13. Rozanski, K., Araguas-Araguas, L. and Gonfiantini, R., Isotopic patterns in modern global precipitation. In *Climate Change in Continental Isotopic Records*, American Geophysical Union Monograph, 1993, vol. 78, pp. 1–36.
14. Longinelli, A. and Selmo, E., Isotopic composition of precipitation in Italy: a first overall map. *J. Hydrol.*, 2003, **270**, 75–88.
15. Clark, I. and Fritz, P., *Environmental Isotopes in Hydrology*, Lewis Publishers, Boca Raton, 1997, p. 328.
16. Simpkins, W. W., Isotopic composition of precipitation in central Iowa. *J. Hydrol.*, 1995, **172**, 185–207.
17. Fritz, P., Drimmie, R. J., Frape, S. K. and O'Shea, K., The isotopic composition of precipitation and groundwater in Canada. In International Symposium on the use of Isotope Techniques in Water Resources Development, IAEA Symposium, Vienna, Austria, 1987, vol. 299, pp. 539–550.
18. Bortolami, G. C., Ricci, B., Susella, G. F. and Zuppi, G. M., Maritime Alps, Pieddmont, Italy. In *Isotope Hydrology*, IAEA Symposium 228, Neuherberg, Germany, 1978, pp. 327–350.
19. Clark, I. D., Fritz, P., Michel, F. A. and Souther, J. G., Isotope hydrogeology and geothermometry of the Mount Meager geothermal area. *Can. J. Earth Sci.*, 1982, **19**, 1454–1473.
20. Fontes, J.-Ch. and Olivry, J. C., Gradient isotopique entre 0 et 4000 m dans les precipitations du Mont Cameroun. *C.R. Reunion Annu. Sci. Soc. Geol. Fr.*, Paris, 1977, p. 171.
21. Duplessy, J.-C., Be, A. W. H. and Blanc, P. L., Oxygen and carbon isotopic composition and biogeographic distribution of planktonic foraminifera in the Indian Ocean. *Palaeogeogr., Palaeoclimatol., Palaeoecol.*, 1981, **33**, 9–46.
22. Ostlund, H. *et al.*, *GEOSECS Atlantic, Pacific and Indian Ocean Expeditions, Vol. 7, Shore-Based Data and Graphics*, International Decade of Ocean Explr., Natl. Science Found., Washington DC, 1987.
23. Craig, H. and Gordon, L., Deuterium and oxygen-18 variation in the ocean and marine atmosphere. In *Stable Isotopes in Oceanographic Studies and Paleotemperatures* (ed. Tongiorgi, E.), Spoleto, 1965, pp. 9–130.
24. Dansgaard, W., Stable isotopes in precipitation. *Tellus*, 1964, **16**, 436–468.
25. Gonfiantini, R., Environmental isotopes in lake studies. In *Hand Book of Environmental Isotope Geochemistry, The Terrestial Environment* (eds Fritz, P. and Fontes, J.-Ch.), Elsevier, 1986, vol. 2, pp. 113–168.
26. Merlivat, L. and Jouzel, J., Global climatic interpretation of the deuterium–oxygen-18 relationship for precipitation. *J. Geophys. Res.*, 1979, **84**, 5029–5033.
27. Gat, J. R., Oxygen and hydrogen isotopes in the hydrologic cycle. *Annu. Rev. Earth Planet. Sci.*, 1996, **24**, 225–262.
28. Gat, J. R., Bowser, C. J. and Kendall, C., The contribution of evaporation from the Great Lakes to the continental atmosphere: estimate based on stable isotope data. *Geophys. Res. Lett.*, 1994, **21**, 557–560.
29. Machavaram, M. V. and Krishnamurthy, Earth surface evaporation process: a case study from the Great Lakes region of the United States based on deuterium excess in precipitation. *Geochim. Cosmochim. Acta*, 1995, **59**, 4279–4283.
30. Edwin Harvey, F., Stable hydrogen and oxygen isotope composition of precipitation in northeastern Colorado. *J. Am. Water Resour. Assoc.*, 2005, 447–459.
31. Indian Meteorological Division (IMD), Nungambakam, Chennai, 2006.
32. Vinayachandran, P. N., Shankar, D., Kurian, J., Durand, F. and Shenoi, S. S. C., Arabian Sea mini warm pool and the monsoon onset vortex. *Curr. Sci.*, 2007, **93**, 203–214.
33. Pankajakshan, T. and Rama Raju, D. V., Intrusion of Bay of Bengal water into Arabian Sea along the west coast of India during north east monsoon. In *Contributions in Marine Sciences* (eds Rao, T. S. S. *et al.*), Dr S. Z. Qasim Sastyabdapurti Felicitation Volume, National Institute of Oceanography, Goa, 1987, pp. 237–244.

ACKNOWLEDGEMENTS. We thank Dr James, CWRDM for help in the analysis of the samples. We also thank DST, New Delhi for financial assistance under IRHPA facility at Kozhikode and Dr S. M. Rao, BARC, Mumbai, for guidance and support. We are grateful to the reviewers for useful suggestions.

Received 2 August 2007; revised accepted 9 March 2009

Principal component and spectral analyses of palaeo-climate time series

K. P. C. Kaladhara Rao* and R. K. Tiwari

National Geophysical Research Institute, Hyderabad 500 007, India

Mathematical modelling and time series analysis techniques are important tools for extracting information from complex geotime series. These techniques also facilitate a fair degree of prediction, which is one of the prime goals of science. The data analysis strategy for such a purpose mainly involves spectral analysis and pattern classification. The aim of pattern classification and frequency analysis is to assign observations or patterns into semantic categories. Traditional statistical methods generally applied during the past years fail to recognize patterns from high dimensional georecords. Principal component analysis (PCA) is a powerful tool in identifying patterns in such records and provides useful means for reducing the number of dimensions without loss of much information. Here we have carried out spectral analysis and PCA of a climate record for approximately 28,000 yrs spanning from 1.15 to 29.78 kyr, off central Japan in the northwest Pacific. Our analysis reveals a dominant oscillation corresponding to the well known 'Heinrich Cycle'. The physical significance of the results has been discussed and the observed cyclic pattern corresponding to the global 'Heinrich Cycle' originating from the North Atlantic and Greenland ice rafting fluctuations has been linked to the Pacific phenomenon and Asian monsoon system.

Keywords: Heinrich Cycle, Last Glacial Maximum, palaeoclimate, principal component analysis, spectral analysis, time series.

THE climate system dynamics is a complex and coupled phenomenon because it is the result of interactions of various components of the land–ocean–atmosphere and cryosphere. Extracting physical information and making

*For correspondence. (e-mail: rao.kaladhar@gmail.com)

valid interpretation[1] using such complex palaeo-climate time series are therefore somewhat difficult, unless appropriate treatment of the data is made. Statistical models have been used for many years in the hope to predict geo-phenomena. As such, geophysical data acquired or measured contain certain unwanted information, which has to be either filtered or simply iterated so as to enable the observer to obtain the required information. The main application of the time series is to identify the hidden trends and periodicities from such complex time series and make a possible prediction. The mathematical methods that have been commonly used to identify periodicities are based on the assumption that the underlying time series is stationary and provides a density spectrum of the data. Hence, spectral analysis helps in identifying periodicities in counted or measured data, when the time series are presumably stationary and also long enough to contain at least four cycles. The highest peak in such a spectrum gives its frequency and power value, together with probability that the spectral peak could occur from the random/stochastic/cyclic processes. Appropriate statistical significance test of dominant spectral mode ensures the physical validity of the interpretation.

During the recent years, principal component analysis (PCA) has been widely used for identifying patterns from the complex geo-time series data. The utility of PCA lies in the fact that it can be used to infer variance of various components from the complex time series. This is particularly useful in applications where the complexity of the climate data makes it difficult to understand the variance of different harmonics. Here we have analysed a climate record for approximately the last ~28,000 yrs spanning from 1.15 to 29.78 kyr, which corresponds to the important geologic period of the Last Glacial Maximum (LGM). Our spectral and PCA analyses reveal a long-term harmonic in the data and provide important climate tele-connections between the oceanic and Indian monsoon system.

The LGM dates back to approximately 21,000 yrs and is the glacial period that witnessed large changes in the greenhouse gases, sea level and ice sheets. In the present study, we used the published down-core data of sea surface temperature (SST) decoded from the planktonic foraminifera Mg/Ca thermometry record off central Japan since the LGM[2]. The Mg/Ca ratio of this planktonic forminiferal calcite has been recently established as one of the significant palaeo-temperature proxy, since it can be used as a sample for oxygen isotope analysis also. The data have been taken from the Mg/Ca ratio and δ^{18}O values of these planktonic forminifera from a sediment core (MD01-2320) from the Kuroshio front, i.e. the boundary between the Kuroshio current and the convergence zone located south of 33°N off Central Japan in the northwest Pacific Ocean have been used to determine the SST data. Application of the Mg/Ca temperature calibration down-core measurements includes magnitude and timing of the SST changes for the past ~30,000 years spanning from 1.15 to 29.78 kyr, particularly northwest Pacific. SST data record and their spectrum are presented in Figure 1.

Before PCA, we applied Fourier spectral analysis to the data to find statistically significant harmonic components, if any. The spectral results also are displayed in Figure 1. In Figure 1, the dotted line shows 95% statistical significance level and it is evident that there is only one strong peak at around 4500 ± 500 kyr, which is statistically significant at 95% confidence level. This might correspond to the well-known Heinrich Cycle. The spectral result shows peak frequency value: 0.9867, peak power: 25.22, *P* (random): 1.493 E–09. 0.01 level: 9.503, 0.05 levels: 7.868. There are other oscillations also, but these are below the confidence level and hence may be treated as statistically and physically irrelevant. It is interesting to note here that similar or almost identical periodicities have also been reported in the Asian monsoon system record[3]. This might suggest a possible tele-connection between the Heinrich Cycle and the East Asian winter monsoon system. By definition, Heinrich events are restricted to the North Atlantic, but there is evidence for 6.00 to 1.5 kyr Heinrich Cycle in the global climate records. This correlates with major changes in the global climate system from North Atlantic and Antarctica. The Greenland temperature also oscillates and coin-

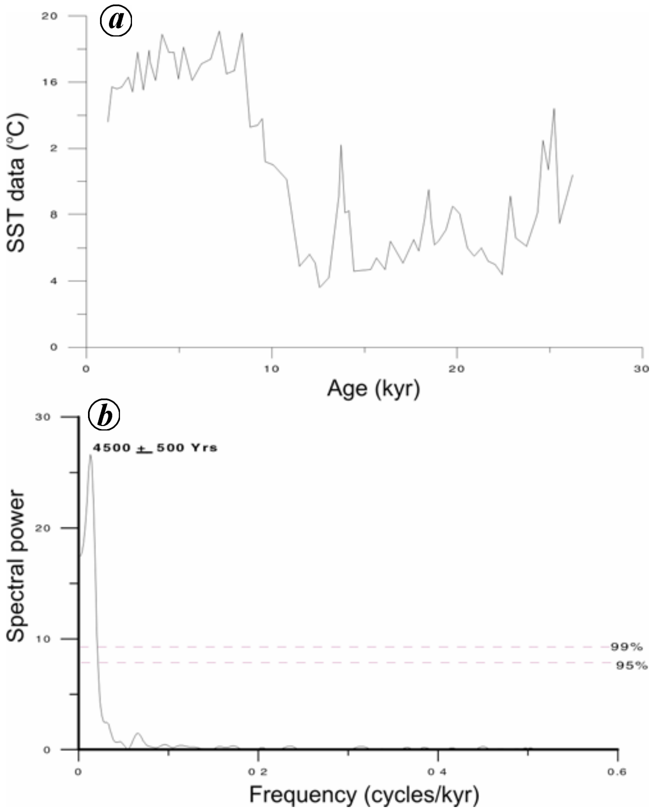

Figure 1. ***a***, Mg/Ca-based sea surface temperature (SST) data (after Sagawa *et al.*[1]). ***b***, Spectral analysis of the SST data. Dotted lines show the 99 and 95% confidence level.

cides with the Heinrich Cycles, with major changes in the global climate dynamics[3–5].

In the PCA we attempt to explain the total variability of correlated variables through the use of orthogonal principal components. The components themselves are merely weighted linear combinations of the original variables.

The first principal component can be expressed as follows:

$$Y_1 = a_{11}X_1 + a_{21}X_2 + \cdots + a_{p1}X_p,$$

or in matrix form

$$Y_1 = a'x.$$

The a_{j1} are scaled such that $a^{T}a_1 = 1$. Y_1 accounts for the maximum variability of the variables of any linear combination. The variance of Y_1 is λ_1.

Next, principal component Y_2 is formed such that its variance, λ_2, is the maximum amount of the remaining variance and that it is orthogonal to the first principal component. That is, $a_1'a_2 = 0$.

One continues to extract components until some stopping criteria are encountered or the until required number of components are formed. It is possible to compute principal components from either the covariance matrix or correlation matrix of the same number of variables. If the variables are scaled in a similar manner, then many researchers prefer to use the covariance matrix. When the variables are scaled different from one another, then using the correlation matrix is preferred. Common stopping criteria when using the correlation matrix are to stop when the variance of a component is less than one.

Generally, the weights used to create the principal components are the eigenvectors of the characteristic equation:

$$(S - \lambda_i I)a = 0 \quad \text{or} \quad (R - \lambda_i I)a = 0.$$

where S is the covariance matrix and R is the correlation matrix. λ_i are the eigenvalues, the variances of the components.

We applied PCA to identify a pattern in the data under observation and expressing the data in such a way in order to highlight their similarities and differences. PCA is also known as the empirical orthogonal function (EOF) analysis, or an orthogonal transformation which transforms the data in question to a new coordinate system such that the greatest variance in the projection of the data comes to lie on the first coordinate called the first principal component, the second greatest variance on the second coordinate and so on[6]. Theoretically speaking, PCA is the optimum transform for a given data in least square terms. In PCA we subtract the mean to obtain the average across each dimension or projection; thereby covariance matrix would be calculated drawing eigenvectors and eigenvalues of the covariance matrix. Thereafter, we choose components and form a feature of the vector deriving a new dataset. Thus, basically we transform our data so that they are expressed in terms of the patterns between them, where the patterns are lines that mostly closely describe the relationships between the data. This is useful because we have now classified our datapoints as a combination of the contributions from each of those lines of the dataset. Thus PCA is a technique used to reduce multidimensional datasets to lower dimension for analysis in exploratory data analysis and for making predictive models involving calculation of the eigenvalue decomposition or singular value decomposition of a dataset, usually after mean centring the data for each attribute. The results of the PCA are usually discussed in terms of component scores and loadings[7,8].

The results of the PCA are displayed in Figure 2. We have considered three principal components for the SST record, viz. PC1, PC2, PC3. Figure 2 shows the patterns of decomposition. The spectral periodicities and principal components indicate that the first component consists of almost 93% of total variance which signifies the low frequency period suggesting that only one cycle of the period ~4500 ± 500 kyr is statistically significant. Both analyses provide a robust time constant in global climate system records, which seems to be of fundamental importance. Spectral analyses of the sea surface data thus pro-

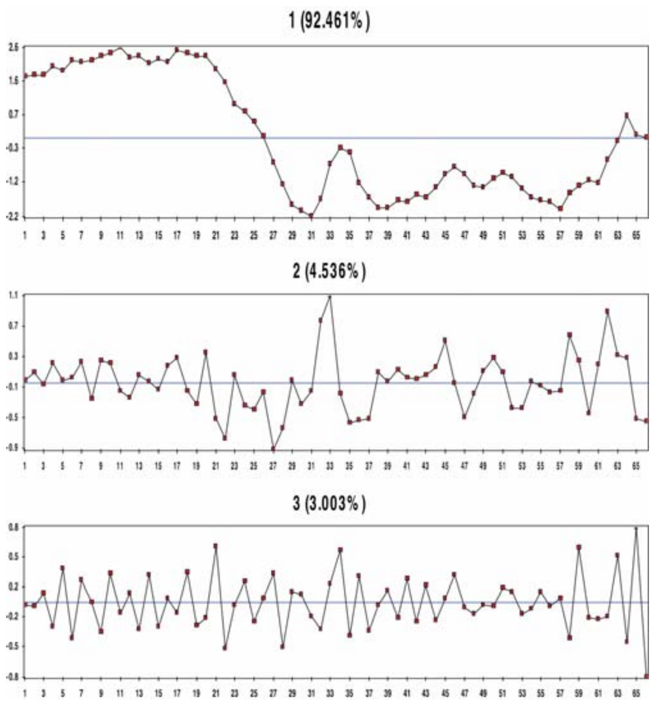

Figure 2. Principal component analysis of SST data showing percentage variance of the three principal components PC1 = 92.461, PC2 = 4.536 and PC3 = 3.003 respectively.

vide good correlation with the PCA coordinates. We used Caterpillar Software developed by the Caterpillar Group, 1997, for the PCA.

It has been observed that the total variance of these principal components is 100%, based on which the time series was reconstructed along with their residuals (Figure 3). The results of the PCA indicate the dominant temporal and spatial variability patterns such as Heinrich Cycle with relevant physical processes. The first principal component (92.461%) explains the majority of variance, the second principal component (4.536%), the third principal component (3.003%) and so on.

The LGM approximately 21,000 yrs ago, is a glacial period and forcing phenomenon with large changes in the greenhouse gases, sea level and ice sheets. The late glacial warming prior to Heinrich Event-I has influenced the ice rafting and large ice sheets were displaced in the northern hemisphere[9]. The freshwater flux originating from the North Atlantic affected the density of sea surface water, which in turn, controlled the large scale convection occurring in the northernmost part of the Atlantic Ocean. The resulting 'overturning' ventilates the deep layers of the global oceanic system[10,11]. This overturning is widely known as the North Atlantic Deep Water (NADW) flux, which belongs to wider global circulation system known as thermo-haline circulation and represented as a conveyer belt that links the world ocean basins. These cold ocean waves influenced the proximal terrestrial climates, affecting remote regions with respect to monsoon and seasons at almost regular intervals of 4000–5000 yrs and are widely known as the Heinrich Cycle. The spectral analysis and PCA of SST data from central Japan in the northwest Pacific, reveals statistically significant quasi-periodicity of 4000–5000 yrs identical to the global Heinrich cycle. In view of the common cyclic mode and physical link as discussed above, our results might suggest that temporal and spatial variability of Japan SST could have been tele-connected to the larger system of coupled land–ocean–atmosphere cryosphere. The results will have significant implication for understanding the link between the global oceanic thermohaline circulation pattern and the Asian monsoon system[12].

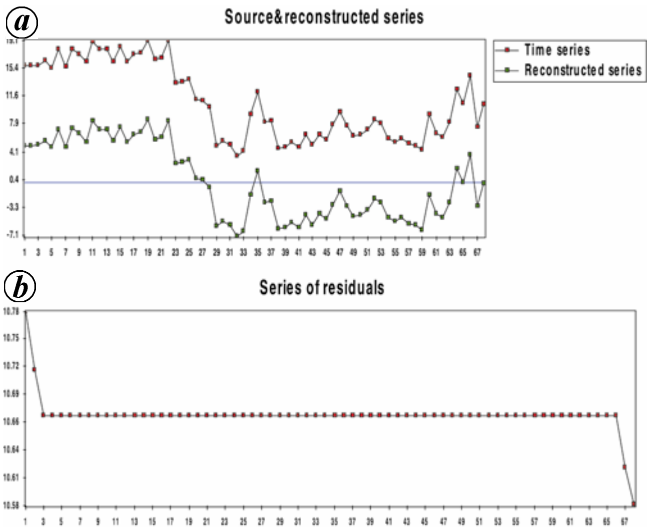

Figure 3. ***a***, Original and reconstructed SST data. ***b***, Residuals of SST data.

1. Wu, A., Hsieh, W. W. and Tang, B., Neural network forecasts of the tropical Pacific sea surface temperatures. *Neural Networks*, 2006, **19**, 145–154.
2. Sagawa, T., Toyoda, K. and Oba, T., Sea surface temperature record off central Japan since the Last Glacial Maximum using planktonic foraminiferal Mg/Ca thermometry. *J. Quat. Sci.*, 2006, **21**, 63–73.
3. Keifer, K. and Kienast, M., Patterns of deglacial warming in the Pacific Ocean: a review with emphasis on the time interval of Heinrich event I. *Quat. Sci. Rev.*, 2005, **24**, 1063–1081.
4. Bard, E., Rostek, F., Turon, J. L. and Gendreau, S., Hydrological impact of Heinrich events in subtropical northeast Atlantic. *Science*, 2000, **289**, 1321–1324.
5. Tangang, F. T., Hsieh, W. W. and Tang, B., Forecasting the equatorial Pacific sea surface temperatures by neural network models. *Climate Dyn.*, 1997, **13**, 135–147.
6. Wilson-Diaz, D., Mriano, A. J., Evans, R. H. and Luther, M. E., A principal component analysis of sea surface temperature in the Arabian Sea. *Deep-Sea Res. II*, 2001, **48**, 1097–1114.
7. Monge Sanz, B. M. and Medrano Marques, N. J., Total ozone time series analysis: a neural network model approach. *Nonlinear Process Geophys.*, 2004, **11**, 683–689.
8. Agarwal, N., Kishawal, C. M. and Pal, P. K., An analogue prediction method for global sea surface temperature. *Curr. Sci.*, 2001, **80**, 49–55.
9. Calvo, E., Pelejero, C., Herguera, J. C., Palanques, A. and Grimalt, J. O., Insolation independence of the southeastern subtropical Pacific sea surface temperature over last 400 kyrs. *Geophys. Res. Lett.*, 2001, **28**, 2481–2484.
10. Broecker, W. S. *et al.*, Origin of Northern Atlantic's Heinrich events. *Climate Dyn.*, 1992, **6**, 265–273.
11. Broecker, W. S., Peteet, D. M. and Rind, D., Does the ocean–atmosphere system have more than one stable mode of operation? *Nature*, 1985, **315**, 21–26.
12. Agnihotri, R., Dutta, K., Bhushan, R. and Somayajulu, B. L. K., Evidence for solar forcing on the Indian monsoon during the last millennium. *Earth Planet. Sci. Lett.*, 2002, **198**, 521–527.

ACKNOWLEDGEMENT. We thank the Director, NGRI, Hyderabad, for permission to publish this paper.

Received 4 August 2008; revised accepted 19 March 2009