Determination of Vertical Thermal Structure from Sea Surface Temperature

PETER C. CHU AND CHENWU FAN

Department of Oceanography, Naval Postgraduate School, Monterey, California

W. TIMOTHY LIU

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

ABSTRACT

A recently developed parametric model by P. C. Chu et al. is used in this paper for determining subsurface thermal structure from satellite sea surface temperature observations. Based on a layered structure of temperature fields (mixed layer, thermocline, and lower layers), the parametric model transforms a vertical profile into several parameters: sea surface temperature (SST), mixed layer depth (MLD), thermocline bottom depth (TBD), thermocline temperature gradient (TTG), and deep layer stratification (DLS). These parameters vary on different timescales: SST and MLD on scales of minutes to hours, TBD and TTG on months to seasons, and DLS on an even longer timescale. If the long timescale parameters such as TBD, TTD, and DLS are known (or given by climatological values), the degree of freedom of a vertical profile fitted by the model reduces to one: SST. When SST is observed, one may invert MLD, and, in turn, the vertical temperature profile with the known long timescale parameters: TBD, TTG, and DLS.

The U.S. Navy's Master Oceanographic Observation Data Set (MOODS) for the South China Sea in May 1932–94 (10 153 profiles) was used for the study. Among them, there are 40 data points collocating and coappearing (same week) with the weekly daytime NASA multichannel SST data in 1986–94. The 40 MOODS profiles were treated as a test dataset. The MOODS dataset excluding the test data is the training dataset, consisting of 10 113 profiles. The training dataset was processed into a dataset consisting of SST, MLD, TBD, TTG, and DLS using the parametric model. SST from the test dataset was used for the inversion based on the known information on TBD, TTG, and DLS. The 40 inverted profiles agreed quite well with the corresponding observed profiles. The rms error is 0.72°C, and the correlation between the inverted and observed profiles is 0.79. This is much better than the simple method of estimating subsurface temperature anomaly from SST anomaly by correlating the two in the training dataset. The possibility of using this method globally is also discussed.

1. Introduction

The most difficult problem in physical oceanography is the lack of in situ observations. With the help of electromagnetic techniques, especially satellite remote sensing, we may obtain global coverage of temporally varying surface data such as sea surface temperature (SST). Can we determine the vertical thermal structure from satellite SST observations? The answer should come from examining the linkage between SST and subsurface thermal structure (Fig. 1).

For a given location this is a linkage between a zero-dimensional variable (SST) and a one-dimensional variable (subsurface thermal structure). The key issue is how to compress a large profile dataset into a small parameter (or coefficient) dataset. The U.S. Navy's Generalized Digital Environmental Model uses several analytical curve fitting functions (submodels) to compress the profile data into a set of coefficients (Teague et al. 1990). The number of coefficients varies among the submodels. For example, the shallow top submodel (0–400 m) contains eight coefficients. The middepth submodel (200–2450 m) contains seven coefficients. Most coefficients represent the feature of the profile.

Recently, a parametric model (Chu et al. 1997a,b, 1999) has been developed for analyzing observed temperature profiles based on a layered structure (mixed layer, thermocline, and deep layer). The output of the parametric model is a set of major physical characteristics of each profile: SST, mixed layer depth (MLD), thermocline bottom depth (TBD), thermocline temperature gradient (TTG), lower layer stratification, and deep layer temperature. The model successfully reproduced Yellow Sea. (Chu et al. 1997a,b) and Beaufort/Chukchi Sea (Chu et al. 1999) historical temperature profiles

Corresponding author address: Dr. Peter C. Chu, Dept. of Oceanography, Naval Postgraduate School, Monterey, CA 93943.
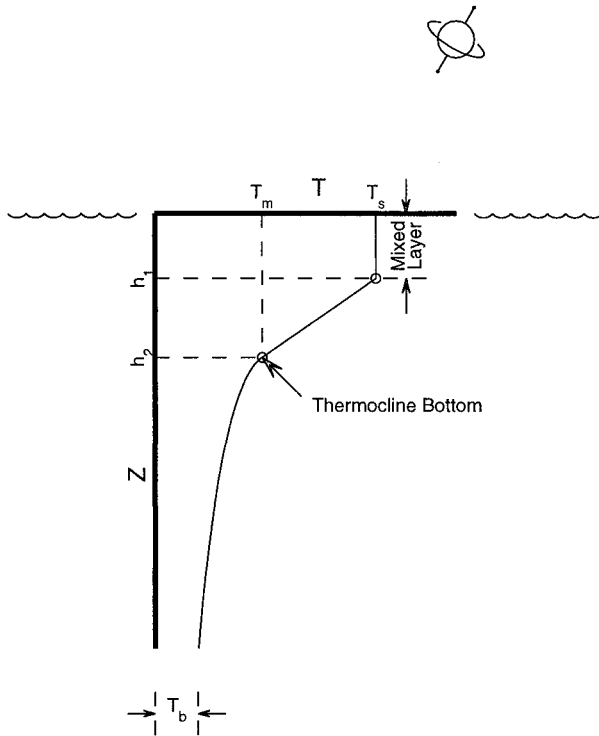E-mail: chu@nps.navy.mil

FIG. 1. A thermal parametric model.

from the Naval Oceanographic Office (NAVO-CEANO)'s Master Oceanographic Observation Data Set (MOODS). Using the parametric model, the inversion of the subsurface thermal structure from satellite SST becomes a relationship between SST and subsurface parameters such as MLD, TBD, and TTG.

2. Thermal parametric model

Keeping the minimal possible degrees of freedom, a temperature profile can be simply depicted by (Fig. 1),

$$T(z) = T_s, \qquad\qquad (0 \geq z \geq -h_1),$$

$$T(z) = T_s + G_{th}(z + h_1), \qquad (-h_1 \geq z \geq -h_2),$$

$$T(z) = T_d + (T_{tb} - T_d) \exp\left[\frac{z_0^w - (z_0 - z - h_2)^w}{H^w}\right],$$
$$(z \leq -h_2), \qquad (1)$$

where T_s, T_{tb}, and T_d are SST, temperature at TBD, and a deep temperature; h_1, h_2, H are MLD, TBD, and a lower layer e-folding scale, respectively; and G_{th} is TTG. The deep temperature T_d is the temperature at the deepest ocean depth such as 5500 m in the climatological data (Levitus and Boyer 1994). For shallow water regions, T_d is, of course, not a real observed value but an extrapolated value to the deepest depth (e.g., 5500 m). We use T_d to keep the data above the bathymetry fitting the parametric model (1).

In this model, the thermocline is featured by a linear profile (constant G_{th}), and the lower layer is characterized by a nonlinear profile. To guarantee $T(z)$ and $T'(z)$ continuous at TBD,

$$T(h_2 + 0) = T(h_2 - 0),$$

$$\frac{dT(h_2 + 0)}{dz} = \frac{dT(h_2 - 0)}{dz} = G_{th}, \qquad (2)$$

we need two additional parameters, z_0 and w.

Differentiation of (1) with respect to z and use of (2) lead to

$$z_0 = \left[\frac{HG_{th}}{w(T_{tb} - T_d)}\right]^{1/(w-1)}. \qquad (3)$$

The parameter w cannot be greater than or equal to 1. Otherwise, z_0 becomes very large and distorts the e-folding decrease of temperature with depth. Also, w cannot be 0. In this study we use $w = 0.5$.

Thus, from a vertical temperature profile we may extract three temperatures (T_S, T_{tb}, T_d), three depths (h_1, h_2, H), and one gradient (G_{th}), seven parameters in total. We require continuity of temperature at TBD, that is,

$$T_S - G_{th}(h_2 - h_1) = T_{tb}. \qquad (4)$$

Therefore, any six of the seven parameters (T_S, T_{tb}, T_d, h_1, h_2, H, G_{th}) determine a vertical profile. Thus, the degrees of freedom of the thermal parametric model are six.

3. The U.S. Navy's MOODS data

The South China Sea (SCS) has a bottom topography (Fig. 2) that makes it a unique semienclosed ocean basin that is seasonally forced by a pronounced monsoon surface wind. Extended continental shelves (less than 100 m deep) exist along the north boundary and across the southwest portion of the basin, while steep slopes with almost no shelf are found along the eastern boundary. The deepest water is confined to an oblate bowl oriented southwest–northeast, centered around 13°N. The maximum depth is around 4500 m.

The MOODS is a compilation of ocean data observed worldwide consisting of (a) temperature-only profiles, (b) both temperature and salinity profiles, (c) sound-speed profiles, and (d) surface temperature (drifting buoy). Due to the shear size (more than six million profiles total for the global ocean) and constant influx of data to NAVOCEANO from various sources, quality control is very important (Chu et al. 1997b, 1998). After quality control, we used a subset of MOODS data in May (1932–94) consisting of 10 153 profiles for the whole SCS (2°–26°N, 99°–123°E).

The temporal and spatial distribution of MOODS data is irregular. Certain periods and areas are very well sampled, while others lack enough observations to gain any meaningful insights. There are some 10–20-day gaps
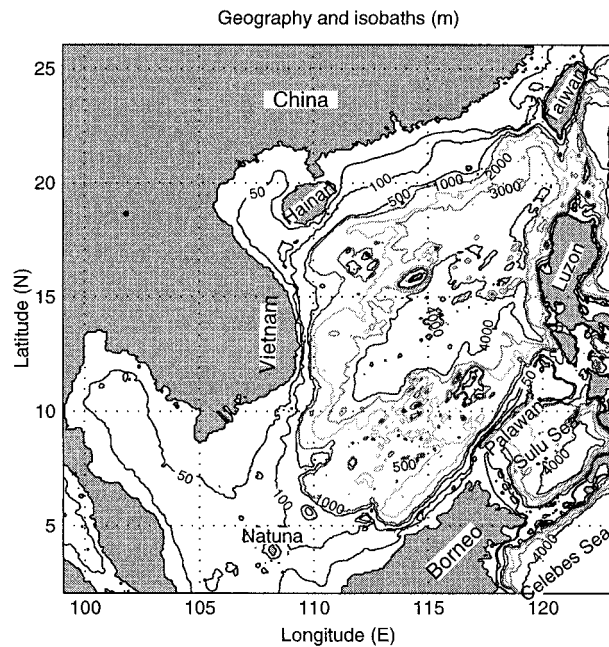
FIG. 2. Geography and isobaths showing the bottom topography of the South China Sea.

FIG. 3. Distribution of MOODS profiles in May 1932–94: (a) historical data and (b) data used for the study.

with no observations in the whole SCS (Chu et al. 1997b). Figure 3 shows the sparsity of profiles in the southeast portion of the study domain and in the coastal region of China continent. Figure 4 indicates a heavily sampled period during the Vietnamese War (1965–69). The maximum number of observations is in 1966 (1230 profiles). The minimum number of observations is in 1935 (three profiles).

In May, for the years 1986–94, there are 40 daytime multichannel SST (MCSST) and MOODS data points that are collocated in the same week, marked by (*) in Fig. 5. Notice that the number of *'s in Fig. 5 is much less than 40. This is due to several data points sharing the same spots. The 40 MOODS profiles were treated as a test dataset. The MOODS dataset excluding the test data is the training dataset, consisting of 10 113 profiles.

4. Mean thermal parameters

We used the May climatological temperature dataset in the SCS with a $1° \times 1°$ horizontal resolution and values located at half-degrees (Levitus and Boyer 1994) at 5500-m depth for T_d, and processed all the training data profiles (10 113) using the parametric model (1). Here, an iteration method illustrated in Chu et al. (1997a) was used. A set of parameters (T_S, T_{tb}, h_2, H, G_{th}, T_d) was obtained for each profile. We averaged the thermal parameters within $1° \times 1°$ grid and took the averaged values as the representative values for the grid cell. These values might not be representative in high gradient and coastal regions. Three types of cells were found in the SCS, representing collocated MCSST and
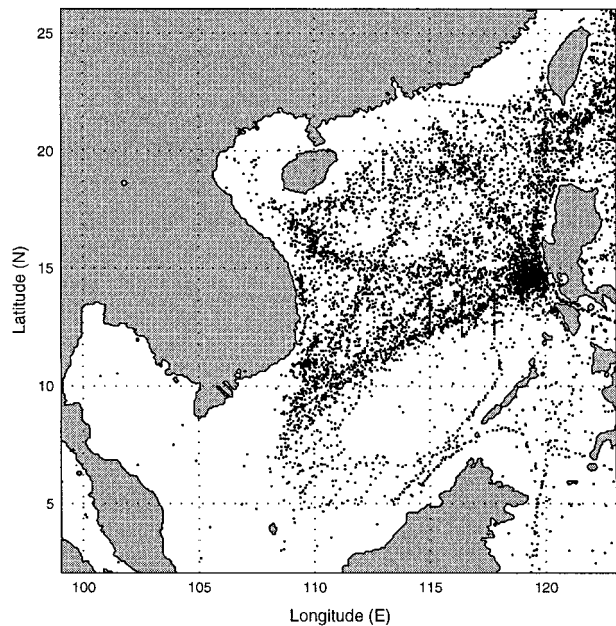
MOODS data points (*), MOODS data points less than 10 (+), and MOODS data points more than 10 (○), as shown in Fig. 5.

Usually early May is the time of SCS summer monsoon onset (Tao and Chen 1987). The thermal parameters obtained from processing the MOODS dataset (May 1932–94) may represent thermal response of SCS to the monsoon onset. Figure 6 shows the mean thermal parameter fields in May averaged over 1932–94. Surface warm water ($\geq29.5°C$) with a maximum temperature 30°C occupies most of the southern half of the SCS (Fig. 6a). The 29.5°C isotherm extended from the southeast corner of the Vietnam coast (near 11°N, 108°E) northeastward to the southwest coast of the Luzon Island
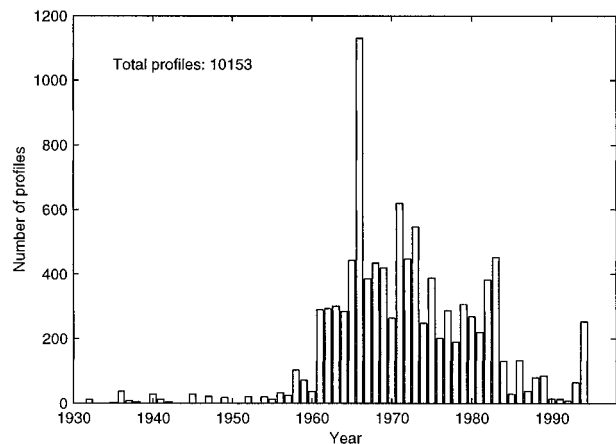

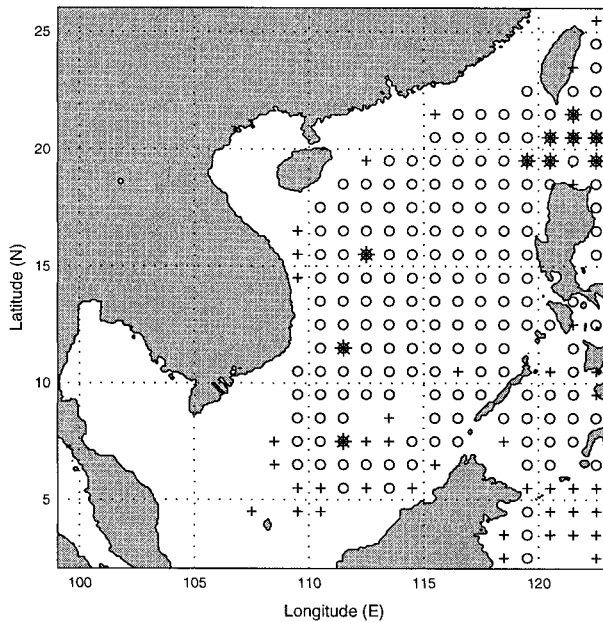FIG. 4. Number of temperature profiles in the South China Sea in May 1932–94.

FIG. 5. Three types of $1° \times 1°$ cells represented by center grid points: collocating and coappearing of MCSST and MOODS data points (*), MOODS data points less than 10 (+), and MOODS data points more than 10 (○).

(near 15°N, 120°E). MLD (h_1) varied from 10 to 40 m and had a latitudinal variation (Fig. 6b). The southern SCS (south of 13°N) was characterized by a deep mixed layer ($h_1 \geq 20$ m) region with a maximum value of 40 m near Palawan Island. This suggests strong turbulent mixing in the southern part of SCS right after the summer monsoon onset. The northern SCS (north of 13°N) has a shallow mixed layer ($h_1 \leq 20$ m) with a depth of 10 m. In the continental shelf regions, TBD (h_2) was quite shallow (≤ 100 m) and in the deep SCS basin, h_2 was deeper (>100 m) with a maximum value of 400 m in the Luzon Strait (Fig.6c), where, however, a weak thermocline (G_{th}) was found with a vertical temperature gradient around 0.04°C m^{-1} (Fig. 6d). Temperature at TBD (T_{tb}) was coldest (12°C) in Luzon Strait and warmest (22°C) in the southern shelf region near Natuna Island (Fig. 6e). The lower layer e-folding thickness (H) represents the stratification in the layer below the thermocline. The smaller the value of H, the stronger the stratification of this layer. In the SCS deep basin, H is quite large (100–200 m), indicating weak stratification below the thermocline (Fig. 6f).

Thus, the SCS thermal response to the monsoon onset can be characterized by a northward advancement of warm surface water, strong turbulent mixing in the southern part with deeper mixed layers, and a relatively uniform deep layer below the thermocline in the SCS deep basin.

5. Regression method

For each grid cell, we compute the mean temperature profile $\overline{T}(z)$ and subtract the mean profile from each

profile in the MOODS training dataset to obtain temperature anomaly $T'(z)$. The simplest method of estimating subsurface $T'(z)$ from SST' is to regress $T'(z)$ with SST':

$$T'(z) = b(z)\text{SST}', \tag{5}$$

where $b(z)$ is the regression coefficient obtained from the training dataset.

6. Multitimescale inverse method

Having the current SST information in the inversion, we need to use the multidecorrelation timescale hypothesis. This hypothesis will reduce the degrees of freedom of the parameter space.

a. Multitimescale hypothesis

The seven parameters vary on different timescales: T_S and h_1 on a short decorrelation timescale; T_{tb}, T_d, h_2, H, and G_{th} on a long decorrelation timescale. The parameters on a long decorrelation timescale are treated as a background dataset, which may be predetermined by historical data. The parameters on the short timescale are determined by the inverse method. If the five parameters on the long timescale are assumed to be predetermined, the degrees of freedom of this model reduces to one. Between the two short timescale parameters T_S and h_1, only one parameter is independent. Usually, we take T_S as the independent parameter. If T_S is given by satellite observation, we can use (4) to determine h_1 and therefore the vertical profile. We call this inverse method the multitimescale method. We use the U.S. Navy's MOODS data for SCS in May to verify this inverse method.

b. Correlation between SST and subsurface parameter anomalies

For each $1° \times 1°$ grid cell, we subtract the mean values (for that cell) from each of the thermal parameters (T_S, T_{tb}, h_1, h_2, H, G_{th}) to obtain the thermal parameter anomalies (T'_S, T'_{tb}, h'_1, h'_2, H', G'_{th}) and to compute the correlation coefficients (Table 1) between T'_S and the subsurface parameter anomalies (h'_1, h'_2, G'_{th}, T'_{tb}, H'). Figure 7 shows the scatter diagram between T'_S and (h'_1, h'_2, G'_{th}, T'_{tb}, H'). Both Table 1 and Fig. 7 indicate that among the subsurface parameters h'_1 has the strongest linear association with T'_S. The significance of the correlation can be evaluated by

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}, \tag{6}$$

which has a t distribution with $n - 2$ degrees of freedom. Here, r is the correlation coefficient, n is the number of samples (10 153). We begin with the usual null hypothesis that there is no linear association between T'_S
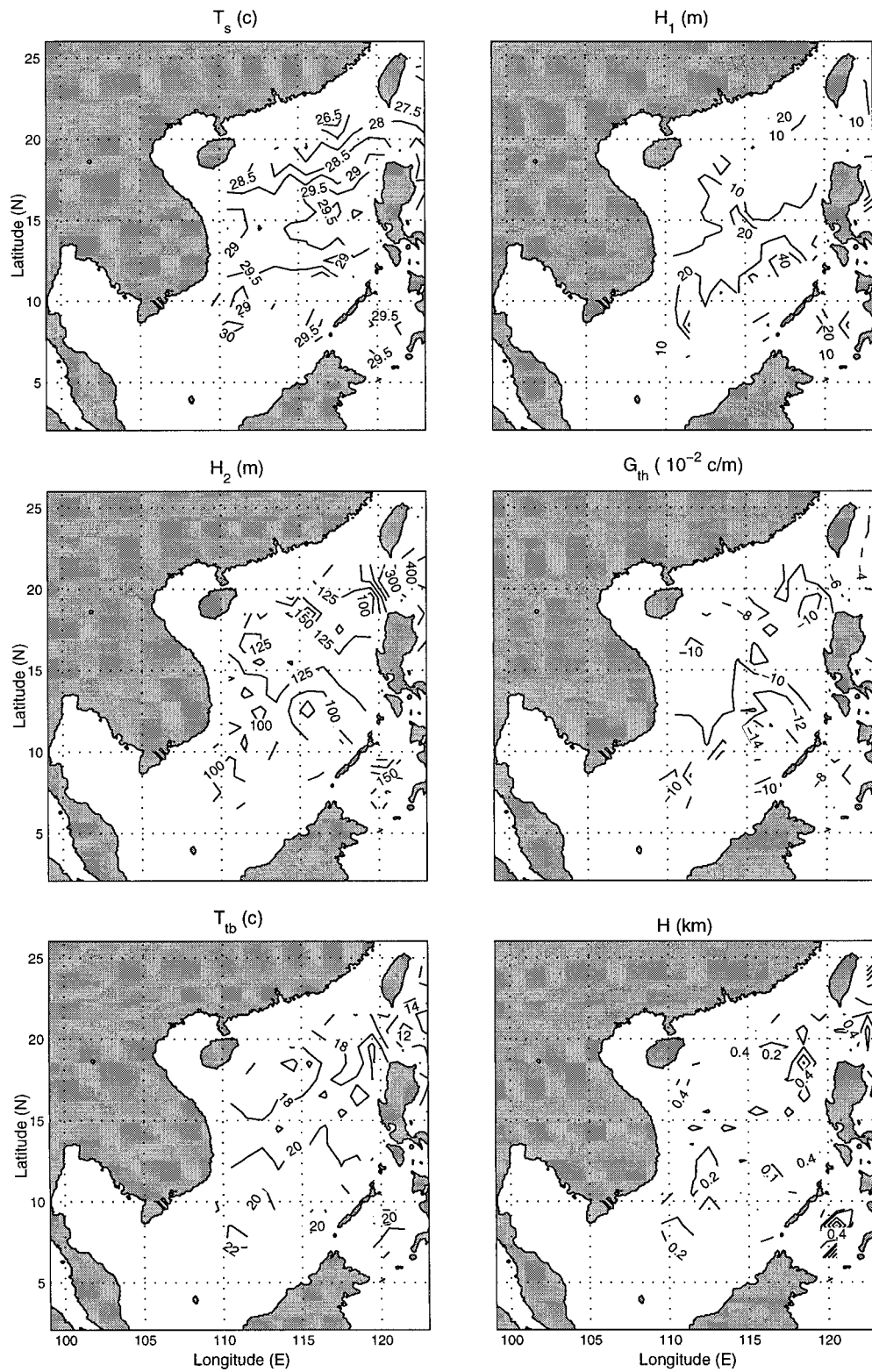
FIG. 6. Horizontal distributions of mean thermal parameters obtained from the training dataset: (a) SST, (b) h_1, (c) h_2, (d) G_{th}, (e) T_{tb}, and (f) H.

TABLE 1. Correlation coefficients with the SST anomaly.

h_1'	h_2'	G_{th}'	T_{tb}'	H'
−0.12	0.03	0.05	−0.01	−0.01

TABLE 2. The t values of various correlation coefficients.

h_1'	h_2'	G_{th}'	T_{tb}'	H'
−15.77	3.98	6.64	−1.33	−1.33

and (h_1', h_2', G_{th}', T_{tb}', H'). The critical t value at significance level of 0.005 ($t_{0.005}$) is 2.576. Three absolute values of t computed by (5) are larger than the critical value (2.576): −15.77 between T_S' and h_1', 6.64 between T_S' and G_{th}', and 3.98 between T_S' and h_2' (Table 2). Thus, we reject the null hypothesis for h_1', h_2', and G_{th}'. Here, we notice that the use of the number of MOODS samples as degree of freedom in the t test gives a very low critical t value (2.576 for $\alpha = 0.005$), which may be caused by the dependence of some MOODS sampling.

Considering various correlation coefficients (Table 1) we may conclude that the correlation between the two short timescale parameters T_S' and h_1' is much stronger than the correlation between T_S' and the long timescale parameters (h_2', G_{th}', T_{tb}', H'). This in turn confirms the multitimescale hypothesis for the SCS thermal parameters.

A negative correlation between T_S' and h_1' might not be true everywhere in the ocean. For example, Tully and Giovando (1963) found it difficult to establish such a relationship at least for a portion of the eastern subarctic Pacific Ocean. However, Chu (1993) pointed out the possibility of such a negative correlation using an analytic ocean mixed layer model for the equatorial Pacific.

c. Inversion

If we take SST from the MOODS test data (40 data points) (Fig. 5) as known values for the short timescale parameter T_S, we may use (a) the background long timescale parameters T_{tb}, h_2, H, and G_{th} to determine h_1 (Figs. 6b–e); or (b) the temperature continuity condition at TBD [Eq. (4)] to determine h_1; plus the May climatological values for T_d (Levitus and Boyer 1994). With all the seven parameters given, we can easily construct vertical profiles $\hat{T}^{(F)}(z)$ by (1).

The 40 inverted profiles agree quite well with the observed profiles; however, the 40 regressed profiles have a larger mismatch with the observed profiles (Fig. 8).

7. Model verification

Any model, including the regression and inverse models presented here, should be verified before claiming any practical usefulness. Usually, the model verification contains two parts: the root-mean-square (rms) error and the correlation coefficient between modeled and observed profiles.

The May climatological profiles (Levitus and Boyer 1994) at the MCSST points are used as the "least-effort" profiles. The standard deviation (SDV) of the climatological profiles (climatological SDV) represents the first criterion for the model validity. If the model rms error is larger than the climatological SDV, the model does not have any practical usefulness. The model becomes valid only if its rms error is smaller than the climatological SDV.

Figure 9a shows the vertical distribution of the model rms and climatological SDV over the whole test data area. The rms errors for both regression and inverse methods increase with depth from the surface to maximum values around 1.8°C near 100-m depth, and then reduce with depth. At all depths except near 100-m depth, the rms errors for the inverse model are much smaller than the rms errors for the regression model, which in turn are smaller than the climatological SDV. The depth of 100 m is approximately the mean TBD (Fig. 6). This implies some difficulty in inverting the temperature at TBD. Overall, the vertically averaged inverse model rms error (around 0.72°C) is smaller than the regression model rms error (around 1.06°C), which in turn is smaller than the climatological SDV (1.51°C).

The correlation coefficients between modeled and observed profiles at all depths represent the second criterion for the model validity. The correlation coefficient for the inverse model varies with depth between 1 and 0.5 and has a vertical mean value of 0.79. Use of (6) leads to $t = 3.559$ for $n = 40$ and $r = 0.5$. This value is much larger than the critical t value (2.576), which means significant correlations at confidence level of 0.005 between the inverted and the observed profiles for all depths.

However, the correlation coefficient for the regression model decreases rapidly from 1 at the surface to 0 near 100-m depth, and then becomes negative below that depth (Fig. 9b), which indicates no significant positive correlations between the regressed and the observed profiles for the sublayer depths.

The small mean rms error (0.72°C) and high positive correlation coefficient (0.79) make this multitimescale inverse method valid for practical use.

8. Limitation of the multitimescale inverse method

The key issue of inverting subsurface thermal structure from SST is to reduce the degree of freedom of the thermal parameter space by multitimescale hypothesis. To apply this method globally, we should first test the validity of this hypothesis. This can be done by the correlation analysis. If correlation between T_S' and h_1' is much stronger than the correlation between T_S' and the other parameters (h_2', G_{th}', T_{tb}', H'), we may confirm the multitimescale hypothesis and use this inverse method
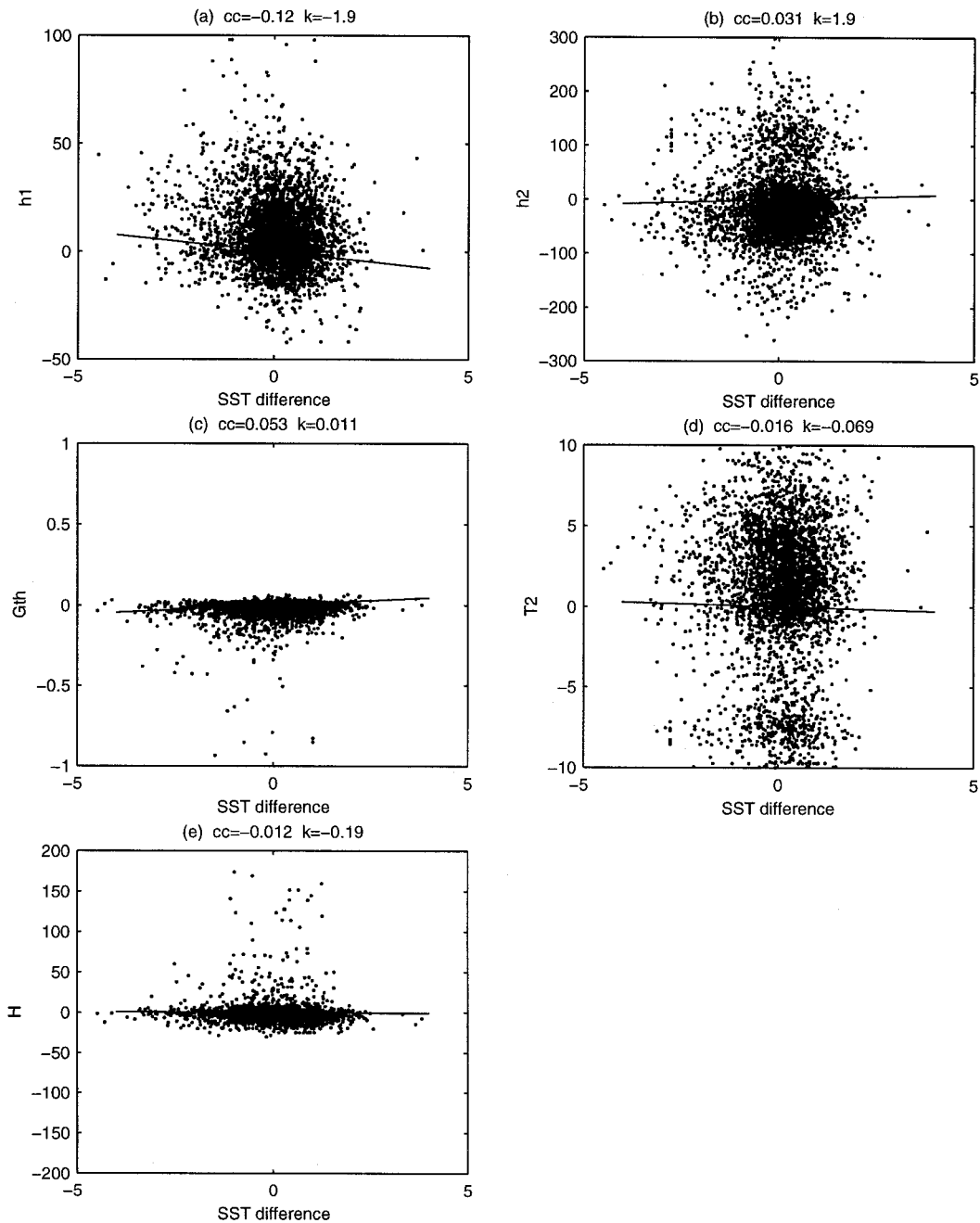
FIG. 7. Scatter diagrams of SST anomaly with various thermal parameter anomalies: (a) h'_1, (b) h'_2, (c) G'_{th}, (d) T'_{tb}, and (e) H'.

for the region. If correlation between T'_s and h'_1 is not significant, such as Tully and Giovando (1963) found in one region of the North Pacific, it is very hard to use this inverse method for that region. Furthermore, we should do the rms error and correlation tests after the inversion to see the real usefulness.

9. Conclusions

1) The thermal parametric model depicted in this paper demonstrates the capability to invert subsurface structure from SST for one oceanic region, the South China Sea (SCS). Based on a multitimescale hypothesis, the long correlation timescale parameters are treated as a background dataset during the inversion, which may be predetermined by historical data. The short timescale parameters are determined by the inverse method. In this study, only SST and MLD are treated as short correlation timescale parameters. After a long timescale parameter dataset is established, we have a one-to-one relation between
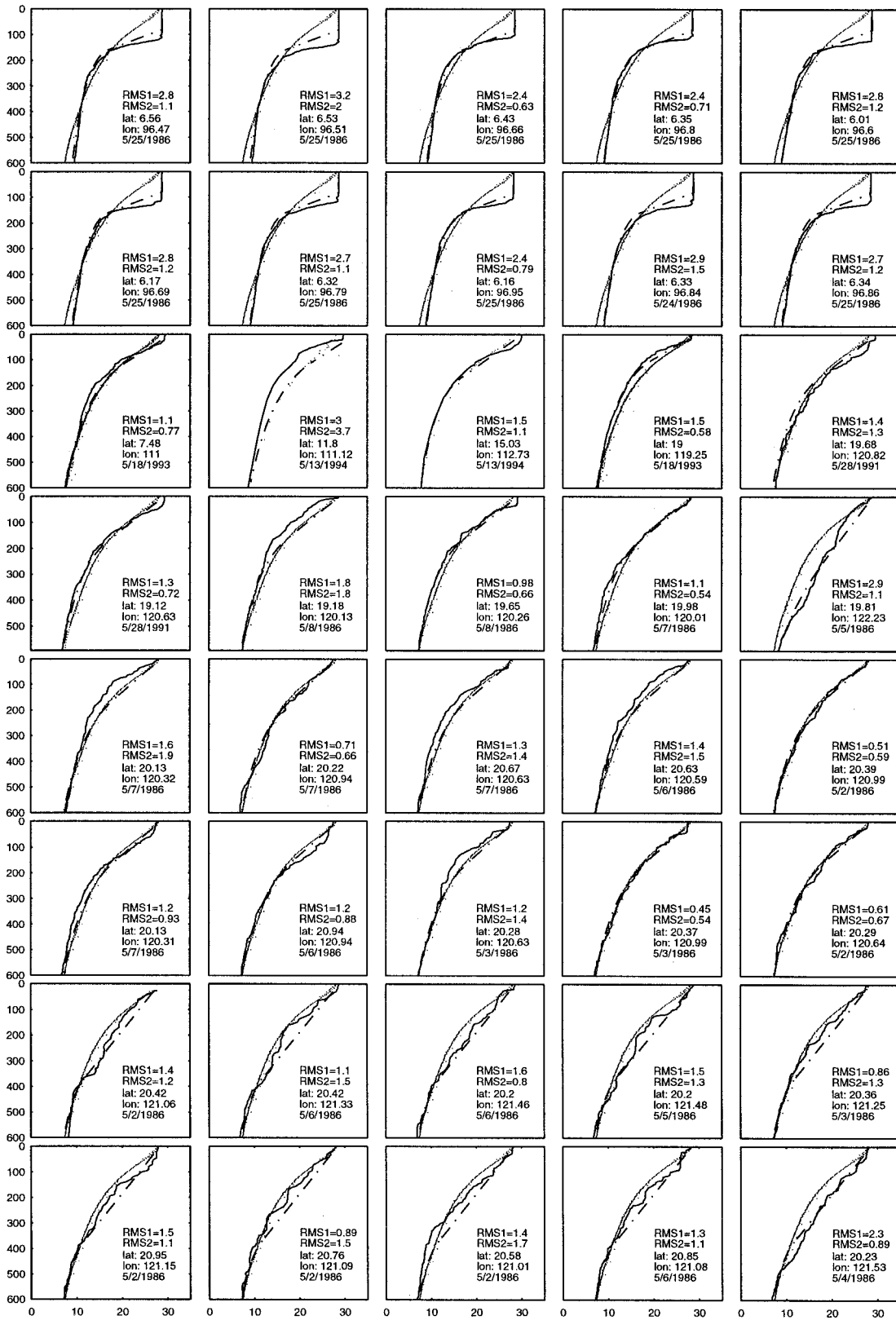
FIG. 8. Comparison between 40 regressed (dotted), inverted (dash–dotted), and observed (solid) profiles.
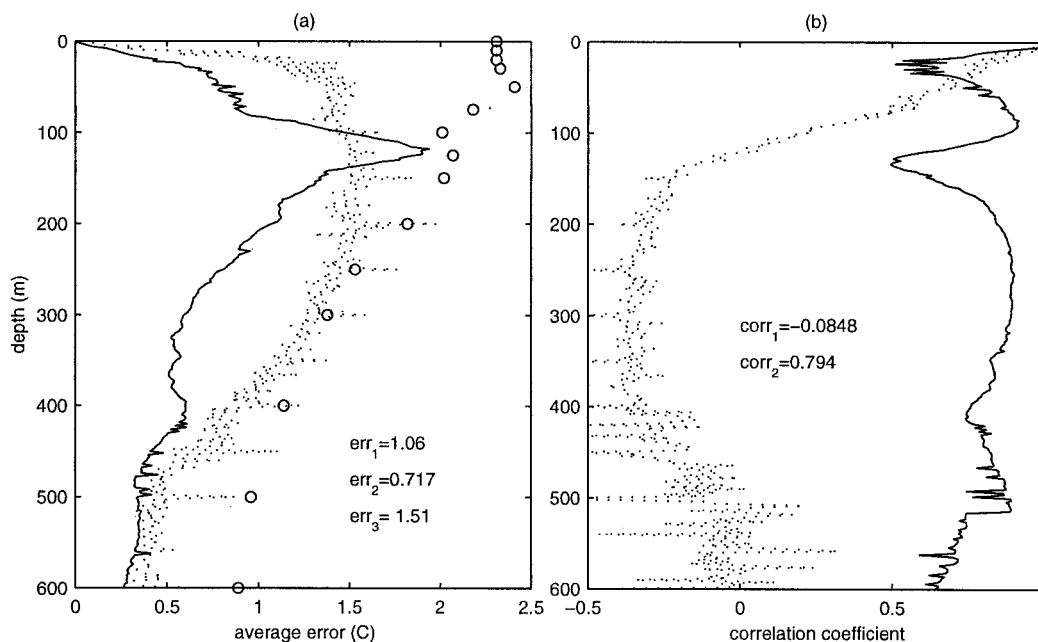
FIG. 9. Vertical distribution of (a) inverse model rms errors (solid), regression model rms errors (dotted), and climatological SDV (circle); and (b) correlation coefficients between observed and inverted (solid), observed and regressed (dotted) profiles.

SST and MLD. Through this relation, MLD is determined from SST. Together with the predetermined long timescale parameters, we can easily obtain the vertical profile for each known SST.

2) The inverse methods proposed here were verified by a dataset from the U.S. Navy's Master Observational Oceanographic Data Set (MOODS) for SCS in May 1932–94. Among the total 10 153 profiles, 10 113 profiles, treated as training data, were processed into a dataset consisting of SST, MLD, TBD, TTG, and DLS using the parametric model. SST of the remaining 40 profiles were used for the inversion based on the known information on TBD, TTG, and DLS. The 40 inverted profiles agreed quite well with the corresponding observed profiles. The rms error is around 0.72°C, and the correlation between the inverted and the observed profiles is 0.79. The improvement of the multitimescale versus the mean inverse methods is in the upper layer from around 1°C at the surface to 0.1°C at 30-m depth.

3) To apply the multitimescale inverse method globally, we should first test the validity of the multitimescale hypothesis. This can be done with the correlation analysis. If the multitimescale hypothesis is valid, we can use this method to invert the subsurface thermal structure from SST. Furthermore, we should do the model verification after the inversion to see the real usefulness.

REFERENCES

Chu, P. C., 1993: Generation of low-frequency unstable modes in a coupled equatorial troposphere and ocean mixed-layer model. *J. Atmos. Sci.,* **50,** 731–749.
——, C. R. Fralick Jr., S. D. Haeger, and M. J. Carron, 1997a: A parametric model for the Yellow Sea thermal variability. *J. Geophys. Res.,* **102,** 10 499–10 507.
——, H. C. Tseng, C. P. Chang, and J. M. Chen, 1997b: South China Sea warm pool detected from the Navy's Master Oceanographic Observational Data Set (MOODS). *J. Geophys. Res.,* **102,** 15 761–15 771.
——, C. W. Fan, J. Kirling, and C. Lozano, 1998: An airborne expandable bathythermograph survey of the South China Sea in May 1995. *J. Geophys. Res.,* **103,** 21 637–21 652.
——, Q. Q. Wang, and R. H. Bourke, 1999: A geometric model for the Beaufort/Chukchi Sea thermohaline structure. *J. Atmos. Oceanic Technol.,* **16,** 613–632.
Levitus, S., and T. Boyer, 1994: *Temperature.* Vol. 4, *World Ocean Atlas, NOAA Atlas NESDIS,* 117 pp.
Tao, S.-Y., and L.-X. Chen, 1987: A review of recent research on the east Asian summer monsoon in China. *Monsoon Meteorology,* C.-P. Chang and T. N. Krishnamurti, Eds., Oxford University Press, 60–92.
Teague, W. J., M. J. Carron, and P. J., Hogan, 1990: A comparison between the generalized digital environmental model and Levitus climatologies. *J. Geophys. Res.,* **95,** 7167–7183.
Tully, J. P., and L. F. Giovando, 1963: Seasonal temperature structure in the eastern subarctic Pacific Ocean. *Marine Distributions,* M. J. Dunbar, Ed., University of Toronto Press, 10–36.