ARTICLE IN PRESS

Quaternary Science Reviews ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Estimating glacial western Pacific sea-surface temperature: methodological overview and data compilation of surface sediment planktic foraminifer faunas

M.-T. Chen[a,*], C.-C. Huang[a], U. Pflaumann[b], C. Waelbroeck[c], M. Kucera[d]

[a]Institute of Applied Geosciences, National Taiwan Ocean University, Keelung 20224, Taiwan
[b]Institut für Geowissenschaften, Universität Kiel, Olshausenstr. 40-60, 24118 Kiel, Germany
[c]Laboratoire des Sciences du Climat et de l'Environnement, Laboratoire mixte CNRS-CEA, Domaine du CNRS, Gif-sur-Yvette, France
[d]Department of Geology, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK

Received 24 October 2003; accepted 15 July 2004

Abstract

We present a detailed comparison of five ''transfer function'' techniques calibrated to reconstruct sea-surface temperature (SST) from planktic foraminifer counts in western Pacific surface sediments. The techniques include the Imbrie–Kipp method (IKM), modern analog technique (MAT), modern analog technique with similarity index (SIMMAX), revised analog method (RAM), and the artificial neural network technique (ANN). The calibration is based on a new database of 694 census counts of planktic foraminifers in coretop samples from the western Pacific, compiled under a cooperative effort within the MARGO (multiproxy approach for the reconstruction of the glacial ocean surface) project. All five techniques were used to reconstruct SST variation in a well-dated Holocene to last glacial maximum interval in core MD972151 from the southern South China Sea (SCS) to evaluate the magnitude of cooling in the western tropical Pacific during the LGM. Our results suggest that MAT, SIMMAX, RAM and ANN show a similar level of performance in SST estimation and produce $\leqslant 1\,^{\circ}$C uncertainties in coretop SST calibrations of the western Pacific. When applying these techniques to the downcore faunal record, the IKM, which performed significantly worst in the calibration exercise, produced glacial SST estimates similar to present-day values, whereas the other four techniques all indicated $\sim 1\,^{\circ}$C cooler glacial SST. Because of their better performance in the calibration dataset, and because of the convergence among the techniques in the estimated magnitude of glacial cooling in the studied core, we conclude that MAT, SIMMAX, RAM and ANN provide more robust planktic foraminifer paleo-SST estimates than traditional IKM techniques in western Pacific paleoceanographic studies.

1. Introduction

Foraminifers growing in the oceans provide some of the richest paleoclimate archives in the world. For the purpose of reconstructing past sea surface temperatures (SSTs), planktic foraminifer faunas are particularly useful paleoclimate proxies because they are widely distributed and are abundant enough for quantitative and statistical analysis. The first extensive application of planktonic foraminifers in paleoclimatology was the pioneering study by CLIMAP (1976) in which the distribution data of planktic foraminifer faunal assemblages and their relationships with SSTs were used for predicting past surface ocean-climate patterns during the last glacial maximum (LGM). Despite an intensive effort to develop new climatic and geochemical proxies for reconstructions of glacial surface ocean conditions, planktic foraminifer faunas continue to play an important role in providing essential information on past SSTs and for inferring water mass and circulation changes.

*Corresponding author. Tel.: +886-2-2462-2192x6503; fax: +886-2-2462-5038.
E-mail address: mtchen@mail.ntou.edu.tw (M.-T. Chen).

0277-3791/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.quascirev.2004.07.013

Tropical oceans, especially those in the western Pacific, serve as a heat engine for Earth's climate and as a vapor source for its hydrological cycle. On annual to inter-annual time scales, the western Pacific has a great impact on global climate through propagation of El Niño—Southern Oscillation (ENSO) events. There is a need to understand more precisely the role that the western Pacific plays in glacial–interglacial cycles and in the more rapid millennial- to centennial-scale oscillations of global climate change (Cane and Clement, 1999; Clement and Cane, 1999; Clement et al., 1999). Until recently, there have been debates on the magnitude of SST cooling during the LGM in different parts of the western Pacific. CLIMAP (1976, 1981) estimated ~2–3 °C LGM cooling in most western equatorial Pacific and western Australian current regions and a relatively large cooling (~4–8 °C) was estimated for the main Kuroshio pathway region near the Pacific margin of Japan. In the South China Sea (SCS) and East China Sea (ECS), where many well-dated, high-resolution paleoceanographic records have been collected and analyzed recently, CLIMAP (1981) had obtained no data which might have been used as a control.

One of the most remarkable results of CLIMAP (1976, 1981) were the reconstructed positive SST anomalies (the LGM warmer than the present) occurring in the mid-latitude gyre regions. These positive anomalies are difficult to reconcile with some terrestrial proxy-based or climate modeling results (Webster and Streeten, 1978; Rind and Peteet, 1985), in which the glacial tropics SSTs were estimated to be as much as 5–6 °C colder than the present. Moreover, geochemically based estimates using tropical coral skeletons Sr/Ca (Beck et al., 1992; Guilderson et al., 1994) have suggested a large glacial cooling, which is consistent with the terrestrial evidence. These disagreements in different types of SST reconstructions still remain to be reconciled.

The Imbrie–Kipp transfer function method (IKM) (Imbrie and Kipp, 1971), the first rigorous environmental calibration technique developed for use with marine microfossils, formed the backbone of CLIMAP LGM SST reconstructions. For reconstructing LGM SSTs in the western Pacific, the CLIMAP project developed a regional IKM transfer function FP-12E (Thompson, 1976, 1981) based on census counts of planktic foraminifers in 165 coretop samples, whereas the SSTs of the western Australian current region near the eastern Indian Ocean were estimated by another IKM transfer function, FI-2 (Hutson and Prell, 1980). The validity of CLIMAP SST reconstructions was later assessed by another, methodologically different, method: the modern analog technique (MAT) (Hutson, 1980; Prell, 1985). In the southwestern tropical Pacific near eastern Australia, a small LGM cooling of ~0–2 °C was reported from a new coretop compilation study using MAT (Barrows et al., 2000). Although the MAT results reduced the difference between the observed and estimated SSTs, this method still suggested the existence of positive anomalies in the western Pacific. It is difficult to assess how reliable the warm glacial SST estimates in the western Pacific are, since the limited coverage and uneven distribution of coretop samples as well as the different levels of clacite preservation in these samples make SST reconstructions from planktic foraminifers in this area rather difficult. To complete a better global map of glacial SSTs and to provide modeling groups with a better basis for simulating glacial and present climate conditions, the western Pacific was identified as a region where there was a great need for improvement in the collection of more high-quality coretop samples and for development of better regional estimation techniques. While a testing of SST estimation techniques against an entire Pacific planktic foraminifer coretop database is presented in Kucera et al. (2004), a separate calibration for the western part of the Pacific is needed, as this testing excludes faunas from equatorial divergence/upwelling regions which may bias the estimation.

Many new SST estimation techniques have been proposed recently to improve traditional techniques and reduce their biases. These new methods include: SIMMAX (modern analog technique using a similarity index) (Pflaumann et al., 1996); revised analog method (RAM) (Waelbroeck et al., 1998); and Artificial Neural Networks (ANN) (Malmgren and Nordlund, 1997; Malmgren et al., 2001). Each of these methods adopts a different approach for improved performance in SST reconstructions: SIMMAX uses a new similarity index and incorporates geographic information, RAM increases the number and range of calibration samples and introduces more rigorous criteria for selection of best analog samples, and ANN establishes highly nonlinear equations which optimize fauna-SST relationships using artificial intelligence techniques. All of these new techniques have been or are being applied to an Atlantic data set and have been demonstrated to reduce the biases of previous SST estimates.

Working within the framework of multiproxy approach for the reconstruction of the glacial ocean surface (MARGO) objectives, the current study was designed to compile a well-organized, high-quality surface sediment planktic foraminifer faunal database for the western Pacific. We used this database to develop better estimates of the SST, based on available, traditional or newly developed techniques. Our aim was to compare and evaluate the performance of these techniques, by analyzing the relationships between the biases and possible factors in the calibration data set. Here we present SST reconstructions using five techniques in a high-quality and well-dated International Marine Past Global Change Study (IMAGES) core from the SCS and provide the SST estimates and the uncertainty ranges of the estimates for the Holocene (0–4 ka) and LGM (19–23 ka) intervals (EPILOG chronozone) (Mix et al., 2001).

2. Data and methods

2.1. Calibration data sets

To obtain a data set of high-quality, well-organized coretop planktic foraminifer fauna data in surface sediments from the western Pacific, we cooperated as one of the MARGO working groups and compiled a western Pacific modern planktic foraminifera data base composed of 694 samples (data available on the internet [http://www.pangaea.de/Projects/MARGO/data.html]). The set of 694 coretop data used in this study has benefited from the contributions of various publication sources: global compilation efforts from Prell et al. (1999) and Ortiz and Mix (1997); East China Sea and Okinawa Trough studies by Ujiie and Ujiie (2000); SCS data collections from Miao et al. (1994), Chen et al. (1998), and Pflaumann and Jian (1999); Southwestern Pacific investigations by Thiede et al. (1997); several individual surveys of the North Pacific (Coulbourn et al., 1980; Ye and Yang, unpublished data; Kiefer et al., 2001); and unpublished data from M. Schulz. The geographical range of the 694 coretop data is from a latitude of 65°N to 65°S and a longitude of 105°E to 180°E (Fig. 1); the samples provide good coverage of the SST range between ~30 and 20 °C, but are relatively sparse in the SST range of 20–5 °C (Fig. 1). The data points are of especially high resolution in the western equatorial Pacific and in the marginal basins of the western Pacific such as the SCS, the East China Sea, and Western Australia. We think that this calibration data set is most appropriate for developing SST estimation techniques for the low- to mid-latitude western Pacific and should be suitable for reconstructing high-resolution SST records from high sedimentation rate sites.

The relative abundance of each planktic foraminifer species is expressed as a percentage of the total faunal count. Descriptive statistics of the relative abundances of 28 species of planktic foraminifers were calculated for the data set (Table 1). The taxonomy of planktic foraminifers used in this study was determined by the MARGO group and generally followed the working schemes of Parker (1962), Be (1967) and Kipp (1976), with some modifications. We combined the two different morphotypes *Globigerinoides sacculifer* (no sac) and *G. sacculifer* (with sac) together as one unit. We also combined *Globorotalia menardii* and *G. tumida* into one category to overcome an identification difficulty that vexed previous studies. In this western Pacific data set, the most important taxonomic problem is the definition and identification of morphologically similar groups consisting of *Neogloboquadrina dutertrei*, *N. pachyderma-dutertrei* intergrade (P-D intergrade)'' (Kipp, 1976), and *N. pachyderma* (right coiling). Most western Pacific coretop studies did not recognize the ''P-D intergrade'' and different studies use different criteria for recognizing *N. dutertrei* and *N. pachyderma* (right coiling); thus, we wanted to minimize this ''taxonomic noise'' during the processing of the coretop data for our SST estimation. We have assumed that most ''P-D intergrade'' specimens identified in previous western Pacific studies represent juvenile forms of *N. dutertrei* and could be placed in the category of *N. dutertrei.* In this study we therefore included specimens of ''P-D intergrade'' with *N. dutertrei* in coretop as well as downcore data. According to the sample processing procedures described in the published studies, all samples used to generate faunal counting data were weighed and sieved through a 150 μm sieve. Census counts of planktonic foraminifers were made on fractions ⩾150 μm for each sample and at least 300 whole specimens were counted for most samples. These coretop samples range in water depth between 42 and 5351 m, a wide interval that overlaps the depth of the regional carbonate lysocline and represents various calcite preservation levels.

Modern SST values (summer, winter and annual) were assigned to each of the 694 coretop samples using the data in the World Ocean Atlas version 2 (WOA, 1998), in the same way as in Kucera et al. (2004), to ensure a common reference base for all SST proxies used by the MARGO project. Temperatures at the sample locations were computed using the WOA 98 sample software (http://www.palmod.uni-bremen.de/~csn/woa-sample.html).

2.2. SST estimation techniques

Five SST estimation techniques (IKM, MAT, SIMMAX, RAM and ANN) were applied to the new calibration data set in the present study, to determine the accuracy of SST estimates produced by different techniques through a validation in a coretop data set with known SST values. The accuracy of SST estimates based on a validation of a coretop data set is the best available measure of the likely performance of a technique in fossil assemblages. For the purpose of the downcore application we have chosen a SCS IMAGES core MD972151 from which many good AMS C^{14} dates and age control data, and planktic foraminifer faunas and alkenone SST are available (Huang et al., 1999; Lee, in preparation; Lee et al., 1999; Huang et al., 2002). Since our calibration data set has good coverage in the SCS, the LGM reconstructions of this core should be highly accurate.

In developing the IKM transfer function we used the standard statistical technique of VARIMAX Q-mode factor analysis (Imbrie and Kipp, 1971) to decompose

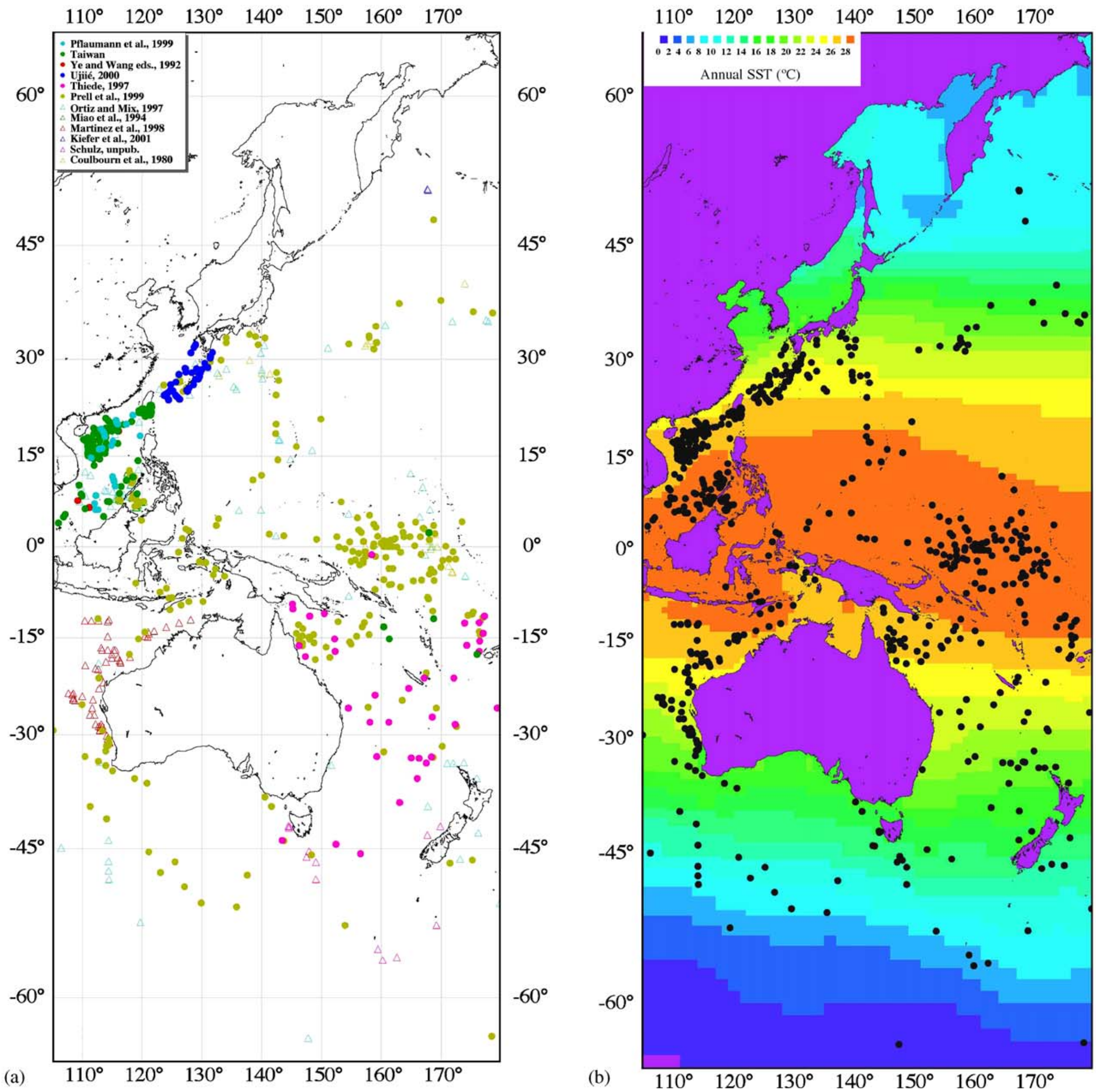

Fig. 1. A western Pacific map indicating the site locations of a newly compiled coretop data set of planktic foraminifer faunal abundances ($N = 694$) and sea-surface temperature (SST, annual average, WOA, 1998) distribution in the western Pacific.

the 694 faunal variation data into a reduced number of factor variables that were independent from each other. Our analyses of the faunal data were based on a log transformation (ln[species percentage + 1]) that was first applied in an Atlantic and eastern equatorial Pacific study (Mix et al., 1999). In the factor analysis, the log transform amplifies the signals of less dominant species with a compression of dominant species abundance variations. The log transformation also has the effect of making the species distribution more Gaussian, which is an essential assumption of the next step in the IKM that involves multiple regression. The log transform is applied to all faunal percentage data (coretop and downcore samples) in this study before calculating factor scores or loadings. Four factors (Table 1) that explain ~89% of the original but log-transformed coretop fauna data are derived from the factor analysis. Planktic foraminifer species *G. ruber*, *G. sacculifer*, *Globigerina bulloides*, *N. pachyderma* (right coiling), *Pulleniatina obliquiloculata*, *G. inflata*, *G. menardii*+ *tumida* and *Globigerinita glutinata* exhibit high factor scores that are highly correlated with these factors. We calibrated the coretop factors loadings of these four factors with their cross-product and squared terms, to annual average, winter and summer SSTs by following a standard procedure of multiple regression techniques

Table 1
Factor score assemblage matrix of 694 coretop fauna percentage data from the western Pacific area

Foraminifer species	Factor 1	Factor 2	Factor 3	Factor 4
O.universa	0.063	0.078	0.003	0.046
G.conglobatus	0.144	−0.013	0.037	0.163
G.ruber	**0.597**	0.030	−0.141	0.291
G.tenellus	0.184	−0.002	−0.107	0.015
G.sacculifer	**0.409**	−0.086	0.076	0.140
S.dehiscenes	−0.053	0.001	0.195	0.056
G.aequilateralis	0.222	−0.040	0.149	0.011
G.calida	0.182	0.006	−0.033	0.035
G.bulloides	0.196	**0.409**	0.027	**−0.540**
G.falconensis	0.088	0.195	−0.094	0.056
B.digitata	0.009	0.021	0.061	0.037
G.rubescens	0.152	−0.024	−0.073	−0.066
T.quinqueloba	−0.001	0.141	−0.056	−0.138
***N.pachyderma* (L)**	−0.069	0.250	−0.010	**−0.443**
N.pachyderma (R)	−0.094	0.377	0.039	0.149
N.dutertrei	0.136	0.196	0.388	0.085
G.conglormerata	0.022	−0.013	0.035	−0.024
G.hexagona	0.026	−0.002	−0.020	0.001
P.obliquiloculata	0.057	−0.066	**0.661**	−0.187
G.inflata	−0.148	**0.661**	0.008	0.392
G.truncatulinoides (L)	−0.025	0.163	−0.029	−0.004
G.truncatulinoides (L)	−0.011	0.144	0.008	0.182
G.crassaformis	0.016	0.065	−0.019	0.044
G.hirsuta	0.005	0.054	−0.014	0.061
G.scitula	0.036	0.046	−0.011	−0.006
G.menardii+tumida	0.016	−0.011	**0.530**	0.151
C.nitida	0.015	−0.004	−0.009	−0.001
G.glutinata	**0.448**	0.132	−0.068	−0.259
Variance (%)	51.52	16.38	18.43	2.40
Cumulative variance	51.52	67.90	86.33	88.74

(VARIMAX solution by CABFAC factor analysis).

Table 2
Multivariate regression coefficients

	Variable	Annual SST	Winter SST	Summer SST
1	F1	36.559	48.199	26.465
2	F2	15.575	22.285	8.441
3	F3	29.943	37.463	22.657
4	F4	12.159	*	18.028
5	F1*F1	−17.477	−23.613	−12.658
6	F2*F2	−15.159	−17.752	−11.399
7	F3*F3	−13.108	−15.275	−11.065
8	F4*F4	−8.674	−19.590	−5.422
9	F1*F2	−23.893	−34.321	−14.089
10	F1*F3	−26.581	−34.093	−21.016
11	F1*F4	−17.218	−9.198	−23.261
12	F2*F3	−10.573	−21.528	*
13	F2*F4	3.345	14.699	*
14	F3*F4	−7.960	*	−10.532
	Intercept	10.796	5.040	16.506

The symbol "*" marks the variables that are not incorporated in the equations.
Note: Annual SST: Annual sea surface temperature. Winter SST: Average sea surface temperature of caloric winter season. Summer SST: Average sea surface temperature of caloric summer season.

(Imbrie and Kipp, 1971), but we constrained the regression by using a "Best Subsets Regression" (MINITAB program) criterion for selecting terms to enter into the equation (Table 2).

We calculated MAT SST for the 694 coretop and MD972151 downcore samples using the same 694 coretops as the calibration data set. The MAT procedure used in this study is the same as that reported in Prell (1985), which introduces a dissimilarity coefficient (the squared chord distance) to measure the dissimilarity between coretop–coretop or coretop–downcore paired species percentage data. Higher values of the squared chord distance indicate higher dissimilarity. The squared chord distance also has the effect of amplifying the signals of less dominant species and reducing the influence of dominant species abundances. MAT SST estimates of coretop or downcore samples are thus calculated from the weighted average of observed SST values of 10 best analog samples. We applied a cut-off dissimilarity value of 0.4, as suggested by Prell (1985), for coretop samples that did not qualify to be included in an SST estimation. Therefore, in some cases, the number of best analogs used to calculate the SST estimate was less than 10.

The SIMMAX approach (SIMMAX28-1900) adopted here is a revised version which has been applied in a small set of SCS coretop samples (Pflaumann and Jian, 1999); this routine calculates the scalar product of the normalized faunal assemblage vectors as a similarity index. This similarity index ranges from 0 to 1, with 0 indicating complete dissimilarity and 1 indicating full similarity. Both coretop and downcore data sets were analyzed by the SIMMAX routine without any threshold in the search for similarity, and with scalar-product-weighted averaging of the SSTs of the 10 best analogs from the western Pacific 694 coretop sample data set. In this study, we also used the geographic distance-weighting procedure for SST estimation (Pflaumann et al., 1996). This is important as coretops of the western Pacific are unevenly distributed. The resultant output for the single core and time slices contains similarity, estimated annual average, and winter and summer SSTs, as well as the minima, maxima and standard deviations of the 10 best analogs for the three SST estimates, respectively.

RAM SST calibrations were performed for the first time on the western Pacific coretop data. The RAM technique (Waelbroeck et al., 1998) was proposed to improve the estimation accuracy and precision of the MAT, adopting the same dissimilarity coefficient (i.e. squared chord distance) but with two important modifications: (1) selecting only good analog coretop samples by examining the rate of increase in dissimilarity; and (2) remapping and interpolating the coretop fauna database into a more homogenous, evenly distributed space as a function of winter and summer SSTs. These modifications should make RAM useful for SST reconstructions in the western Pacific, where less evenly distributed and/or less well-preserved samples may limit the selection of good analogs for downcore estimates. In this study, a threshold value of the dissimilarity coefficient of 0.6 was applied to reject poor analog coretop samples. The 694 western Pacific coretop samples were expanded by 466 "virtual coretops" by the RAM remapping procedure (with a grid step of 0.4 °C and an interpolation radius of 0.5 °C) in a winter and summer SST, and annual environmental space, respectively. In this paper, we use the latest RAM02 software with the following parameters for all regional calibrations: initial $\alpha = 0.1$, $\beta = 10$, $\gamma = 0.25$ °C, $R = 0.3$ °C, as explained in Kucera et al. (this volume).

A fundamentally different approach for SST calibration, ANN uses sophisticated algorithms to search for a relationship between coretop fauna abundance distributions and SSTs. The general principles and architecture of a back propagation (BP) neural network and its application in reconstructions of past environmental conditions from assemblage counts have been described by Malmgren and Nordlund (1997) and Malmgren et al. (2001). A trained ANN can be best compared to a long, complicated, recurrent mathematical formula transforming species abundances into a desired variable(s). An ANN can successfully learn very complex, nonlinear relationships; this technique is not as dependent on the size, coverage and balance of the calibration dataset as are modern analog techniques (MAT, SIMMAX, RAM). Unfortunately, it is virtually impossible to interpret the meaning of the coefficients associated with individual neurons and thus to understand how the network assigns output variables values to unknown samples.

The networks used in this paper were trained on a database with counts of 28 species in 1111 Pacific coretops. The details of the training results are presented in Kucera et al. (2004). It is important to point out that the ANN SST estimation presented in this study was based on whole Pacific database training, which may reduce the comparability to the other techniques.

3. Comparisons of SST estimations

3.1. Modern SST calibrations

We calibrated (or trained) abundances of 28 species of planktic foraminifers in 694 coretops to annual average, winter, and summer SSTs following the procedures of the five individual SST estimation techniques (IKM, MAT, SIMMAX, RAM, ANN) described above. The calibration results are examined here by comparing scatter plots of observed vs. estimated SSTs (Fig. 2). The success of the calibration can also be evaluated by calculating the correlation (R^2) between the observed and estimated SSTs as well as the root mean squared error (RMS) of the residuals of the estimated SSTs (Table 3). To simplify the presentation of our results, only the annual average SSTs estimated by these five different techniques are presented for comparison.

Based on their success in predicting observed SSTs from coretop samples, the different techniques can be divided into two distinct groups. For all SST estimations, the IKM technique produced by far the lowest correlation and the highest RMS value (Fig. 2; Table 3). The IKM showed relatively large scattering in observed vs. estimated SST plots and gave a magnitude of ~1.3–1.4 °C of the uncertainties of the estimates.

MAT, SIMMAX, RAM and ANN appeared to yield similarly accurate results. The relatively high correlations and small RMS values (Fig. 2; Table 3) produced by the four techniques indicated good predictability in estimating SSTs in the 694 coretop data. Applying these four techniques to the coretops gives somewhat lower estimation uncertainties of ~0.8–0.9 °C, which represents a significant improvement over the IKM. Among the four techniques, our results suggest also that RAM produces a much lower RMS value (~0.85 °C) than MAT (~0.97 °C), SIMMAX (~0.95 °C), and ANN (~0.96 °C). Kucera et al. (this volume) have pointed out that the low error values produced by RAM reflect the fact that the leaving-one-out validation procedure in the RAM02 software is implemented only after the two-dimensional interpolation and thus underestimates the full error rate. The difference is, however, negligible compared with the ~0.5 °C estimation uncertainty difference between the IKM and MAT/SIMMAX/RAM/ANN groups.

We also noticed that in estimating winter and summer SSTs, these five different techniques yielded results similar to those for estimating annual average SSTs (Table 3). The IKM showed large errors in estimating the SSTs, while the other four (MAT, SIMMAX, RAM, and ANN) exhibited smaller errors with better predictability than the IKM. All five techniques produced relatively larger uncertainties when estimating winter SSTs (1.2–1.7 °C) versus summer SSTs (0.7–1.3 °C).

3.2. Coretop evaluation of estimation biases

To evaluate potential biases in the calibration of the five SST estimation techniques, which might lead to inaccurate downcore estimates, we examined the relationship of SST residuals (the difference between estimated minus observed SSTs) with observed SST (Fig. 3), latitudinal distribution (Fig. 4), and water depth of coretops (Fig. 5). Techniques (MAT, SIMMAX, RAM) adopting dissimilarity (or similarity) indices for comparing coretop faunal samples were also

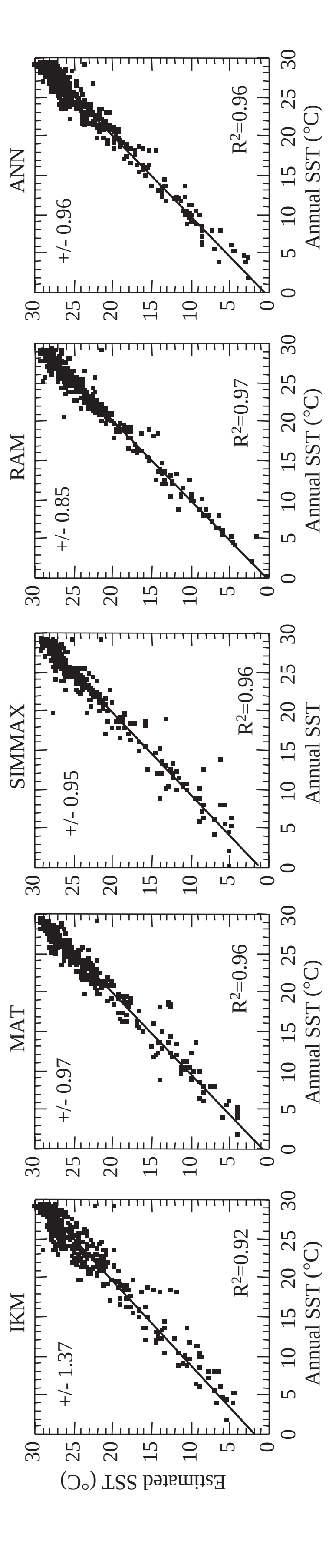

Fig. 2. Scatter diagrams of observed vs. estimated SST (annual average) produced by IKM, MAT, SIMMAX, RAM and ANN for 694 western Pacific coretops. The ANN was trained on the whole Pacific database ($N = 1111$, Kucera et al., 2004).

Table 3
Statistical results for IKM, MAT, SIMMAX, RAM, and ANN

Methods	Annual SST		Winter SST		Summer SST	
	R^2	RMS error	R^2	RMS error	R^2	RMS error
IKM	0.92	1.37	0.89	1.73	0.91	1.35
MAT	0.96	0.97	0.96	1.14	0.96	1.00
SIMMAX	0.95	0.96	0.96	1.12	0.96	0.90
RAM	0.97	0.85	0.95	1.17	0.98	0.72
ANN	0.96	0.96	0.94	1.27	0.98	0.74

Note: Annual SST: Annual sea surface temperature. Winter SST: Average sea surface temperature of caloric winter season. Summer SST: Average sea surface temperature of caloric summer season.

evaluated by examining the relationship of the SST residuals with different levels of dissimilarity (or similarity), since the lack of good analog samples in the coretop data set might also bias the SST estimates (Fig. 6).

No significant correlation ($R^2 \leqslant 0.01$) was found between SST residuals and observed SST, water depth, or latitudes given by the five different techniques. This suggests that the variables did not bias the calibration or training estimation results for SST. Although the correlation was not significant, IKM, MAT, and SIMMAX SST estimates appeared to be too high in SST < 15 °C and too low in SST > 25 °C (Fig. 3). These patterns suggest that these three techniques may tend to underestimate the full range of SST variability in downcore faunal records. In evaluating the bias patterns associated with observed SSTs, the RAM and ANN estimates exhibited a noticeable improvement at the cold and warm ends of the SST range. These are indications that RAM and ANN might provide better reconstructions in the full range of possible SST variation.

When we examined the latitudinal distribution of the SST residuals by the five different techniques (Fig. 4), it also appeared that IKM, MAT, and SIMMAX SST estimates have large uncertainties, mostly at latitudes poleward of 30° north and south. The large errors (primarily overestimates) expressed in high latitudes indicate that IKM, MAT, and SIMMAX techniques are more susceptible to lacking good analogs in high latitudes. The 694 coretop database used here might not represent faunal distribution patterns well at higher latitudes, due to relatively few published data and also poorer preservation of carbonate sediments at high latitudes of the northwest Pacific (Fig. 1). These three techniques appear to have been forced to use fauna factors or analogs of lower latitudes for estimating the high-latitude coretop SSTs. Several large SST errors in the MAT and SIMMAX estimates associated with large dissimilarity (Fig. 6) support the idea that the MAT and SIMMAX estimates were biased by the no-analog

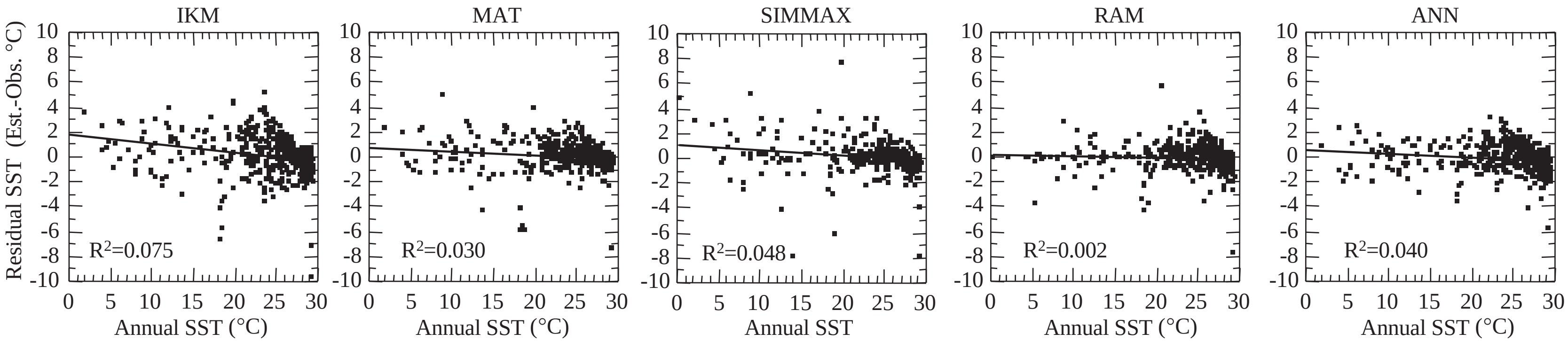

Fig. 3. Scatter diagrams of observed SST vs. ΔSST (estimated minus observed values) produced by IKM, MAT, SIMMAX, RAM and ANN for 694 western Pacific coretops.

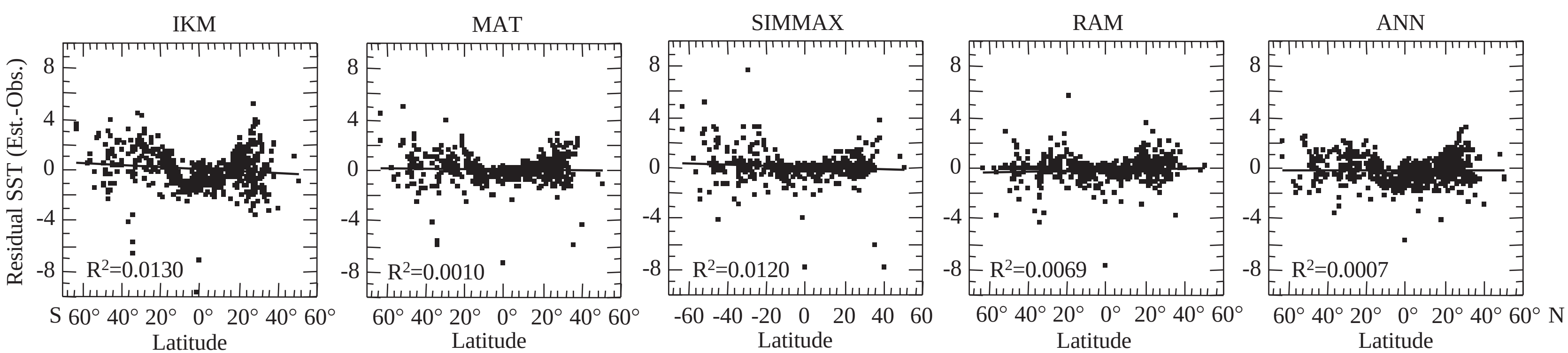

Fig. 4. Scatter diagrams of latitude vs. ΔSST (estimated minus observed values) produced by IKM, MAT, SIMMAX, RAM and ANN for 694 western Pacific coretops.

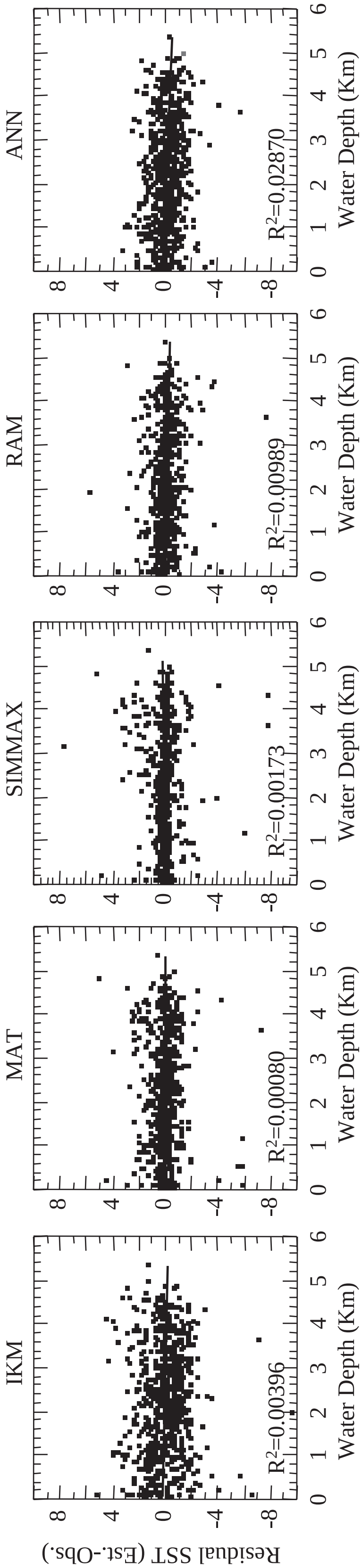

Fig. 5. Scatter diagrams of water depth vs. ΔSST (estimated minus observed) predicted by IKM, MAT, SIMMAX, RAM and ANN for 694 western Pacific coretop data.

conditions. In contrast, RAM and ANN appear to be more successful in estimating the high-latitude coretop SSTs, since there were only relatively small errors at the two ends (Fig. 3). For ANN, this could be in part due to the larger calibration database on which the neural networks were trained (Kucera et al., 2004).

Changes in different levels of carbonate preservation were previously thought to produce biases in SST estimates based on planktic foraminifer faunal assemblages (Thompson, 1976; Le, 1992; Miao and Thunell, 1994). Poor carbonate preservation might preferentially eliminate some species living primarily in warm surface water (such as *G. ruber* and *G. sacculifer*) with relatively thin and delicate tests. With the dissolution of these warm water species, SST estimate techniques might be biased toward colder estimates. The five different SST estimation techniques appear to be less affected by the dissolution problem; we observed no significant correlation between SST residuals and water depth, and no large error was observed from coretops located at deeper water depths (Fig. 5).

3.3. Downcore evaluation of estimation biases

We evaluated SST estimations by these five different techniques on a high-resolution planktic foraminifer fauna record of core MD972151 (8°43.73′N 109°52.17′E, water depth 1589 m) taken from the southern SCS during the 1997 IMAGES cruise (Chen et al., 1998). This record combines planktic foraminifer isotope stratigraphy, AMS C^{14} age model (Lee et al., 1999) and paleomagnetic stratigraphy (Lee, in preparation) with high-resolution planktic foraminifer fauna abundance data (Huang et al., 2002) and alkenone SST data (Huang et al., 1999). SST estimates based on the five above-described transfer function techniques are shown for a Holocene (0–4 kya) as well as an LGM (19–23 kya) window of the record, with comparisons to alkenone SST estimates and $\delta^{18}O$ variations (Fig. 7).

The observed annual average SST at this core site is 28.0 °C. All five techniques appear to have succeeded in reconstructing the absolute value of the observed SST based on the Holocene part of the record, although the stratigraphy of this core suggests that the age of the coretop sediments reaches ~1 kya. While applying SIMMAX and RAM to the downcore record, we excluded the coretop at the site MD972151. All five SST estimates fluctuated within a range of 27–29 °C and indicated a few small cooling events in the late Holocene window. While the MAT, SIMMAX, RAM and ANN SST estimates agreed well with each other and variations fell into a narrow range of ~27–28 °C, the IKM yielded warmer estimates by ~1–1.5 °C than all the other techniques and the alkenone SST. The IKM SST estimates may be biased by a low communality (~0.7) of

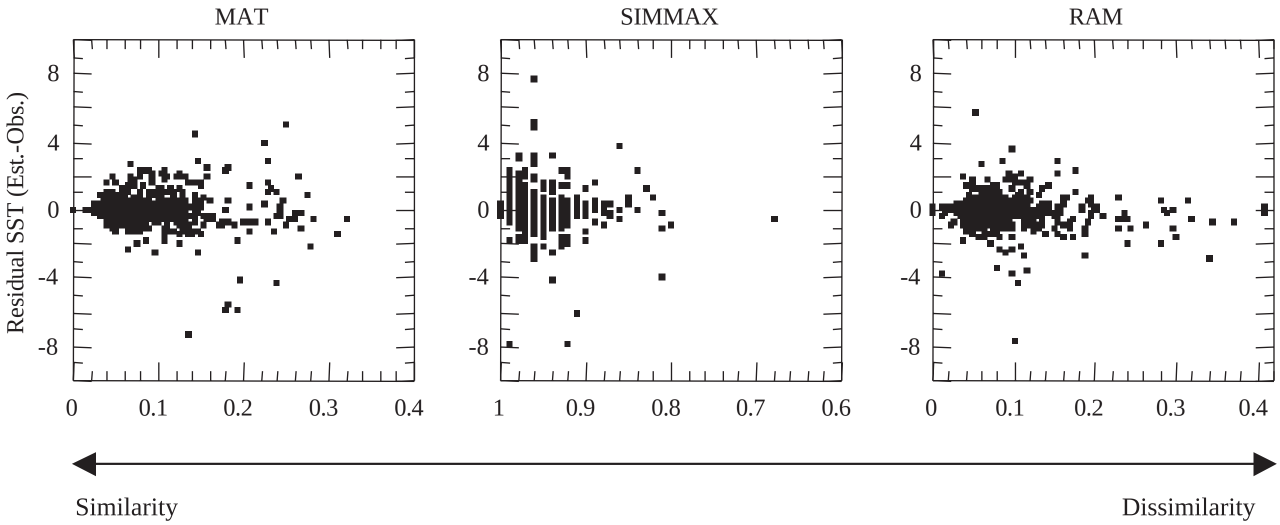

Fig. 6. Scatter diagrams of dissimilarity or similarity coefficients vs. ΔSST (estimated minus observed values) produced by MAT, SIMMAX and RAM for 694 western Pacific coretops.

the downcore samples (Fig. 7). The low communality indicates that the faunal assemblages in core MD972151 are not well explained by the factor model (Table 1) used in generating the transfer function. The appearance of local or high diversity of faunal assemblages in the southern SCS might not be represented well in the factor model of all western Pacific data. In contrast, all the other techniques appeared to overcome the problem, since they exhibited either low dissimilarity (or high similarity) or had a small standard deviation of estimates based on different partitions of the calibration database.

In the LGM window, all estimation techniques yielded much greater variability than in the Holocene. The IKM SST estimates remained within a range of 28–29 °C, but differed from the other techniques by +1.5–2 °C (Fig. 7). This appeared as ‘‘tropical stability’’ in glacial oceans if we accept the IKM SST estimate for the LGM in the southern SCS. In disagreement with the IKM, all other techniques produced much cooler SST estimates in the LGM. MAT, SIMMAX, RAM and ANN, as they did in the Holocene window, showed resemblance and consistently yielded LGM SSTs in the range of ~26.5–27 °C, although there are some indications of the presence of minor no-analog faunas in the LGM (relatively high dissimilarity and standard deviation) (Fig. 7). If these estimates are considered more reliable, then the faunal data would suggest ~1 °C cooling in the glacial southern SCS. In the LGM window, the alkenone SST shows a large offset of ~3 °C to the IKM and ~1–1.5 °C to the other techniques. As compared to the alkenone estimates of the Holocene, the organic geochemical method yields a ~2 °C glacial cooling in the southern SCS.

4. Discussion and conclusions

Our tests of five different SST estimation techniques against a western Pacific coretop data set indicated that better predictions could be obtained with MAT, SIMMAX, RAM and ANN than with IKM. The superiority of the other techniques over IKM when predicting coretop SSTs has been reported in a number of previous studies (Pflaumann et al., 1996; Ortiz and Mix, 1997; Waelbroeck et al., 1998; Malmgren et al., 2001). One source of the large errors in the IKM estimates might be the application of a factor model to faunal percentage data. A factor model is efficient in summarizing multivariate fauna data and translating that data into a few independent, oceanographically interpretable factor ‘‘assemblages’’. However, the composition of the factor assemblages might not reflect real ecological associations of the faunas in oceans. This discrepancy may have led to systematic biases in the IKM estimates. A newly revised version of IKM (Mix et al., 1999) where downcore faunas are included in the construction of the factor model appears to improve the predictability of IKM estimates, but this technique was not tested in this study.

MAT, SIMMAX and RAM obtained SST estimates from the observed values of SSTs from a set of coretop analogs, which were selected based on calculations of similarity/dissimilarity coefficients. The different coefficients adopted by these techniques may have resulted in slightly different estimation results. Since all similarity/dissimilarity coefficients adopted by the techniques have the common effect of amplifying the signals of less dominant species and of reducing the influence of dominant species abundances, we suspect that the use

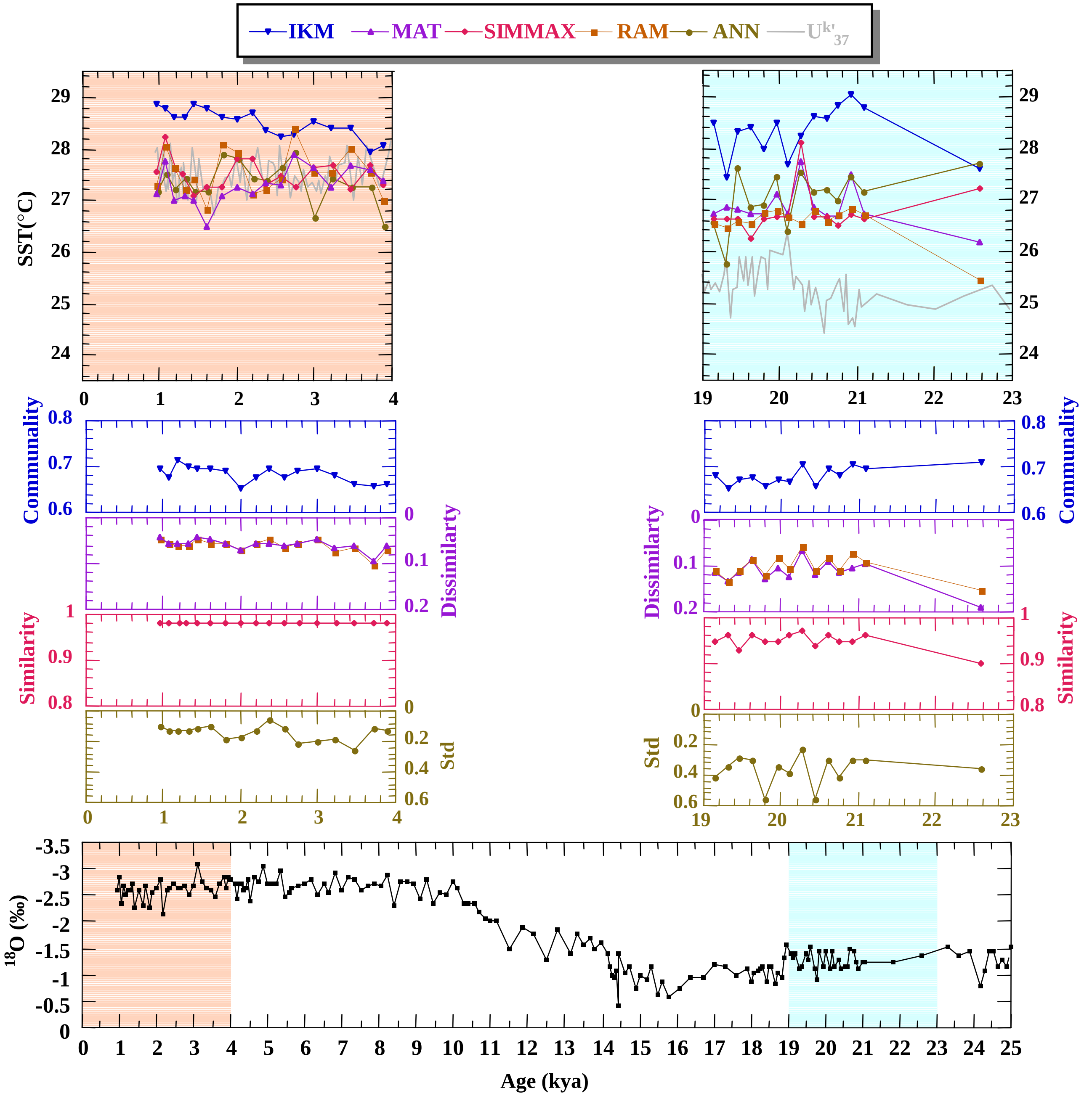

Fig. 7. Holocene and LGM SST reconstructions in core MD972151 (southern South China Sea) by IKM, MAT, SIMMAX, RAM and ANN and by planktic foraminifer isotope stratigraphy (Lee et al., 1999). The SST reconstructions presented here are compared with alkenone SST estimates from the same core (Huang et al., 1999).

of different coefficients is not a major factor affecting the accuracy of these estimations. While MAT only takes a simple average of the best 10 analogs, SIMMAX and RAM are considered to yield better estimations because they adopt more advanced weighting or truncating functions. In the SIMMAX estimations, a coretop sample from the downcore record was excluded from the analog searching because it would introduce a strong bias in the geographical distance weighting procedure (Pflaumann et al., 1996). The very good estimation ability indicated by the RAM results is only partly due to the analog searching that was based on an interpolated, expanded database (coretop itself exclusive), which may have significantly improved the sparseness of the coretop data in the western Pacific. On the other hand, it may mainly result from the leaving-one-out technique which RAM applied after virtual samples were added, and which explains the generally lower prediction errors of this technique (Kucera et al., this volume).

The ANN technique offered the same high accuracy and precision in coretop SST estimations as RAM. In our study, ANN demonstrated an ability to learn the very complicated relationships between faunas and SSTs and to produce very accurate estimates. ANN, although its training procedures are extremely complex and time-consuming, provided the same potential as RAM for estimating highly accurate SSTs in the western Pacific. Future work will be needed to compare the estimation results of ANN with training on western Pacific coretop data only.

SST reconstructions based on alkenones (Pelejero et al., 1999) and Mg/Ca data (Lea et al., 2000; Stott et al., 2002; Visser et al., 2003; Rosenthal et al., 2003) suggest a general cooling of ~2–3 °C in the western tropical Pacific marginal seas or open ocean during the LGM. In contrast to the general agreement of LGM alkenone and Mg/Ca SST data, atmosphere–ocean coupled GCMs produce a relatively wide range of estimations of 1–6 °C for LGM cooling on the surface ocean of the western tropical Pacific (Weaver et al., 1999; Bush and Philander, 1999; Hewitt et al., 2001; Shin et al., 2003). Although different atmospheric and/or oceanic processes were incorporated in these GCMs, all modeled estimates of surface ocean temperatures in the western tropical Pacific were interpretable by specific climatic mechanisms. If we consider the faunal SST estimates to be more reliable when different techniques yield similar results, the downcore data of MD972151 from the southern SCS suggest a ~1 °C cooling of the western tropical Pacific during the LGM. This ~1 °C cooling is supported by MAT, SIMMAX, RAM and ANN (Fig. 7), and thus it is robust and also independent of the methodology of SST estimation techniques that were used.

The contribution of the SCS core to determining the magnitude of glacial cooling in the western Pacific assumes that the SCS is a good representative of tropical Pacific climate. However, recent studies (Kienast et al., 2001; Kiefer and Kienast, 2004) suggest that SST in the SCS appears strongly influenced by terrestrial and/or local climate conditions and therefore deviates from open ocean conditions in the western tropical Pacific. Nevertheless, if this ~1 °C glacial cooling is accepted, our study implies that much larger variations in sea surface salinity can be extracted from the planktic foraminifer oxygen isotope data which show ~1.6‰ difference between the Holocene and LGM (Fig. 7). This is consistent with the current though that large changes in hydrological cycles analogous to ENSO are a dominant feature of the tropical Pacific (Cane and Clement, 1999; Clement and Cane, 1999; Clement et al., 1999; Stott et al., 2002; Koutavas et al., 2002; Visser et al., 2003). Although a ~1 °C discrepancy exists between our faunal and alkenone SST estimates of the LGM in core MD972151, this discrepancy may result from differential sensitivity of the faunal and alkenone methods, or from different optimum growing seasons of planktic foraminifers and coccoliths, and/or complex sedimentation patterns in the southern SCS. In any case, our results provide a more conservative estimate of LGM cooling in the western tropical Pacific and will have to be considered in further assessments of surface ocean climate in the glacial tropics.

Acknowledgements

This research was supported by the National Science Council (NSC92-2611-M-019-016), Academia Sinica (Asian Paleoenvironmental Changes (APEC) Projects), and the National Taiwan Ocean University, Republic of China. We thank Thorsten Kiefer and Tim Barrows for their constructive reviews.

References

Barrows, T.T., Juggins, S., De Deckker, P., Thiede, J., Martinez, J.I., 2000. Sea-surface temperatures of the southwest Pacific Ocean during the last glacial maximum. Paleoceanography 15, 95–109.

Be, A.W.H., 1967. Foraminifera families: Globigerinide and Globorotaliidae. Fiches d'Idendification du Zooplancton. Fraser, J.H. Cons., Int. Explor. Mer, Charlottenlund: Sheet 118.

Beck, J.W., Edwards, R.L., Ito, E., Taylor, F.W., Recy, J., Rougerie, F., Joannot, P., Heinin, C., 1992. Sea-surface temperature from coral skeletal strontium/calcium ratios. Science 257, 644–647.

Bush, A.B.G., Philander, S.G.H., 1999. The climate of the last glacial maximum: results from a coupled atmosphere-ocean general circulation model. Journal of Geophysical Research 104, 24509–24525.

Cane, M., Clement, A.C., 1999. A role for the tropical Pacific coupled ocean-atmosphere system on Milankovitch and millennial timescales. Part II: Global Impacts, American Geophysical Union, pp. 373–384.

Chen, M.T., Ho, H.W., Lai, T.D., Zheng, L., Miao, Q., Shea, K.S., Chen, M.P., Wang, P., Wei, K.Y., Huang, C.Y., 1998. Recent planktonic foraminifers and their relationships to surface ocean hydrography of the south China sea. Marine Geology 146, 173–190.

CLIMAP Project Members, 1976. The surface of the ice-age Earth. Science 191, 1131–1137.

CLIMAP Project Members, 1981. Seasonal reconstructions of the Earth's surface at the last glacial maximum. Geological Society of America Map and Chart Series, MC-36, Geological Society of America, Boulder, CO.

Clement, A.C., Cane, M., 1999. A role for the tropical Pacific coupled ocean-atmosphere system on Milankovitch and millennial timescales. Part I: a modeling study of tropical Pacific variability, American Geophysical Union, pp. 363–371.

Clement, A.C., Seager, R., Cane, M.A., 1999. Orbital controls on the El Niño/Southern Oscillation and the tropical climate. Paleoceanography 14 (4), 441–456.

Coulbourn, W.T., Parker, F.L., Berger, W.H., 1980. Faunal and solution patterns of planktonic foraminifera in surface sediments of the North Pacific. Marine Micropaleontology 5, 329–399.

Guilderson, T., Fairbanks, R.G., Rubenstone, J.L., 1994. Tropical temperature variations since 20,000 years ago: modulating interhemispheric climate change. Science 263, 663–665.

Hewitt, C.D., Broccoli, A.J., Mitchell, J.F.B., Stouffer, R.J., 2001. A coupled model study of the last glacial maximum: was part of the North Atlantic relatively warm? Geophysical Research Letters 28, 1571–1574.

Huang, C.-Y., Wang, C.-C., Zhao, M., 1999. High-resolution Carbonate Stratigraphy of IMAGES Core MD972151 from South China sea. The Journal of Terrestrial, Atmospheric, and Oceanic Sciences 10, 225–238.

Huang, C.C., Chen, M.T., Lee, M.Y., Wei, W.Y., Huang, C.Y., 2002. Planktic foraminifer faunal sea surface temperature records of the past two glacial terminations in the South China Sea near Wan-An shallow (IMAGES core MD972151). Western Pacific Earth Sciences 2, 1–4.

Hutson, W.H., 1980. The Agulhas Current during the Late Pleistocene: analysis of modern faunal analogs. Science 207, 64–66.

Hutson, W.H., Prell, W.L., 1980. A paleoecological transfer function, FI-2, for Indian Ocean planktonic foraminifera. Paleontology 54, 381–399.

Imbrie, J., Kipp, N.G., 1971. A new micropaleontological method for quantitative paleoclimatology: application to a Late Pleistocene Caribbean core. In: Turekian, K.K. (Ed.), The Late Cenozoic Glacial Ages. Yale University Press, New Haven, pp. 71–181.

Kiefer, T., Kienast, M., 2004. The deglacial warming in the Pacific Ocean: a review with emphasis on Heinrich Event 1. Quaternary Science Reviews, this issue (doi: 10.1016/j.quascirev. 2004.02.021).

Kiefer, T., Sarnthein, M., Erlenkeuser, H., Grootes, P.M., Roberts, A.P., 2001. North Pacific response to millennial-scale changes in ocean circulation over the last 60 kyr. Paleoceanography 16, 179–189.

Kienast, M., Steinke, S., Stattegger, K., Calvert, S.E., 2001. Synchronous tropical South China Sea SST change and Greenland warming during deglaciation. Science 291, 2132–2134.

Kipp, N.G., 1976. New transfer function for estimating past sea-surface conditions from sea-bed distribution of planktonic foraminiferal assemblages in the North Atlantic. Memoir of Geological Society of America 3–41.

Koutavas, A., Lynch-Stieglitz, J., Marchitto Jr., T.M., Sachs, J.P., 2002. El Niño-like pattern in ice age tropical Pacific sea surface temperature. Science 297, 226–230.

Kucera, M., Weinelt, M., Kiefer, T., Pflaumann, U., Hayes, A., Weinelt, M., Chen, M.-T., Mix, A.C., Barrows, T., Cortijo, E., Duprat, J., Waelbroeck, C., 2004. Reconstruction of sea-surface temperatures from assemblages of planktonic foraminifera: multi-technique approach based on geographically constrained calibration datasets and its application to glacial Atlantic and Pacific Oceans. Quaternary Science Reviews, this issue (doi: 10.1016/ j.quascirev.2004.07.014).

Le, J., 1992. Palaeotemperature estimation methods: sensitivity test on two western equatorial Pacific cores. Quaternary Science Reviews 11, 801–820.

Lea, D.W., Pak, D.K., Spero, H.J., 2000. Climate impact of Late Quaternary equatorial Pacific sea surface temperature variations. Science 289, 1719–1724.

Lee, T.-Q., Environmental magnetic record of deep-sea core MD972151 from the southwestern South China Sea, in preparation.

Lee, M.-Y., Wei, K.-Y., Chen, Y.-G., 1999. High resolution oxygen isotope stratigraphy for the last 150,000 years in the southern South China Sea: core MD972151. The Journal of Terrestrial, Atmospheric, and Oceanic Sciences 10, 239–254.

Malmgren, B.A., Nordlund, U., 1997. Application of artificial neural networks to paleoceanographic data. Palaeogeography, Palaeoclimatology, Palaeoecology 136, 359–373.

Malmgren, B.A., Kucera, M., Nyberg, J., Waelbroeck, C., 2001. Comparison of statistical and artificial neural network techniques for climating past sea surface temperatures from planktonic foraminifer. Paleoceanography 16, 1–11.

Miao, Q., Thunell, R.C., 1994. Glacial-Holocene carbonate dissolution and sea surface temperatures in the South China and Sulu Seas. Paleoceanography 9, 269–290.

Mix, A.C., Bard, E., Schneider, R., 2001. Environmental processes of the ice age: land, oceans, glaciers (EPILOG). Quaternary Science Reviews 20, 627–657.

Mix, A.C., Morey, A.E., Pisias, N.G., Hosterler, S.W., 1999. Foraminiferal faunal estimates of paleotemperature: circumventing the no-analog problem yields cool ice age trophics. Paleoceanography 14 (3), 350–359.

Ortiz, J.D., Mix, A.C., 1997. Comparison of Imbrie–Kipp transfer function and modern analog temperature estimates using sediment trap and core top foraminiferal faunas. Paleoceanography 12 (2), 175–190.

Parker, F.L., 1962. Planktonic foraminiferal species in Pacific sediments. Micropaleontology 8, 219–254.

Pelejero, C., Grimalt, J.O., Sarnthein, M., Wang, L., Flores, J.-A., 1999. Molecular biomarker record of sea surface temperature and climatic change in the South China Sea during the last 140,000 years. Marine Geology 156, 109–121.

Pflaumann, U., Jian, Z., 1999. Modern distribution patterns of planktonic foraminifera in the South China Sea and western Pacific: a new transfer technique to estimate regional sea-surface temperatures. Marine Geology 156, 41–83.

Pflaumann, U., Duprat, J., Pujol, C., Labeyrie, L.D., 1996. SIMMAX: a modern analog technique to deduce Atlantic sea surface temperatures from planktonic foraminifera in deep-sea sediments. Paleoceanography 11 (1), 15–35.

Prell, W.L., 1985. The stability of low-latitude sea-surface temperatures: and evaluation of the CLIMAP reconstruction with emphasis on the positive SST anomalies. United States Department of Energy, Office of Energy Research, TR025, US, Government Printing Office, vol. 1–2, pp. 1–60.

Prell, W., Martin, A., Cullen, J., Trend, M., 1999. The Brown University Foraminiferal Data Base. IGBP PAGES/World Data Center-A for Paleoclimatology Data Contribution Series # 1999-027.

Rind, D., Peteet, D., 1985. Terrestrial conditions at the last glacial maximum and CLIMAP sea-surface temperature estimates: are they consistent? Quaternary Research 24, 1–22.

Rosenthal, Y., Oppo, D.W., Linsley, B.K., 2003. The amplitude and phasing of climate change during the last deglaciation in the Sulu Sea, western equatorial Pacific, Geophysical Research Letters, 30(8), 2002GL016612.

Shin, S.-I., Liu, Z., Otto-Bliesner, B., Brady, E.C., Kutzbach, J.E., Harrison, S.P., 2003. A simulation of the last glacial maximum climate using the NCAR-CCSM. Climate Dynamics 20, 127–151.

Stott, L., Poulsen, C., Lund, S., Thunell, R., 2002. Super ENSO and global climate oscillations at millennial time scales. Science 297, 222–226.

Thiede, J., Nees, S., Schulz, H., De Deckker, P., 1997. Organic surface conditions recorded on the sea floor of the southwest Pacific Ocean through the distribution of foraminifers and biogenic silica. Palaeogeography, Palaeoclimatology, Palaeoecology 131, 207–239.

Thompson, P.R., 1976. Planktonic foraminiferal dissolution and the progress towards a Pleistocene equatorial Pacific transfer function. Journal of Foraminiferal Research 6, 208–227.

Thompson, P.R., 1981. Planktonic foraminifera in the western north Pacific during the past 150 000 years: comparison of modern and fossil assemblages. Palaeogeography, Palaeoclimatology, Palaeoecology 35, 241–279.

Ujiie, Y., Ujiie, H., 2000. Distribution and oceanographic relations of modern planktonic foraminifera in the Ryukyu Arc region, northwest Pacific Ocean. Journal of Foraminiferal Research 30 (4), 336–360.

Visser, K., Thunell, R., Stott, L., 2003. Magnitude and timing of temperature changes in the Indo-Pacific warm pool during deglaciation. Nature 421, 152–155.

Waelbroeck, C., Labeyrie, L., Duplessy, J.-C., Guiot, J., Labracherie, M., Leclair, H., Duprat, J., 1998. Improving past sea surface temperature estimates based on planktonic fossil faunas. Paleoceanography 13 (3), 272–283.

Webster, P.N., Streeten, N., 1978. Late Quaternary ice age climates of tropical Australia, interpretation and reconstruction. Quaternary Research 10, 279–309.

WOA, 1998. World Ocean Atlas 1998, version 2, http://www.nodc.noaa.gov/oc5/woa98.html. Technical report, National Oceanographic Data Center, Silver Spring, Maryland.