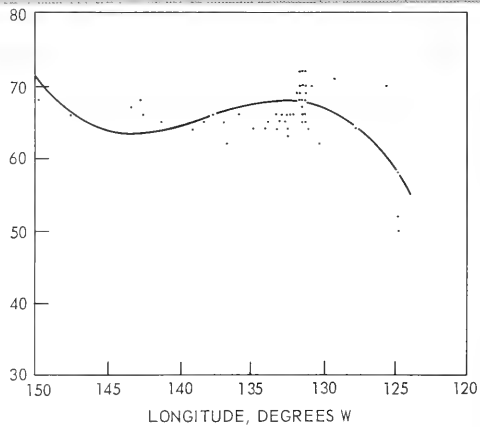
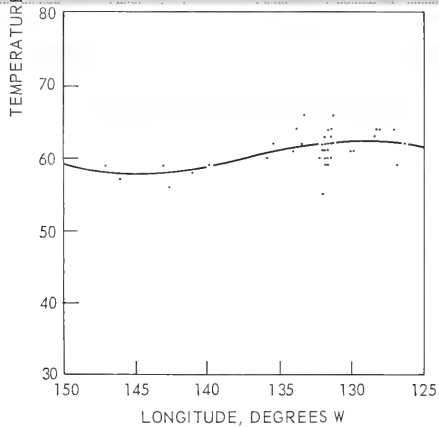
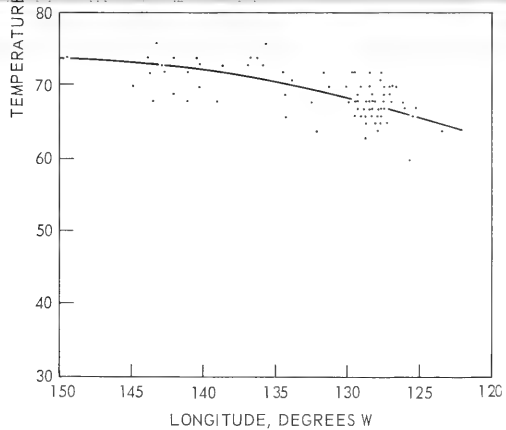
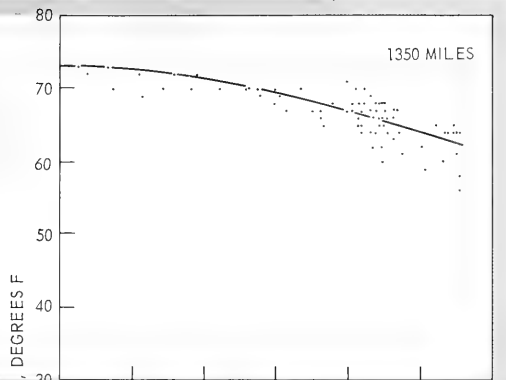
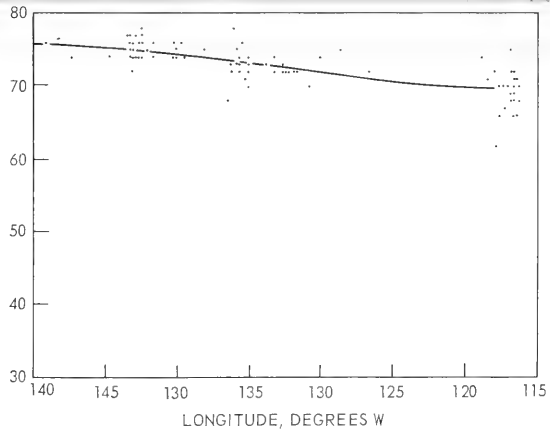
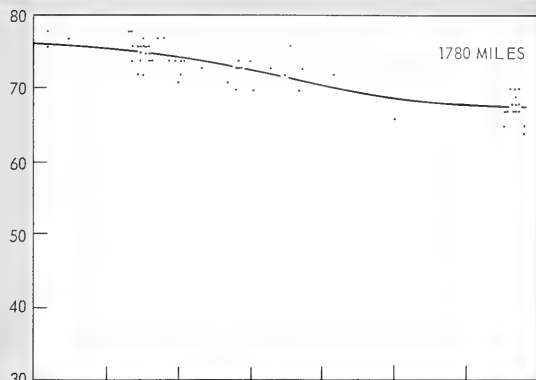


Figure 3 (Continued)

## 36 DEGREES N LATITUDE



## 30 DEGREES N LATITUDE



SEPTEMBER

1956

1957

The variability in the  $R$  and the related  $100R^2$  is considerably more in longitude than in the seasonal or latitudinal analysis, while the variability in the standard deviation is about the same in longitude and in the latitudinal analysis. The variability in the  $R$  and  $100R^2$  is related to the fact that for the samples with the smaller  $R$  and  $100R^2$  there is little change in temperature with longitude. In other words, there is not much systematic variability to explain. For the samples with the larger  $R$  and  $100R^2$  values, the systematic longitudinal variation is relatively greater. Thus, these statistics are not comparable, since they are not independent, for any given sample, of the overall change in temperature with longitude. On the other hand the standard deviation, which is a measure of the random variability in the variation of the temperature, may be compared from sample to sample.

It is concluded that in the geographical area under consideration the surface temperature is less sensitive to longitudinal change than to latitudinal and seasonal change, that there are considerable differences between latitudinal strips, and that equation (3) is flexible enough to describe these differences.

## Interactions

An examination of the data in figure 2 shows that the temperature variation with latitude differs from one longitude to another, indicating that there are interactions between latitude and longitude. In addition, there are indications that there are interactions between latitude and day, and longitude and day. It is necessary, therefore, to include such interactions in the model if the model is to be realistic.

A trial-and-error approach was used to obtain information on the characteristics of the interactions to be included in the model. Many combinations of the main-effect terms were tried and discarded. The following terms appear to be the most important for the area under consideration:

$$a_{12}LD + a_{13}LD^3 + a_{14}LD^5 \quad \text{latitude by day}$$

$$a_{15}GD + a_{16}GD^3 + a_{17}GD^5 \quad \text{longitude by day}$$

$$\left. \begin{aligned} a_{18}LG + (a_{19}G^2 + a_{20}G^3)L + \\ (a_{21}L^2 + a_{22}L^3)G \end{aligned} \right\} \quad \text{latitude by longitude}$$

## Summary

Equations (1), (2), and (3) together with the above interaction terms were combined to form a 22-variable regression model with 23 coefficients to be determined by least squares:

Surface temperature =

$$\begin{array}{l}
 a_0 - a_1 D - a_2 D^2 - a_3 D^3 - a_4 D^4 - a_5 D^5 - \quad \text{day-of-year} \\
 a_6 L - a_7 L^2 - a_8 L^3 - \quad \text{latitude} \\
 a_9 G - a_{10} G^2 - a_{11} G^3 - \quad \text{longitude} \\
 \hline
 a_{12} LD - a_{13} LD^3 - a_{14} LD^5 - \quad \text{latitude by day} \\
 a_{15} GD - a_{16} GD^3 - a_{17} GD^5 - \quad \text{longitude by day} \\
 \hline
 a_{18} LG - \quad \text{interactions} \\
 (a_{19} G^2 - a_{20} G^3)L - \quad \text{latitude by longitude} \\
 (a_{21} L^2 - a_{22} L^3)G
 \end{array} \quad \left. \vphantom{\begin{array}{l} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{array}} \right\} \quad (4)$$

where  $D$ ,  $L$ , and  $G$  are the day-of-year, latitude, and longitude, respectively. This model is applicable to an area off the west coast of the United States extending from  $20^\circ\text{N}$  to  $58^\circ\text{N}$  and from the coast to  $150^\circ\text{W}$ .

## RESULTS OF REGRESSION ANALYSIS

Equation (4), or a modification of it, was fitted to surface temperature observations taken by a bathythermograph in the areas A, B, C, and E shown in figure 4. In addition it was fitted to the large area shown extending from  $30^{\circ}$  to  $49^{\circ}$  N and seaward about 650 miles. In the latter area the sea-surface temperature observations used were made in the four shaded 1-degree-longitude strips (B, C, D, and E) shown in figure 4. The measurements were treated as a single sample drawn from the large area.

Equation (4) was developed as a model for the largest area. Since the other areas cover smaller intervals of latitude and longitude, the terms in equation (4) that involve the higher orders of these variables are omitted. A 3-month overlap in time was used in making the least-squares fit to control the behavior of the fifth-degree polynomial. Thus, the data used covered an 18-month period extending from 1 April of a given year to 1 October of the following year, and the resulting equation was used to estimate the temperature during the included fiscal year.

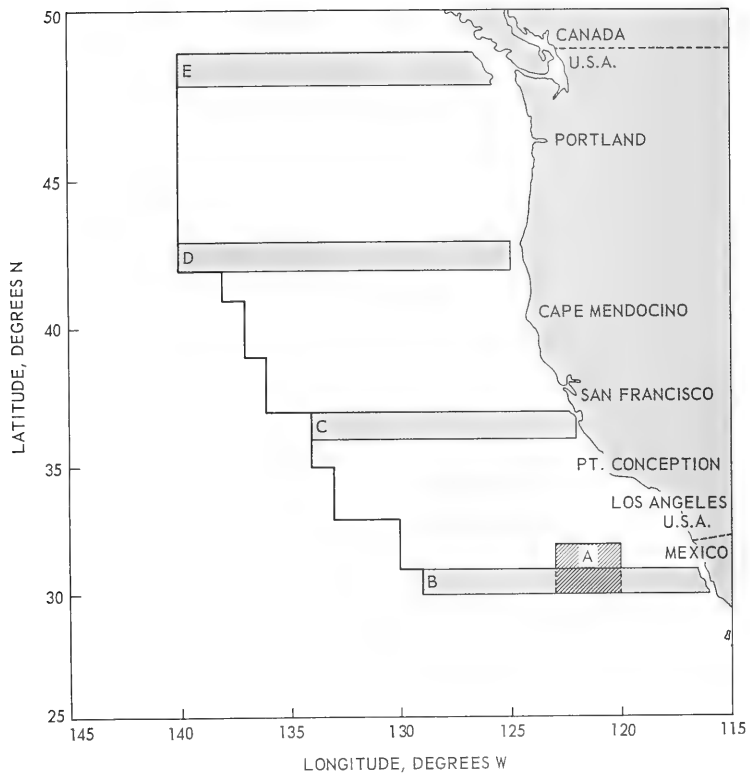


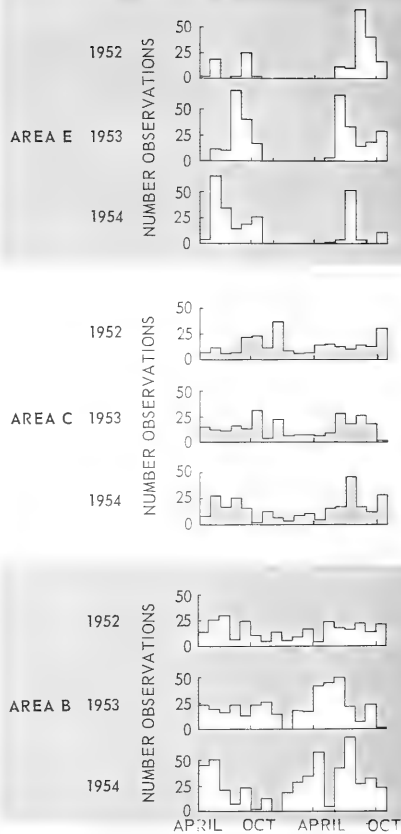
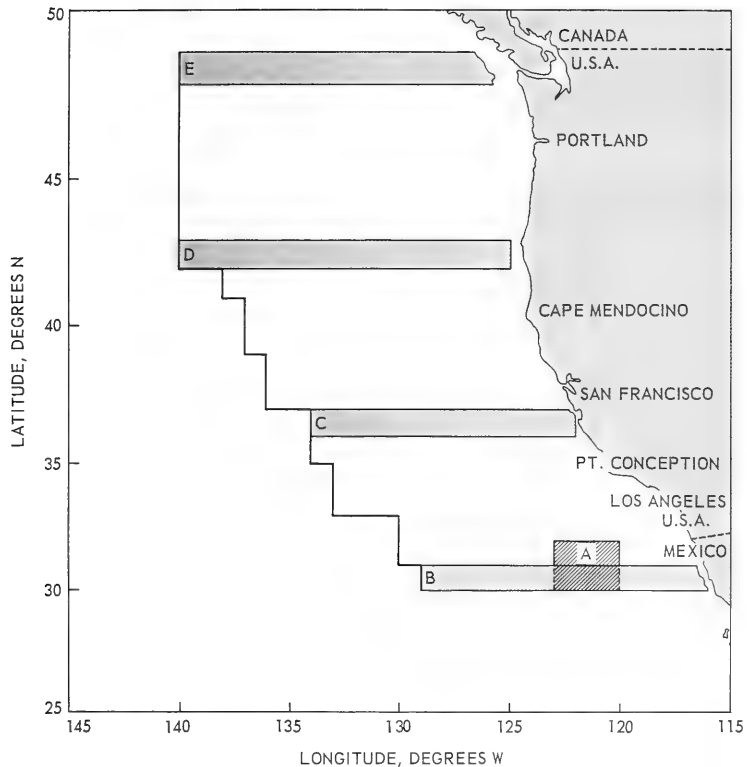
Figure 4. Area locator chart.



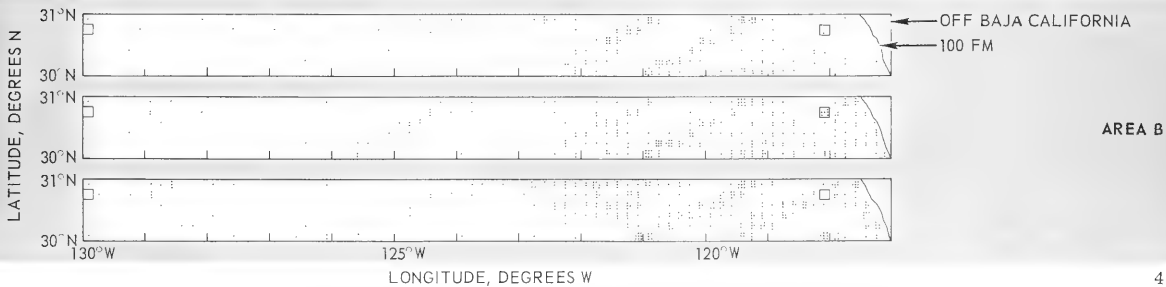
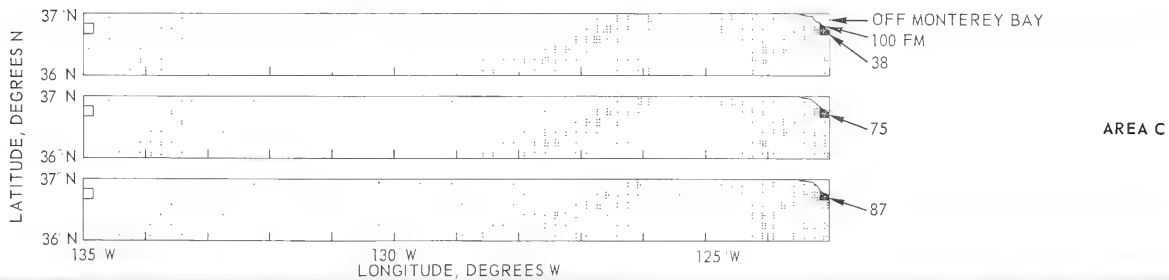
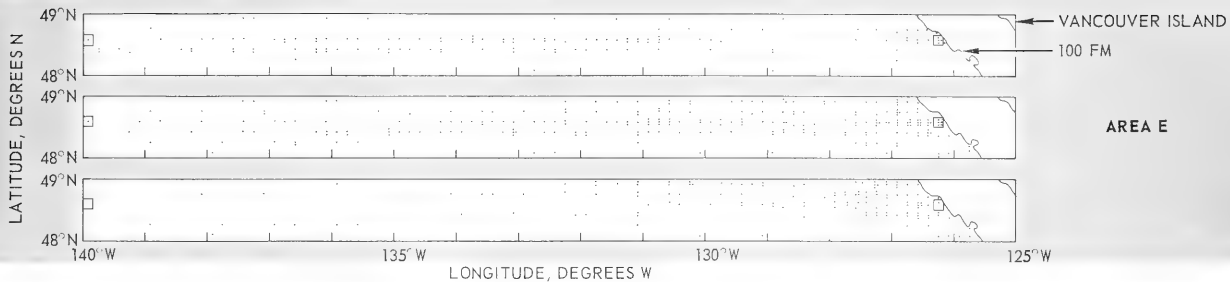
Figure 5 is an example of how the data are distributed with respect to space and time. Data taken in waters less than 100 fathoms in depth were excluded from the analysis. The data distributions for areas B, C, and E for the years 1952, 1953, and 1954 are included. The histograms on page 40 show the monthly distributions and the figures on page 41 the geographical distributions. The nonuniform data distribution with respect to time and space is obvious. Geographically, most of the data are in the area nearest shore. The number of data decreases rapidly to the west, many 10-minute-square areas containing no data. Area C is notable for its nonuniform spatial distribution. Temporally, most of the data were taken during the spring and summer. Area E is notable for its nonuniform temporal distribution.

The distributions of the observations in the other data sets used in this study exhibit similar characteristics.

Equation (4) was least-squares fitted to 5 years of data taken in Area A from 1951 to 1955 inclusive and in areas B, C, and E from 1950 to 1954 inclusive. Each data set consisted of sea-surface temperatures as recorded in degrees Fahrenheit on bathythermograms taken during the indicated year and in the indicated area. The results of the 20 individual regression analyses are presented in figure 6. For each analysis the number of observations, the multiple correlation coefficient, the percent of the variance explained by regression, and the standard deviation of the observations about regression are shown. The last analysis utilized the data taken in the four 1-degree-latitude strips. As indicated in figure 7 these measurements were treated as a single sample drawn from the area  $30^{\circ}$  to  $49^{\circ}$  N and extending



40 Figure 5. Space-time distribution of sea-surface temperature observations in Areas B, C, and E.



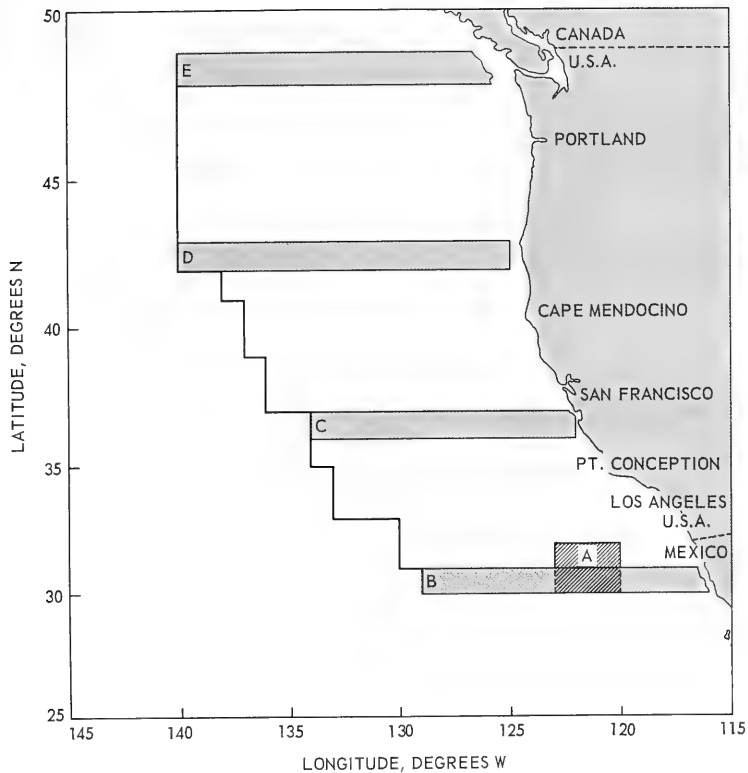


Figure 4. Area locator chart.

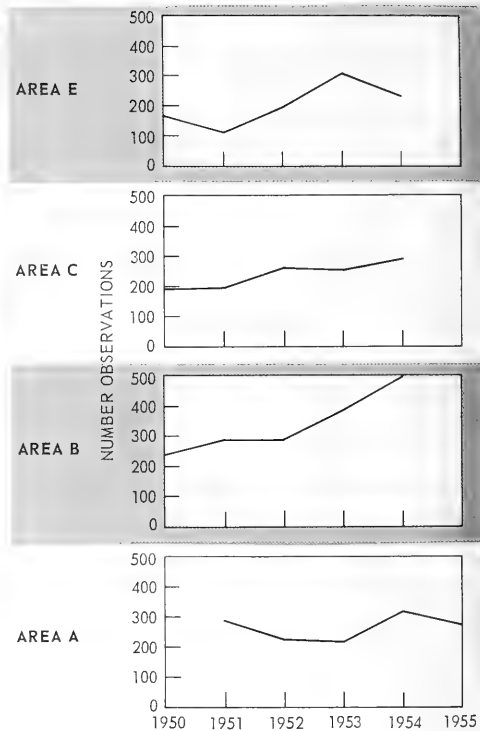
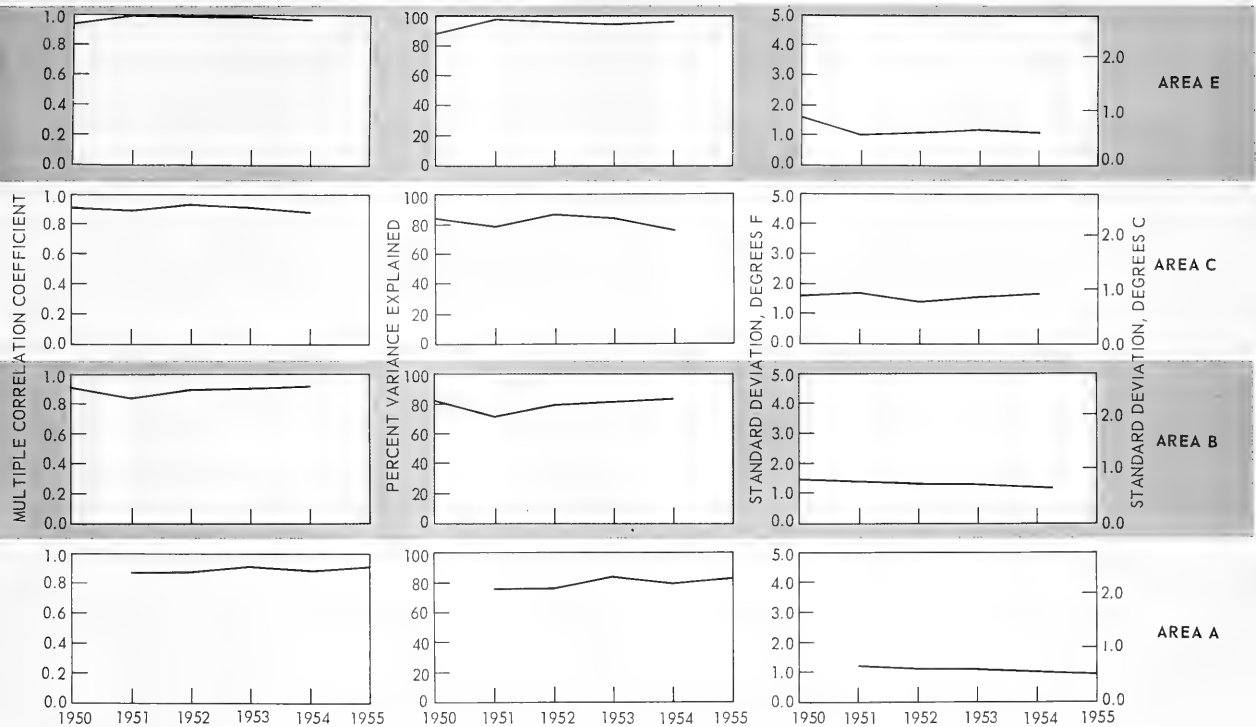


Figure 6. Statistical results for regression analyses on sea-surface temperature data taken in Areas A, B, C, and E.



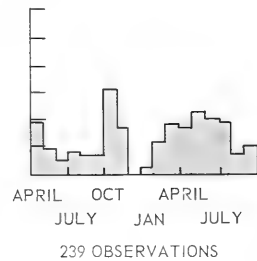
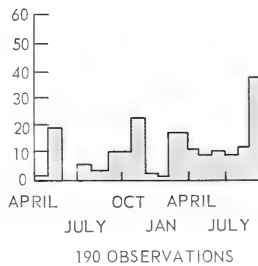
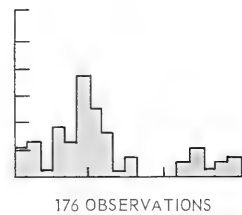
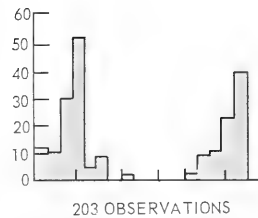
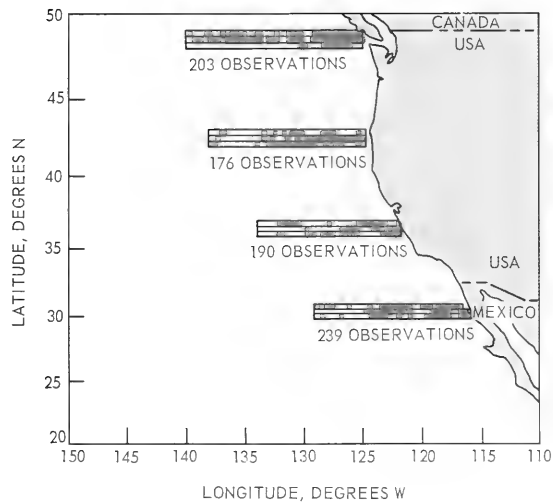


Figure 7. Location of data used for large-area regression model.

seaward about 650 nautical miles and for an 18-month time period extending from 1 April 1949 to 1 October 1950. The total number of observations in each area and their temporal and spatial distribution are also shown. In the shaded areas one to 20 observations were made and in the unshaded areas no observations were made. Equation (4) was fitted to the data. The statistical results were as follows:

Number of observations: Area B	239
Area C	190
Area D	176
Area E	203
Total	808
$100R^2$ , percent variance explained by regression	85.7
$R$ , multiple correlation coefficient	0.93
$\sigma$ , standard deviation in degrees Fahrenheit of the observations about regression	1.9

Figure 8 shows the location in time and space of 971 temperature observations made in this area during fiscal year 1950. The observations were not used in obtaining the regression equation but were used as a control to see how well the regression equation could estimate independently observed sea-surface temperatures. The difference between the observed temperature and that estimated from regression was obtained. The results are summarized in figure 9. The standard deviation of the differences was  $2.3^\circ\text{F}$  compared to  $1.9^\circ\text{F}$  for the regression equation.

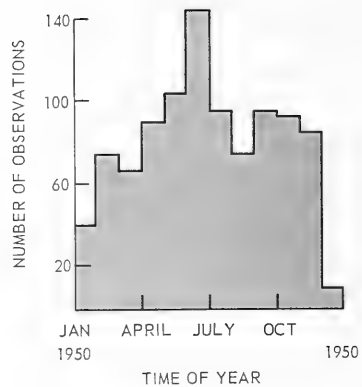
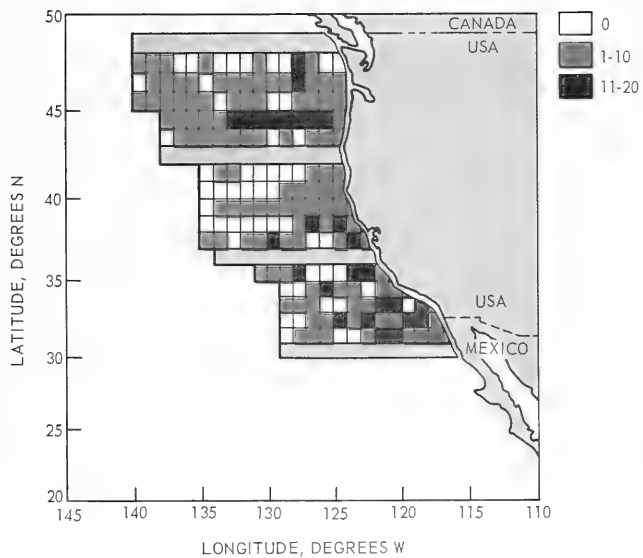


Figure 8. Time-space distribution of data used for sea-surface temperature comparisons.



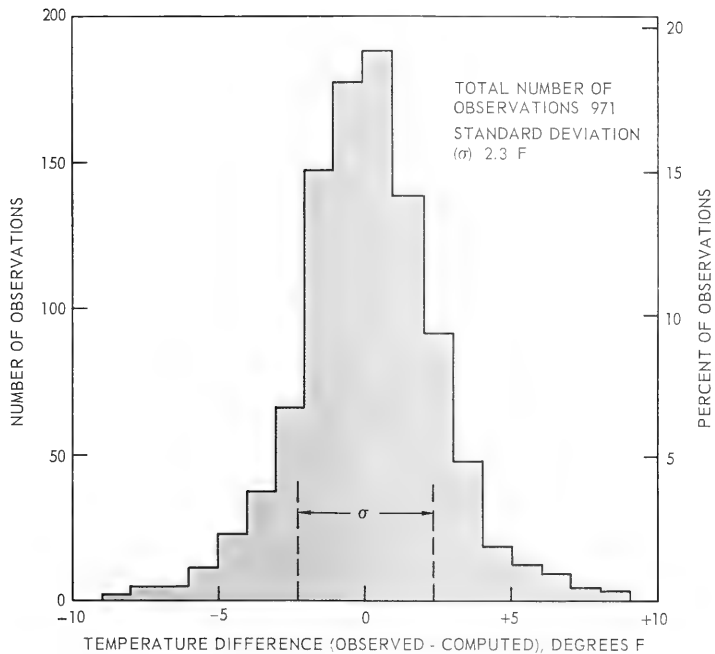


Figure 9. Comparison between observed and computed sea-surface temperatures.

Regression analysis appears to have considerable potential as a technique for estimating sea-surface temperatures. However, the physical reality of the estimates must also be considered, since it is always possible to improve the statistical measures of goodness of a regression-model estimate by merely adding additional terms to the model. From a physical viewpoint these terms may be nonsense terms.

Regression analyses of Area B (1954), Area C (1953), and Area E (1952) are examined to illustrate the physical reality of the model. The difference between observed values of temperature and the values obtained from regression will be considered as a function of water depth, time, latitude, and longitude.

It was noted in studying the results of some of the earlier analyses that many of the large differences between the temperatures obtained from regression and those obtained by observation occurred in the shallower water adjacent to the coast line. This finding was not surprising, since transient and local effects, which are not accounted for in the model, should have their maximum influence on the temperature in such areas.

Figures 10 and 11 present a qualitative histogram analysis of the effect of water depth for the three regressions. Figure 10, for each area, contains three histograms. The shaded portion of the histogram on the left shows the distribution of differences for data taken in water depths less than 100 fathoms, while the unshaded histograms include all data used in the regression analysis.

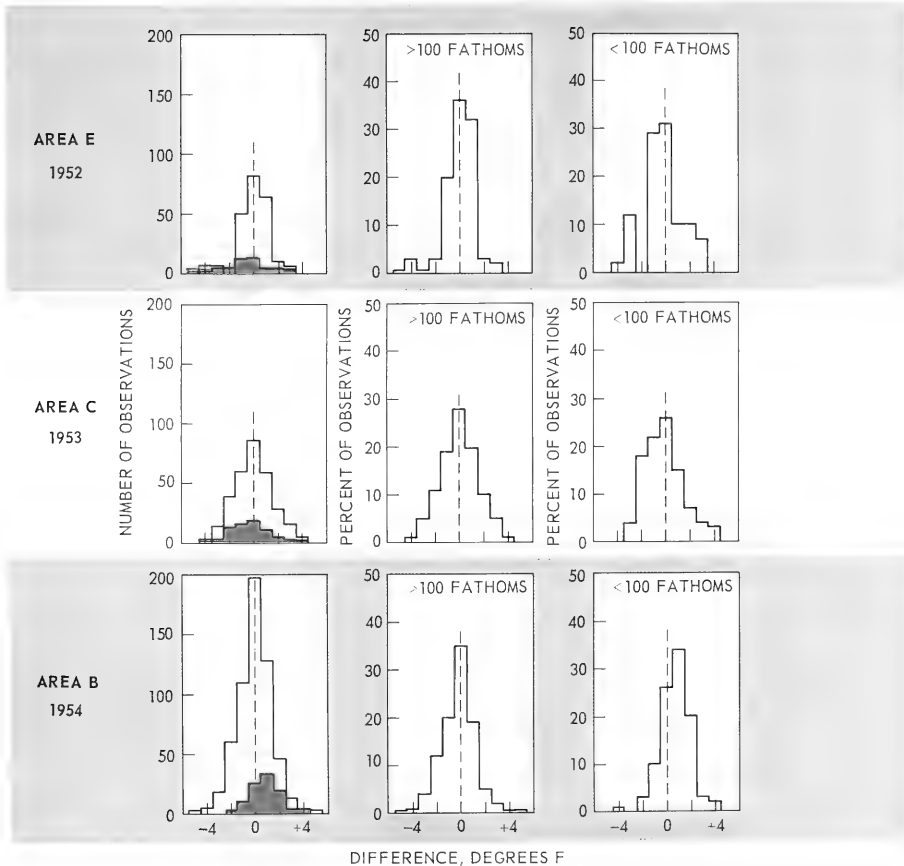


Figure 10. Differences between observed and computed sea-surface temperatures for water depths greater than and less than 100 fathoms.

The center histograms show the distribution of differences for observations taken in water depths greater than 100 fathoms and the histograms on the right show the differences for observations taken in water depths less than 100 fathoms.

Figure 11 contains histograms of the differences between observation and regression as a function of the percent of total observations taken in water of less than 100 fathoms.

An examination of the histograms, particularly those of figure 11, leads to the not unexpected qualitative conclusion that differences between regression and observation for the observations taken in water depths less than 100 fathoms are greater than for those taken in deeper water. The conclusion suggests that variables other than latitude, longitude, and day-of-year are important in determining the distribution of temperature in shallower water.

Figure 12 shows the differences between the observed temperature and that computed from regression as a function of day-of-year for each of the three areas. In addition a two-standard-deviation interval is shown on the right. From a seasonal point of view the differences seem to be randomly distributed about zero difference. However, for time periods of a few days to tens of

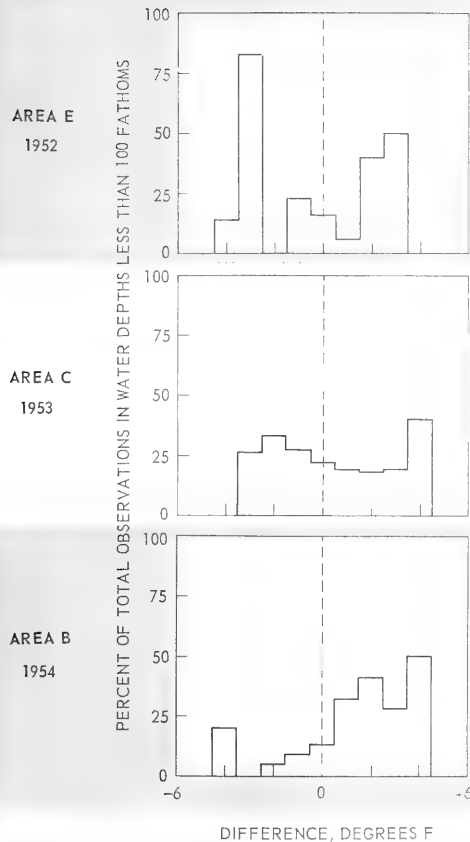
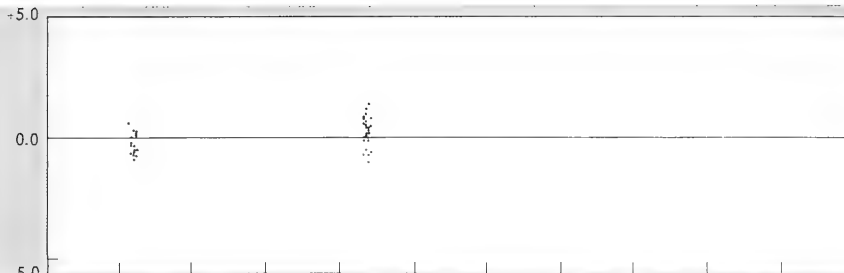
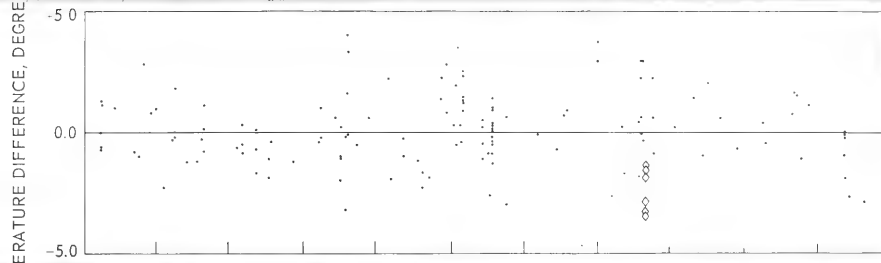


Figure 11. Differences between observed and computed sea-surface temperatures as a function of percent of observations made in water depths less than 100 fathoms.

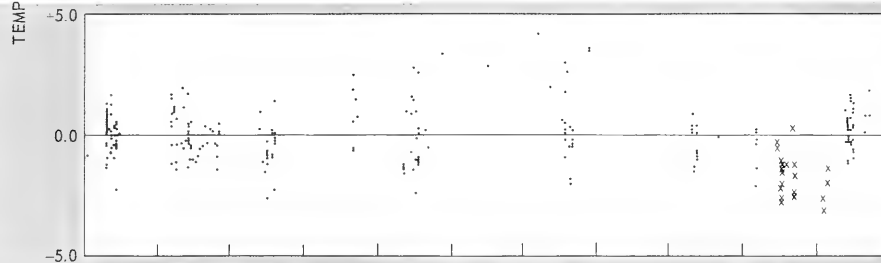
AREA E (1952)



AREA C (1953)

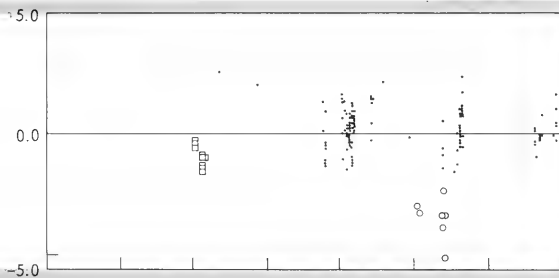


AREA B (1954)

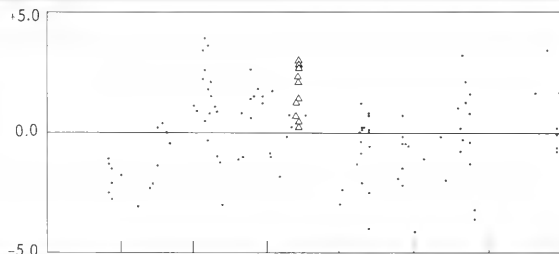


1 APRIL 1 JULY 1 OCTOBER 1 JANUARY 1 MARCH  
TIME, DAYS

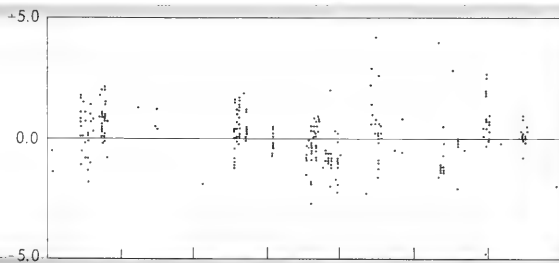
Figure 12. Differences between observed and computed sea-surface temperature as a function of day-of-year.



AREA E (1952)



AREA C (1953)



AREA B (1954)

1 MARCH 1 APRIL 1 JULY 1 OCTOBER

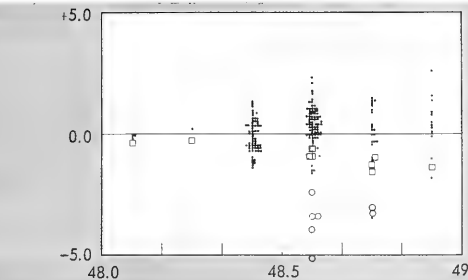
- PERIOD 1
- PERIOD 2
- △ PERIOD 3
- ◇ PERIOD 4
- × PERIOD 5

days, the differences are not randomly distributed. As examples, five such time periods are noted on figure 12. It is noted that period four persisted for only a few days while period five appears to have persisted for a month or more. It is concluded that equation (4) does estimate the seasonal variation, but does not, as expected, estimate the shorter-period temperature variations related to short-period transient phenomena.

Figure 13 shows the differences as a function of 10-minute-latitude intervals. These differences appear to be randomly distributed about zero difference. The numbered short-time-period samples shown on figure 12 are also indicated on this figure, giving an indication as to the latitudinal extent of these short-period anomalies.

Figure 14 shows the differences as a function of 10-minute longitudinal intervals. Again the differences appear to be randomly distributed about zero difference. Also the numbered short-period samples are shown, giving an indication of the longitudinal



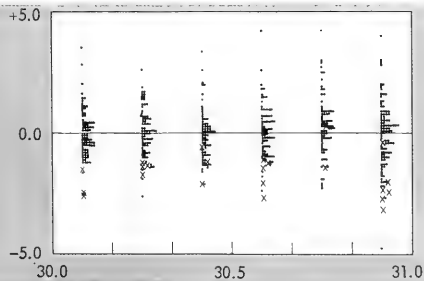
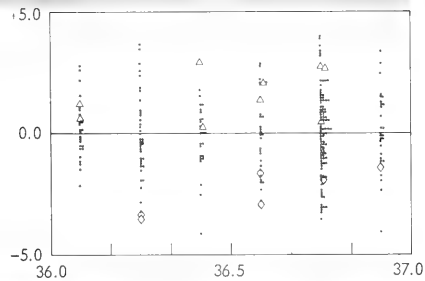


AREA E (1952)

- PERIOD 1
- PERIOD 2
- △ PERIOD 3
- ◇ PERIOD 4
- X PERIOD 5

TEMPERATURE DIFFERENCE,  
DEGREES F

AREA C (1953)

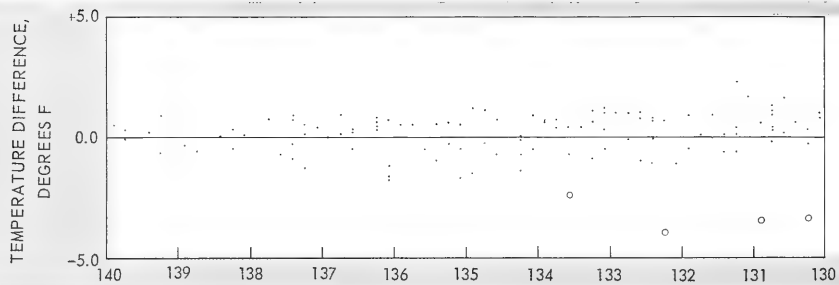


AREA B (1954)

LATITUDE, DEGREES N

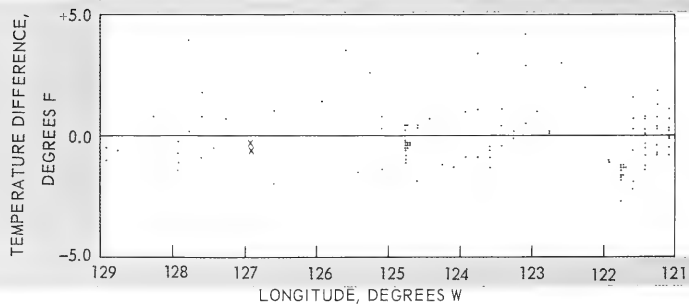
Figure 13. Differences between observed and computed sea-surface temperature as a function of latitude.

AREA D (1952)



AREA C (1953)

AREA B (1954)



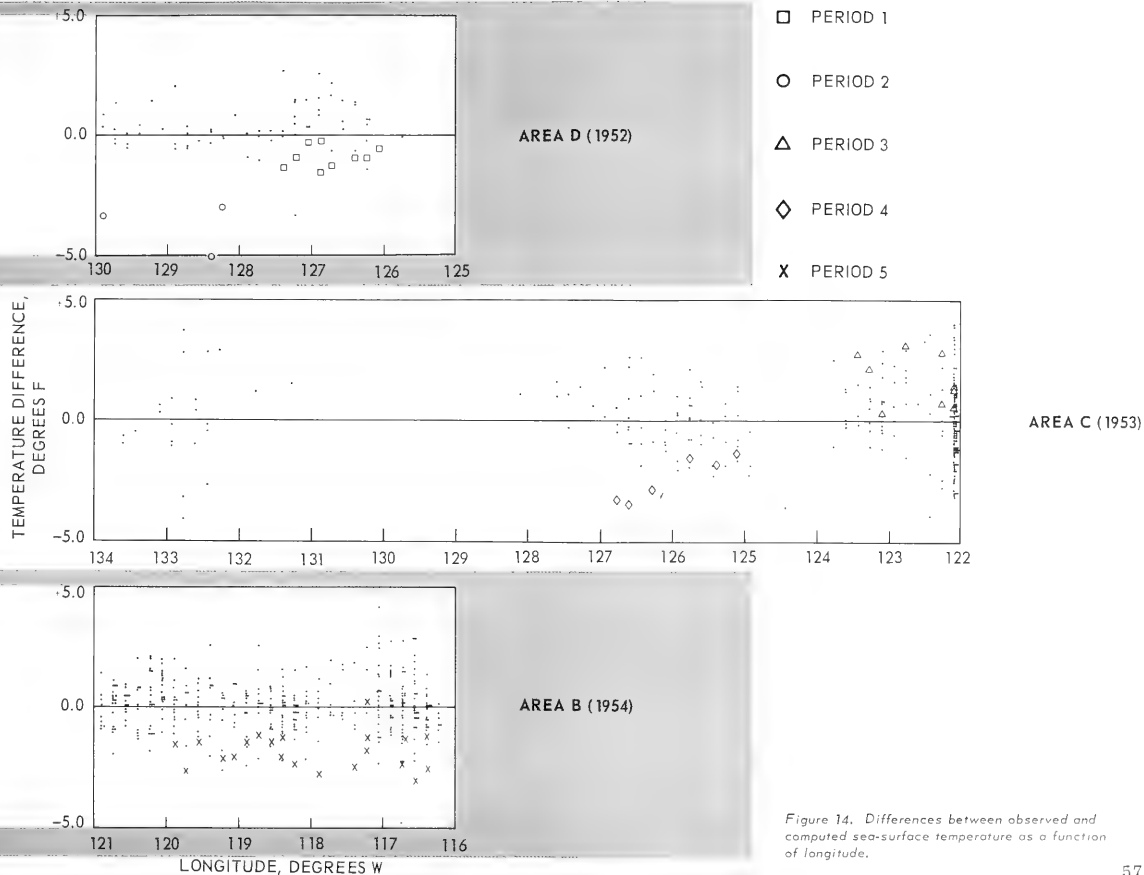


Figure 14. Differences between observed and computed sea-surface temperature as a function of longitude.

extent of the short-period anomalies. The spatial extent of the anomalies, as shown by figures 13 and 14, may cover 1 degree of latitude and several degrees of longitude. For example, numbered sample five shows that the observed temperature was lower than the regression estimate over the 1 degree of latitude and over approximately 4 degrees of longitude for a time period of about 1 month.

For most of the numbered samples the data are located in that part of the latitude strip nearest the coast, where, oceanographically, the variation in temperature is expected to be most erratic, because the number of mechanisms there that affect temperature is greatest.

Figures 15 through 17 are graphical representations of the regression models for these samples. In each figure the regression equation is shown at the top. The narrow strip at the bottom shows the geographical distribution of the observations. The graphs above the strip show the variation of sea-surface tempera-

ture as a function of day-of-year for each of the 10-minute-by-10-minute shaded areas. The dots are measured temperatures within the area. To the right of the strip is a histogram showing the distribution of the observed data sample by months; to the right of the histogram the variation of temperature with longitude for the first day of the month is shown; and to the far right is shown the distribution of the differences between observed and estimated temperature. Pertinent statistics are given at the top of each figure.

A qualitative study of these figures does not reveal any contradictions to generally accepted characteristics of the variation of sea-surface temperature in the areas covered by the analyses.

It is concluded that the regression model, as expressed in equation (4), is a physically acceptable estimator of seasonal and spatial variations in sea-surface temperature in the areas covered by these analyses.

$$T_5 = \alpha_0 + \alpha_1 D + \alpha_2 D^2 + \alpha_3 D^3 + \alpha_4 D^4 + \alpha_5 D^5 +$$

$$\alpha_6 L + \alpha_7 L^2 + \alpha_8 L^3 +$$

$$\alpha_9 G + \alpha_{10} G^2 + \alpha_{11} G^3 +$$

$$\alpha_{12} LD +$$

$$\alpha_{15} GD + \alpha_{16} GD^3 + \alpha_{17} GD^5 +$$

$$\alpha_{18} LG + \alpha_{19} LG^2 + \alpha_{20} LG^3$$

DAY-OF-YEAR  
 LATITUDE  
 LONGITUDE  
 LATITUDE × DAY-OF-YEAR  
 LONGITUDE × DAY-OF-YEAR  
 LATITUDE × LONGITUDE

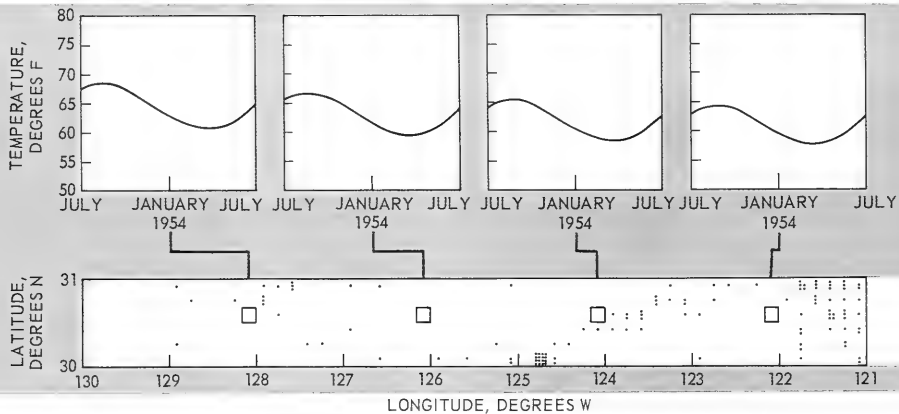
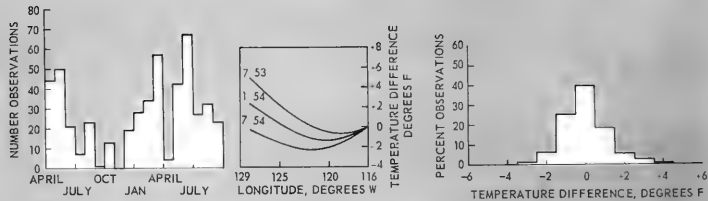
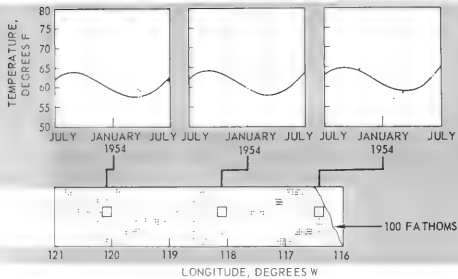


Figure 15. Graphical representation of the regression model for the 30° to 31° latitudinal strip.

<b>STATISTICS</b>	FISCAL YEAR	1954
	NUMBER OBSERVATIONS	497
	PERCENT VARIANCE EXPLAINED	82.7
	MULTIPLE CORRELATION COEFFICIENT	0.91
	STANDARD DEVIATION	1.2°F



$$T_S = a_0 + a_1 D + a_2 D^2 + a_3 D^3 + a_4 D^4 + a_5 D^5 +$$

$$a_6 L + a_7 L^2 + a_8 L^3 +$$

$$a_9 G + a_{10} G^2 + a_{11} G^3 +$$

$$a_{12} LD +$$

$$a_{15} GD + a_{16} GD^3 + a_{17} GD^5 +$$

$$a_{18} LG + a_{19} LG^2 + a_{20} LG^3$$

DAY-OF-YEAR  
 LATITUDE  
 LONGITUDE  
 LATITUDE  $\times$  DAY-OF-YEAR  
 LONGITUDE  $\times$  DAY-OF-YEAR  
 LATITUDE  $\times$  LONGITUDE

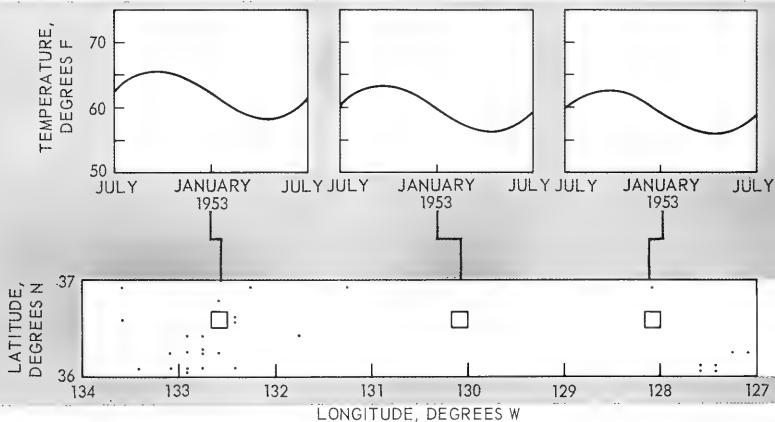
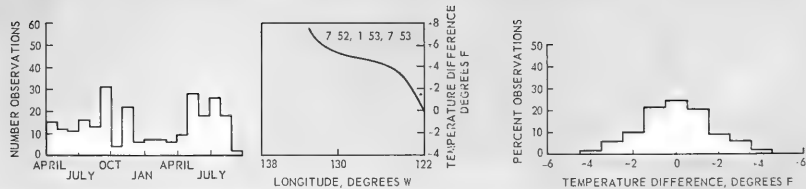
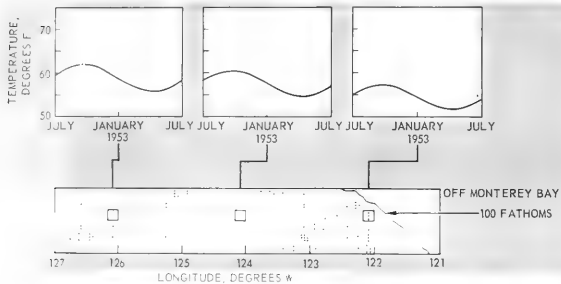


Figure 16. Graphical representation of the regression model for the 36° to 37° latitudinal strip.



<b>STATISTICS</b>	FISCAL YEAR	1953
	NUMBER OBSERVATIONS	251
	PERCENT VARIANCE EXPLAINED	82.7
	MULTIPLE CORRELATION COEFFICIENT	0.91
	STANDARD DEVIATION	1.6 F



$$T_s = \sigma_0 + \sigma_1 D + \sigma_2 D^2 + \sigma_3 D^3 + \sigma_4 D^4 + \sigma_5 D^5 + \sigma_6 L + \sigma_7 L^2 + \sigma_8 L^3 + \sigma_9 G + \sigma_{10} G^2 + \sigma_{11} G^3 + \sigma_{12} LD + \sigma_{15} GD + \sigma_{16} GD^3 + \sigma_{17} GD^5 + \sigma_{18} LG + \sigma_{19} LG^2 + \sigma_{20} LG^3$$

DAY-OF-YEAR

LATITUDE

LONGITUDE

LATITUDE × DAY-OF-YEAR

LONGITUDE × DAY-OF-YEAR

LATITUDE × LONGITUDE

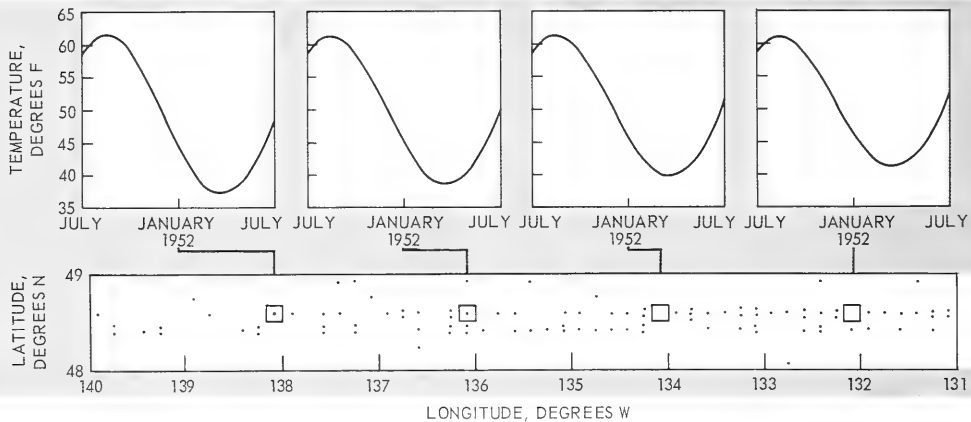
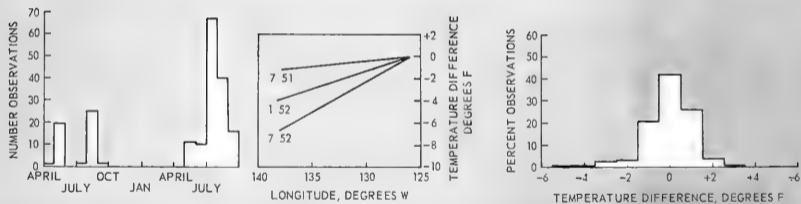
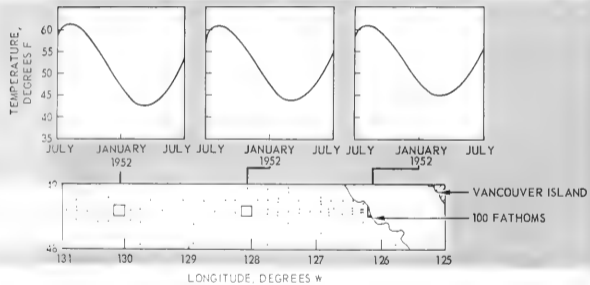


Figure 17. Graphical representation of the regression model for the 48° to 49° latitudinal strip.

<b>STATISTICS</b>	FISCAL YEAR	1952
	NUMBER OBSERVATIONS	171
	PERCENT VARIANCE EXPLAINED	25.6
	MULTIPLE CORRELATION COEFFICIENT	0.98
	STANDARD DEVIATION	1.1 F





## DISCUSSION

Several interesting observations are suggested by a study of the 21 regression analyses considered above.

### Time/Space Consistency

Perhaps the most obvious is the consistency in the statistical parameters with respect to both year and area. In any analysis technique this consistency is important. If the results from year to year and area to area varied widely, then the statistical model would have little predictive potential. It is noted that: (1) The multiple-regression correlation coefficients vary from 0.84 to 0.99, with 50 percent of the coefficients in the 0.88-to-0.91-percent interval. (2) The percent of variance explained by regression varied from 71 to 97 percent, with 50 percent in the 79-to-86-percent interval. (3) The standard deviations varied from 0.9 to 1.9°F, with over 50 percent in the 1.1-to-1.5°F interval. The median standard deviation was 1.2°F. Since the surface-temperature data were obtained from bathythermograms, the data have an instrumental error from 0.5° to perhaps 1.0°F, the magnitude selected depending upon the reader's personal feeling regarding bathythermogram accuracy. The instrumental error represents the noise in the data. The difference between the instrumental error and the standard deviation, about 0.5 to 1.0°F, could possibly be a systematic variation not considered in the

model. The model does not include the diurnal variation, a systematic variation of about this magnitude. It is anticipated that the inclusion of this variable would decrease the standard deviation to a value near the instrumental error. Thus, it appears reasonable to suggest that a simple statistical model, such as equation (4), using bathythermogram data, will describe the seasonal and spatial variation of sea-surface temperatures to one standard deviation of something less than  $1^{\circ}\text{F}$ .

## Time/Space Distribution

The temporal and spatial distribution of the observations is of interest. A study of the distribution of the data (fig. 5) suggests that temporally the distribution is the most unsatisfactory in Area E, since no observations were made from October to March; and that spatially it is most unsatisfactory in Area C, since in one 10-minute square, near the 100-fathom contour, 32, 75, and 87 observations were taken in 1952, 1953, and 1954, respectively. The observations represent 12, 30, and 30 percent of the total data taken in their respective years.

An examination of the variation of the statistical measures shown in figure 6 suggests that the data distributions are not unsatisfactory, as originally thought, but are quite satisfactory. An examination of the regression model supports this contention. The model for the day-to-day variation is a fifth-degree polynomial. If this model truly represents the seasonal variation of surface temperature, then it is necessary only to have observations

during oceanographic summer and winter, since in order for the model to fit the data taken during the seasonal extremes, it must, by nature of the model, fit the data taken during the periods of spring warming and autumn cooling. Thus, the taking of additional data during the latter seasons neither adds to nor detracts from the results obtained. Similar reasoning applies to the spatial distribution of data. If a third-degree polynomial describes the longitudinal variation of surface temperature, then it is necessary only to have a few observations distributed over the area to determine the shape of the polynomial. Again, the taking of additional observations is unnecessary. In support of this observation, an additional fit to the data taken in 1954 was made to the same data shown in figure 6, except that only 22 observations picked randomly from the original 87 observations, taken in the 1-degree square under consideration, were used. The results follow:

<u>Area C (1954)</u>	<u>Data Set 1</u>	<u>Data Set 2</u>
Number of observations	286	221
$100R^2$ , percent variance explained	75.0	74.9
$R$ , multiple correlation coefficient	0.87	0.87
$\sigma$ , standard deviation in degrees F	1.8	1.7

The almost identical results of the two regression analyses suggest that the additional 65 observations used in the first analysis did not contribute any additional information, and that the abnormal spatial distribution of data did not distort the statistical analysis.

The implication is important. A sampling interval based on a model representing the distribution of the variable should be used in place of the fixed time interval employed in the classical manner with an area grid. The oceanographic problem becomes one of searching for adequate models. It is believed that adequate models can be developed for many oceanic areas through examination of historical data and present knowledge of oceanic dynamics.

## Data Screening

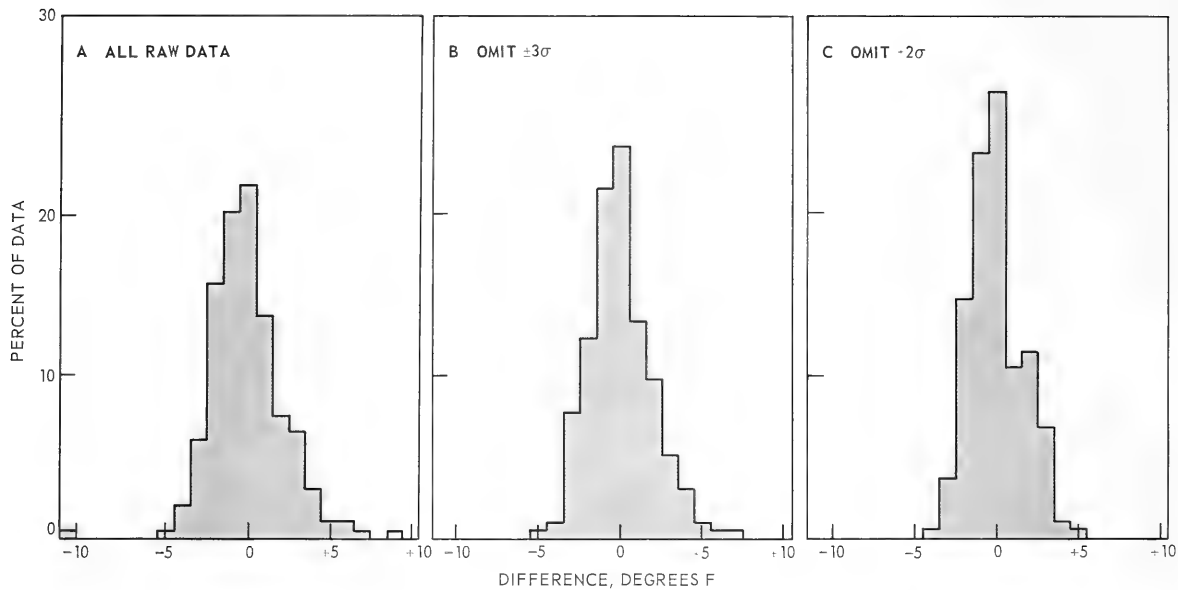
It may be possible to use regression techniques to identify and eliminate erroneous data from data samples. Saur<sup>13</sup> discusses this problem. Erroneous measurements are particularly troublesome when the conventional space/time methods of data summarization are used, since the number of observations in any given cell is generally small, and erroneous data can often distort arithmetic means and standard deviations of discrete samples. Since the average for any given cell is independent of the average in any other space/time cell, the contribution of an erroneous observation tends to be maximized. Frequently a biased average results that must be compensated for in some manner, generally subjective, in the subsequent contour analysis. Thus, the problem of editing out erroneous observations is of considerable importance. Provided the regression model is reasonably valid, the regression technique may offer an effective, and objective, method for editing



out the erroneous observations. Since space and time are treated simultaneously in the regression model rather than separately as in the conventional space/time averaging approach, a single observation is not overly weighted in the averaging process.

Data taken in the  $30^{\circ}$  N-latitude strip for fiscal year 1950 will be used to illustrate this editing technique. The original set of raw data contained 199 observations taken in the 18-month period centered on fiscal year 1950. The left-hand section of figure 18 shows the statistical results obtained by fitting equation (4) to this complete data set. In addition, a histogram of the differences between the sea-surface temperature obtained from the regression equation and the observed value is presented. It is noted that there are three differences greater than  $\pm 3$  standard deviations and eight differences greater than  $\pm 2$  standard deviations. The original data for these 11 observations were examined and in all cases real errors were found. The correspondence suggests that gross errors in data sets may be detected by means of a regression model and eliminated by rejecting data whose differences are greater than some multiple of the standard deviation. The center and right-hand sections present results obtained by rejecting data whose differences were greater than  $\pm 3$  and  $\pm 2$  standard deviations, respectively. If, for one reason or another, it is not desirable to eliminate the erroneous data, the regression technique affords a method of rapidly identifying the data badly in error. Once identified, data can be examined, corrected, and salvaged for subsequent analysis.

Several such analyses were made on different data sets and in all cases the data identified by large differences were found to contain real errors.



	A	B	C
NUMBER OBSERVATIONS	199	196	191
PERCENT VARIANCE EXPLAINED	74.3	78.7	84.0
MULTIPLE CORRELATION COEFFICIENT	0.86	0.89	0.92
STANDARD DEVIATION	2.3°F	1.9°F	1.6°F

Figure 18. An example of the use of a regression equation to edit raw sea-surface temperature data.

## Anomaly Detection

Regression models could be used as anomaly detectors. In this application a model, such as equation (4), could be used to remove the systematic variations in latitude, longitude, and day-of-year. Through a study of the differences between observation and regression (anomalies), information on nonsystematic and other systematic space/time changes would be obtained. The anomalous variations could be examined in terms of causes and mechanisms. This application of regression models was alluded to in the discussion of figures 12 to 14, in which it was noted that the differences revealed short-period, small-area, nonsystematic anomalies.

An additional example of this use of regression models may be found in the differences associated with the data used in figure 9. It is well known that upwelling of cold water occurs off the coast of California from about  $30^{\circ}\text{N}$  to  $45^{\circ}\text{N}$  from March to July. The phenomenon is associated with the north-northwest winds that prevail off the coast of California during these months. The upwelling results in summer and autumn surface temperatures considerably lower than those expected on a seasonal basis alone. The colder-than-expected surface temperatures are centered in the vicinity of  $35^{\circ}\text{N}$  and  $40^{\circ}\text{N}$ . Figure 19 shows the differences for July between the observed surface temperatures and surface temperatures computed from regression. A negative sign means the observed temperature was lower than estimated. The anomalous effect of upwelling on surface temperature is obvious. Parenthetically it is noted that since it is known that these anomalies are the result of a north-northwest wind pattern, they could, in principle, be removed by introducing the wind vector as an independent variable in the regression equation.

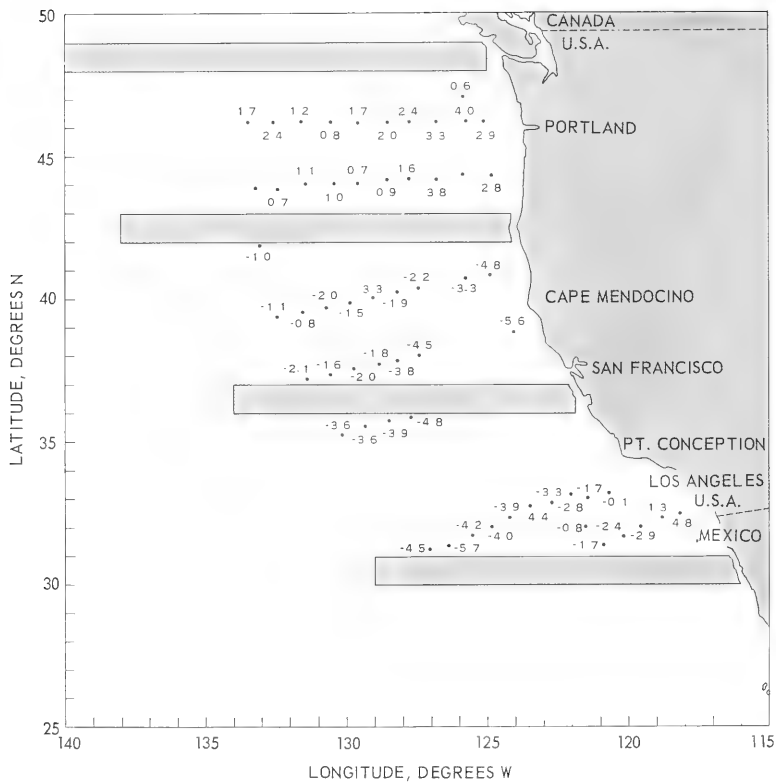


Figure 19. July differences between observed and computed sea-surface temperatures illustrating the use of a regression model for anomaly detection.

## Summarization of Historical Data

Regression techniques could also supply a new approach to summarizing historical sea-surface temperature data. It is assumed that it is possible to develop a realistic regression model. In this study the model was derived by a pseudo-objective method involving a trial-and-error approach to determining the interaction terms. Before regression modeling can become completely satisfactory as a method of sea-surface temperature summarization, it will be necessary to develop objective methods of determining the main effects and their interactions. If it is assumed that a physically acceptable regression model can be developed, it might still be asked how such a model can yield estimates of the day-to-day and location-to-location sea-surface temperature. An unpublished NEL study suggests that an 8-to-10-year time-series record of sea-surface temperatures is long enough to produce reliable long-term estimates that are independent of the time period of observation. Thus, if there is available a 10-year record of sea-surface temperatures covering the area for which the regression model was developed, the regression equation can be fitted to each year of data to provide 10 yearly sets of regression coefficients. A sample could then be drawn from each yearly distribution and combined into a composite sample which could be considered a sample drawn from the 10-year time period. The regression equation could then be fitted to this composite sample

to produce a regression equation that would represent the temporal and spatial variation of sea-surface temperature independently of year-to-year effects. This would be analogous to the climatic charts of the meteorologist.

In the absence of any other information this regression equation will give the best estimate of sea-surface temperature and its variance for any latitude, longitude, and day-of-year. Year-to-year variations, which of course do exist, are neglected. To improve on this estimate it is necessary to consider the year-to-year variation. This might be done as follows: Assume that some observations have been made during the past several months over the area. The composite surface could be adjusted to the new data by a least-squares adjustment of the origin of the regression equation to pass the surface through the currently observed data. The adjusted surface will then be the best estimate of sea-surface temperature for any future day.

For any particular day of the year contour charts, such as illustrated by figure 20, could if desired be prepared. This particular chart was prepared, using equation (4) fitted to the data taken in fiscal year 1950 in the large area, for 8 November 1950. A comparison of this chart with a chart prepared using more classical techniques shows excellent agreement.

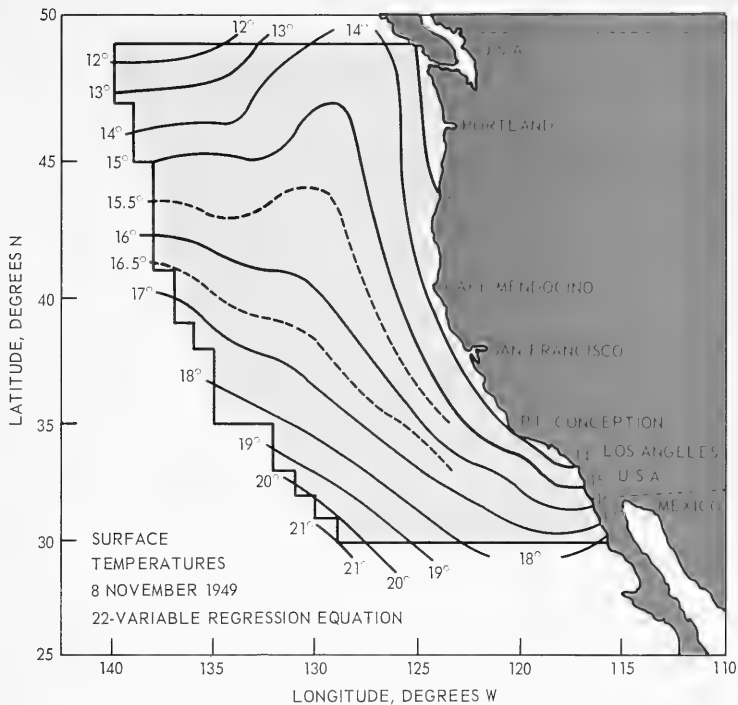


Figure 20. Sea-surface temperature contours for 8 November 1949 computed from a regression model.





## CONCLUSIONS

1. A regression model considering latitude, longitude, and day-of-year as the independent variables together with empirically determined interaction terms, was found capable of estimating the seasonal variation of sea-surface temperatures off the west coast of the United States, in water depths greater than 100 fathoms, to one standard deviation of something less than  $1^{\circ}$  F.

2. From both a statistical and a physical viewpoint relatively simple regression models have a considerable potential as estimators of seasonal and spatial variation in sea-surface temperature.

3. The analysis suggests that more information than previously suspected can be obtained from a given number of observations provided realistic regression models can be developed. This has important implications with regard to sampling. A sampling interval based on the model can be used in place of the fixed time interval employed in the classical manner with an area grid. The oceanographic problem becomes one of searching for adequate models. It is indicated such models can be derived for many ocean areas from the present archive of oceanic temperature data.

4. On the assumption that the regression model is reasonably valid, the regression technique has the potential of being an effective, and objective, method for identifying and editing raw temperature data for erroneous observations.

5. When used to remove the seasonal and spatial variation in a set of sea-surface temperature data, a regression model, such as discussed in this study, may be used to detect and isolate temperature anomalies.

6. Finally, this study suggests regression techniques may be used as a new approach to summarizing archived sea-surface temperature data that is more objective and amenable to computer usage than presently used methods.



## RECOMMENDATIONS

As an outgrowth of this study the following investigations are indicated:

1. Determination of how large an area can be covered by one regression surface.
2. Determination of optimum sampling procedures and sample sizes.
3. Extension of regression models to include depth as an independent variable.
4. Application of regression modeling techniques to describing the distribution of other oceanic parameters, such as salinity.



## REFERENCES

1. Navy Electronics Laboratory Report 965, Underwater Sound Velocity Studies Off the West Coast of North America: Near-Surface Sound-Velocity Distribution from 30°-32°N and 120°-123°W, by E. R. Anderson, 2 May 1960
2. Navy Electronics Laboratory Report 1256, Sea Surface Temperature Estimation; An Empirical Study of the Effect of Missing Data on Regression and Autocorrelation Analysis of Time Series of Data, by C. J. Van Vliet, 5 January 1965
3. Krümmel, O., Handbuch der Ozeanographie, Bd. 1, Stuttgart, J. Engelhorn, 1907
4. Schott, G., Geographie des Atlantischen Ozeans, Hamburg, Boysen, 1926
5. Schott, G., Geographie des Indischen und Stillen Ozeans, Hamburg, Boysen, 1935
6. Böhnecke, G., Temperatur, Salzgehalt und Dichte an der Oberfläche des Atlantischen Ozeans, Berlin, Gruyter, 1936 (Deutsche Atlantische Expedition auf dem Forschungs- und Vermessungsschiff "Meteor", 1925-1927. Bd. V)
7. U. S. Weather Bureau, Atlas of Climatic Charts of the Oceans, Washington, D. C., 1938 (Weather Bureau No. 1247)

8. U. S. Hydrographic Office, World Atlas of Sea Surface Temperatures, Washington, D. C., 1944 (Hydrographic Office Publication No. 225)
9. Chief of Naval Operations NAVAER 50-1C-528, U.S. Navy Marine Climatic Atlas of the World, v. 1: North Atlantic Ocean.  
1 November 1955
10. Chief of Naval Operations NAVAER 50-1C-529, U.S. Navy Marine Climatic Atlas of the World, v. 2: North Pacific Ocean.  
1 July 1956
11. Bureau of Commercial Fisheries Biological Laboratory, Stanford, Monthly Mean Charts, Sea Surface Temperature, North Pacific Ocean, April 1962 (U.S. Fish and Wildlife Service Circular 134)
12. Pyle, R. L., Sea Surface Temperature Regime in the Western North Atlantic 1953-1954, American Geographical Society, 1962 (Serial Atlas of the Marine Environment. Folio 1)
13. Saur, J. F. T., "A Study of the Quality of Sea Water Temperatures Reported in Logs of Ships' Weather Observations," Journal of Applied Meteorology, v.2, p. 417-425, June 1963

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D		
1. SECURITY CLASSIFICATION OF TITLE, BODY OF ABSTRACT AND INDEXING ANNOTATION MUST BE ENTERED WHEN THE OVERALL REPORT IS CLASSIFIED		
2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED		
2b. GROUP UNCLASSIFIED		
3. REPORT TITLE SEA-SURFACE TEMPERATURE ESTIMATION		
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Research and Development Report January 1960 to June 1964		
5. AUTHOR(S) (First name, middle initial, last name)		
E. R. Anderson		
6. REPORT DATE 9 March 1967	7a. TOTAL NO. OF PAGES 84	7b. NO. OF REFS 13
8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S) 1427	
b. PROJECT NO. SR 104 03 01 Task 0586 (NEL L40571)	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned (list report))	
c.		
d.		
10. DISTRIBUTION STATEMENT Distribution of this document is unlimited.		
11. SUPPLEMENTARY NOTES 12. SPONSORING MILITARY ACTIVITY Naval Ship Systems Command Department of the Navy		
13. ABSTRACT A regression model considering latitude, longitude, and day-of-year as the independent variables proves capable of estimating the seasonal variation of sea-surface temperature off the west coast of the United States to a standard deviation of less than 1° F. Time/space interpolation promises to increase the yield of information from a given number of observations. Applications are seen in data editing and summarizing of archived information.		

DD FORM 1473 (PAGE 1)

5/76 0101-807-6601

UNCLASSIFIED

Security Classification

UNCLASSIFIED

Security Classification

KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Ocean Surface - Temperature - Mathematical Analysis						

DD FORM 1473 (BACK)  
1 NOV 68 (PAGE 2)

UNCLASSIFIED

Security Classification







Navy Electronics Lab., San Diego, Calif.  
Report 1427

SEA-SURFACE TEMPERATURE ESTIMATION, by E. R. Anderson,  
84 p., 9 Mar 67.

UNCLASSIFIED

A regression model considering latitude, longitude, and day-of-year as the independent variables proves capable of estimating the seasonal variation of sea-surface temperature off the west coast of the United States to a standard deviation of less than 1°F. Time/space interpolation promises to increase the yield of information from a given number of observations. Applications are seen in data editing and summarizing of archived information.

I. Ocean Surface -  
Temperature - Mathematical  
Analysis

I. Anderson, E. R.

SR 104 03 01, Task 0586  
(NEL L40571)

This card is UNCLASSIFIED

Navy Electronics Lab., San Diego, Calif.  
Report 1427

SEA-SURFACE TEMPERATURE ESTIMATION, by E. R. Anderson,  
84 p., 9 Mar 67.

UNCLASSIFIED

A regression model considering latitude, longitude, and day-of-year as the independent variables proves capable of estimating the seasonal variation of sea-surface temperature off the west coast of the United States to a standard deviation of less than 1°F. Time/space interpolation promises to increase the yield of information from a given number of observations. Applications are seen in data editing and summarizing of archived information.

I. Ocean Surface -  
Temperature - Mathematical  
Analysis

I. Anderson, E. R.

SR 104 03 01, Task 0586  
(NEL L40571)

This card is UNCLASSIFIED

Navy Electronics Lab., San Diego, Calif.  
Report 1427

SEA-SURFACE TEMPERATURE ESTIMATION, by E. R. Anderson,  
84 p., 9 Mar 67.

UNCLASSIFIED

A regression model considering latitude, longitude, and day-of-year as the independent variables proves capable of estimating the seasonal variation of sea-surface temperature off the west coast of the United States to a standard deviation of less than 1°F. Time/space interpolation promises to increase the yield of information from a given number of observations. Applications are seen in data editing and summarizing of archived information.

I. Ocean Surface -  
Temperature - Mathematical  
Analysis

I. Anderson, E. R.

SR 104 03 01, Task 0586  
(NEL L40571)

This card is UNCLASSIFIED

Navy Electronics Lab., San Diego, Calif.  
Report 1427

SEA-SURFACE TEMPERATURE ESTIMATION, by E. R. Anderson,  
84 p., 9 Mar 67.

UNCLASSIFIED

A regression model considering latitude, longitude, and day-of-year as the independent variables proves capable of estimating the seasonal variation of sea-surface temperature off the west coast of the United States to a standard deviation of less than 1°F. Time/space interpolation promises to increase the yield of information from a given number of observations. Applications are seen in data editing and summarizing of archived information.

I. Ocean Surface -  
Temperature - Mathematical  
Analysis

I. Anderson, E. R.

SR 104 03 01, Task 0586  
(NEL L40571)

This card is UNCLASSIFIED

## INITIAL DISTRIBUTION LIST

CHIEF OF NAVAL MATERIAL  
MAT 0331  
COMMANDER, NAVAL SHIP SYSTEMS COMMAND  
SHIPS 0343  
SHIPS 0355  
SHIPS 1610  
SHIPS 1622B  
SHIPS 1633  
SHIPS 2021 (2)  
SHIPS 20413  
COMMANDER, NAVAL AIR SYSTEMS COMMAND  
AIR 5330  
AIR 5401  
AIR 600  
COMMANDER, NAVAL ORDNANCE SYSTEMS COMMAND  
ORD 53C  
ORD 0322  
ORD 0332  
ORD 0332  
COMMANDER, NAVAL FACILITIES ENGINEERING  
COMMAND  
FAC 42310  
COMMANDER, NAVAL ELECTRONIC SYSTEMS COMMAND  
TECHNICAL LIBRARY  
COMMANDER, NAVAL SHIP ENGINEERING CENTER  
CODE 6120  
CODE 6170  
CODE 6179B  
CODE 6179C03  
CODE 8360  
CHIEF OF NAVAL PERSONNEL  
PERS 11B  
CHIEF OF NAVAL OPERATIONS  
OP-385  
OP-311  
OP-312P  
OP-322C  
OP-345  
OP-377  
OP-3774  
OP-701  
OP-702C  
OP-71  
OP-713  
OP-715 (10)  
OP-716  
OP-885  
OP-9224C1  
CHIEF OF NAVAL RESEARCH  
CODE 416  
CODE 418  
CODE 427  
CODE 439  
CODE 466  
CODE 468  
COMMANDER IN CHIEF  
US PACIFIC FLEET  
CODE 93  
US ATLANTIC FLEET  
COMMANDER OPERATIONAL TEST AND EVALUATION  
FORCE  
KEY WEST TEST AND EVALUATION DETACHMENT  
EMPTY COMMANDER OPERATIONAL TEST AND  
EVALUATION FORCE, PACIFIC  
COMMANDER CRUISER-DESTROYER FORCE  
US PACIFIC FLEET  
CODE 425  
US ATLANTIC FLEET  
COMMANDER SUBMARINE FORCE  
US PACIFIC FLEET  
CODE 21  
US ATLANTIC FLEET  
COMMANDER ANTISUBMARINE WARFARE FORCE  
US PACIFIC FLEET  
COMMANDER FIRST FLEET  
COMMANDER SECOND FLEET  
COMMANDER TRAINING COMMAND  
US PACIFIC FLEET  
US ATLANTIC FLEET  
OFFICE OF THE OCEANOGRAPHER OF THE NAVY  
COMMANDER OCEANOGRAPHIC SYSTEM PACIFIC  
COMMANDER SUBMARINE DEVELOPMENT GROUP TWO  
COMMANDER SERVICE FORCE, US ATLANTIC FLEET  
COMMANDER, DESTROYER DEVELOPMENT GROUP,  
PACIFIC  
COMMANDER FLEET AIR WINGS, ATLANTIC FLEET

NAVAL AIR DEVELOPMENT CENTER  
LIBRARY  
NAVAL MISSILE CENTER  
TECHNICAL LIBRARY  
PACIFIC MISSILE RANGE  
CODE 3255  
NAVAL AIR TEST CENTER  
WEAPONS SYSTEMS TEST DIVISION  
NAVAL ORDNANCE TEST STATION  
CHINA LAKE  
CODE 713  
PASADENA ANNEX  
LIBRARY  
FLEET COMPUTER PROGRAMMING CENTER  
ATLANTIC  
TECHNICAL LIBRARY  
NAVAL WEAPONS LABORATORY  
KEL  
LIBRARY  
PEARL HARBOR NAVAL SHIPYARD  
CODE 2468  
PORTSMOUTH NAVAL SHIPYARD  
CODE 242L  
SAFET SOUND NAVAL SHIPYARD  
CODE 246  
SAN FRANCISCO NAVAL SHIPYARD  
HUNTERS POINT DIVISION  
PHILADELPHIA NAVAL SHIPYARD  
CODE 249C  
CHARLESTON NAVAL SHIPYARD  
CODE 240  
NAVAL RADIOLOGICAL DEFENSE LABORATORY  
CODE 222A  
DAVID TAYLOR MODEL BASIN  
LIBRARY  
NAVY MINE DEFENSE LABORATORY  
CODE 716  
NAVAL TRAINING DEVICE CENTER  
HUMAN FACTORS LABORATORY  
CODE 365H  
NAVY UNDERWATER SOUND LABORATORY  
LIBRARY  
CODE 905  
CODE 930  
ATLANTIC FLEET ASW TACTICAL SCHOOL  
LIBRARY  
NAVY MARINE ENGINEERING LABORATORY  
CODE 257  
NAVAL CIVIL ENGINEERING LABORATORY  
LIBRARY  
NAVY RESEARCH LABORATORY  
CODE 2027  
CODE 4320  
CODE 5440  
NAVAL ORDNANCE LABORATORY  
CORONA  
TECHNICAL LIBRARY  
SILVER SPRING, MD.  
DIVISION 730  
DIVISION 880  
NAVY UNDERWATER SOUND REFERENCE LABORATORY  
LIBRARY  
REAR ADMIRAL JUMPET TWO  
FLEET ASW SCHOOL  
TACTICAL LIBRARY  
FLEET SONAR SCHOOL  
NAVAL UNDERWATER WEAPONS RESEARCH AND  
ENGINEERING STATION  
LIBRARY  
OFFICE OF NAVAL RESEARCH BRANCH OFFICE  
PASADENA  
CHIEF SCIENTIST  
SCIENTIFIC LIAISON OFFICER  
SCRIPPS INSTITUTE OF OCEANOGRAPHER  
BOSTON  
CHICAGO  
SAN FRANCISCO  
LONDON  
NAVAL SHIP MISSILE SYSTEMS ENGINEERING  
STATION  
CODE 383  
NAVAL AIR TECHNICAL TRAINING CENTER  
CHIEF OF NAVAL AIR TRAINING  
TRAINING RESEARCH DEPARTMENT

NAVAL PERSONNEL RESEARCH ACTIVITY  
SAN DIEGO  
WEATHER RESEARCH FACILITY  
NAVAL OCEANOGRAPHIC OFFICE  
CODE 1640  
SUPERVISOR OF SHIPBUILDING, US NAVY  
GROTON, CONN.  
CODE 249  
NAVAL POSTGRADUATE SCHOOL  
DEPT. OF ENVIRONMENTAL SCIENCES  
LIBRARY  
FLEET NUMERICAL WEATHER FACILITY  
NAVAL APPLIED SCIENCE LABORATORY  
CODE 9200  
CODE 9832  
NAVAL ACADEMY  
ASSISTANT SECRETARY OF THE NAVY  
(RESEARCH AND DEVELOPMENT)  
NAVAL SECURITY GROUP  
GJ3  
AIR DEVELOPMENT SQUADRON ONE  
YK-1  
SUBMARINE FLOTILLA ONE, US PACIFIC FLEET  
DEFENSE DOCUMENTATION CENTER (30)  
DEPARTMENT OF DEFENSE RESEARCH AND  
ENGINEERING  
WEAPONS SYSTEMS EVALUATION GROUP  
DEFENSE ATOMIC SUPPORT AGENCY  
DOCUMENT LIBRARY SECTION  
NATIONAL OCEANOGRAPHIC DATA CENTER  
CODE 240  
COAST GUARD OCEANOGRAPHIC UNIT  
NATIONAL ACADEMY OF SCIENCES/  
NATIONAL RESEARCH COUNCIL  
COMMITTEE ON UNDERSEA WARFARE  
COAST GUARD HEADQUARTERS  
OSR-2  
ARCTIC RESEARCH LABORATORY  
WOODS HOLE OCEANOGRAPHIC INSTITUTION  
DOCUMENT LIBRARY LO-206  
COAST AND GEODETIC SURVEY  
ROCKVILLE, MD.  
WASHINGTON SCIENCE CENTER - 23  
WASHINGTON, D. C.  
FEDERAL COMMUNICATIONS COMMISSION  
RESEARCH DIVISION  
US WEATHER BUREAU  
DIRECTOR, METEOROLOGICAL RESEARCH  
LIBRARY  
CENTRAL INTELLIGENCE AGENCY  
OSR-30-STANDARD DISTRIBUTION  
NATIONAL SEVERE STORMS LABORATORY  
NATIONAL BUREAU OF STANDARDS  
BOULDER, COLO.  
GEOLOGICAL SURVEY LIBRARY  
DENVER SECTION  
BUREAU OF COMMERCIAL FISHERIES  
LA JOLLA, CALIF.  
TUNA RESOURCES LABORATORY LA JOLLA  
WASHINGTON, D. C.  
BRANCH OF THE FISHERIES  
WOODS HOLE, MASS.  
BIOLOGICAL LABORATORY LIBRARY  
HONOLULU, HAWAII  
FISH AND WILDLIFE SERVICE LIBRARY  
STANFORD, CALIF.  
BIOLOGICAL LABORATORY  
ABERDEEN PROVING GROUND  
TECHNICAL LIBRARY  
ARMY RESEARCH AND DEVELOPMENT ACTIVITY  
ELECTRONIC WARFARE DIVISION  
ELECTRONIC DEPARTMENT  
REDSTONE SCIENTIFIC INFORMATION CENTER  
ARMY MISSILE COMMAND  
DOCUMENT SECTION  
ARMY ELECTRONICS RESEARCH AND DEVELOPMENT  
LABORATORY  
ARMY ELECTRONICS COMMAND  
MANAGEMENT & ADMINISTRATIVE SERVICES DEPT  
AMSEL-RO-MAT  
FRANKFORD ARSENAL  
HUMAN FACTORS LABORATORY  
SHUPA 86400/702-4  
WHITE SANDS MISSILE RANGE  
STEMS-10-E

FORT HUACHUCA  
32D USASAC  
AFS FORCE HEADQUARTERS  
DIRECTOR OF SCIENCE AND TECHNOLOGY  
AFRSTA  
AIR UNIVERSITY LIBRARY  
AULT31-5028  
STRATEGIC AIR COMMAND  
OAST  
AIR FORCE EASTERN TEST RANGE  
ARPC TECHNICAL LIBRARY - MU-155  
AIR PROVING GROUND CENTER  
PEEPS-12  
WRIGHT-PATTERSON AIR FORCE BASE (1)  
SYSTEMS ENGINEERING GROUP (RTD)  
SERIP  
AIR FORCE SECURITY SERVICE  
ESSOESS  
ELECTRONICS SYSTEMS DIVISION  
EST1  
UNIVERSITY OF MICHIGAN  
OFFICE OF RESEARCH ADMINISTRATION  
NORTH CAMPUS  
CODLEY ELECTRONICS LABORATORY  
RADAR AND OPTICS LABORATORY  
UNIVERSITY OF CALIFORNIA-SAN DIEGO  
MARINE PHYSICAL LABORATORY  
SCRIPPS INSTITUTION OF OCEANOGRAPHY (1)  
LIBRARY  
LIBRARY OF MIAMI  
THE MARINE LABORATORY LIBRARY  
MICHIGAN STATE UNIVERSITY  
LIBRARY-DOCUMENTS DEPARTMENT  
COLUMBIA UNIVERSITY  
HUDSON LABORATORIES  
LAMONT GEOLOGICAL OBSERVATORY  
DARTMOUTH COLLEGE  
RADIOPHYSICS LABORATORY  
CALIFORNIA INSTITUTE OF TECHNOLOGY  
JET PROPULSION LABORATORY  
HARVARD COLLEGE OBSERVATORY  
HARVARD UNIVERSITY  
GORDON MCKAY LIBRARY  
LYMAN LABORATORY  
OREGON STATE UNIVERSITY  
DEPARTMENT OF OCEANOGRAPHY  
UNIVERSITY OF WASHINGTON  
DEPARTMENT OF OCEANOGRAPHY  
PHYSICS-OCEANOGRAPHY LIBRARY  
APPLIED PHYSICS LABORATORY  
NEW YORK UNIVERSITY  
DEPARTMENT OF METEOROLOGY AND OCEANOGRAPHY  
TUFTS UNIVERSITY  
INSTITUTE FOR PSYCHOLOGICAL RESEARCH  
OHIO STATE UNIVERSITY  
ANTENNA LABORATORY  
UNIVERSITY OF ALASKA  
GEOGRAPHIC INSTITUTE  
UNIVERSITY OF RHODE ISLAND  
NARRAGANSETT MARINE LABORATORY  
LISINE FISHERY  
YALE UNIVERSITY  
BINGHAM OCEANOGRAPHIC LABORATORY  
FLORIDA STATE UNIVERSITY  
OCEANOGRAPHIC INSTITUTE  
UNIVERSITY OF HAWAII  
HAWAII INSTITUTE OF GEOPHYSICS  
ELECTRICAL ENGINEERING DEPARTMENT  
ASU COLLEGE OF FISH  
DEPARTMENT OF OCEANOGRAPHY  
THE UNIVERSITY OF TEXAS  
DEFENSE RESEARCH LABORATORY  
ELECTRICAL ENGINEERING RESEARCH LABORATORY  
PENNSYLVANIA STATE UNIVERSITY  
ORDNANCE RESEARCH LABORATORY  
STANFORD RESEARCH INSTITUTE  
NAVAL WARFARE RESEARCH CENTER  
MASSACHUSETTS INSTITUTE OF TECHNOLOGY  
ENGINEERING LIBRARY  
LINCOLN LABORATORY  
LIBRARY, A-082  
FLORIDA ATLANTIC UNIVERSITY  
DEPARTMENT OF OCEAN ENGINEERING  
THE JOHNS HOPKINS UNIVERSITY  
APPLIED PHYSICS LABORATORY  
DOCUMENT LIBRARY  
INSTITUTE FOR DEFENSE ANALYSES  
DOCUMENT LIBRARY