12.740 Paleoceanography
Spring 2008

Paleo-ecological temperature estimation

12.740 Lecture 4 Spring 2008

Paleo-ecological temperature estimation

Paleo-ecological temperature estimation is based on the empirical observation that certain species of marine organisms live and leave fossils on the seafloor that correspond to temperature patterns in oceanic surface waters. Experimental manipulations of organisms show that many have preferred temperature ranges. Downcore records confirm known glacial/interglacial fluctuations.

Experimental growth-T relationships
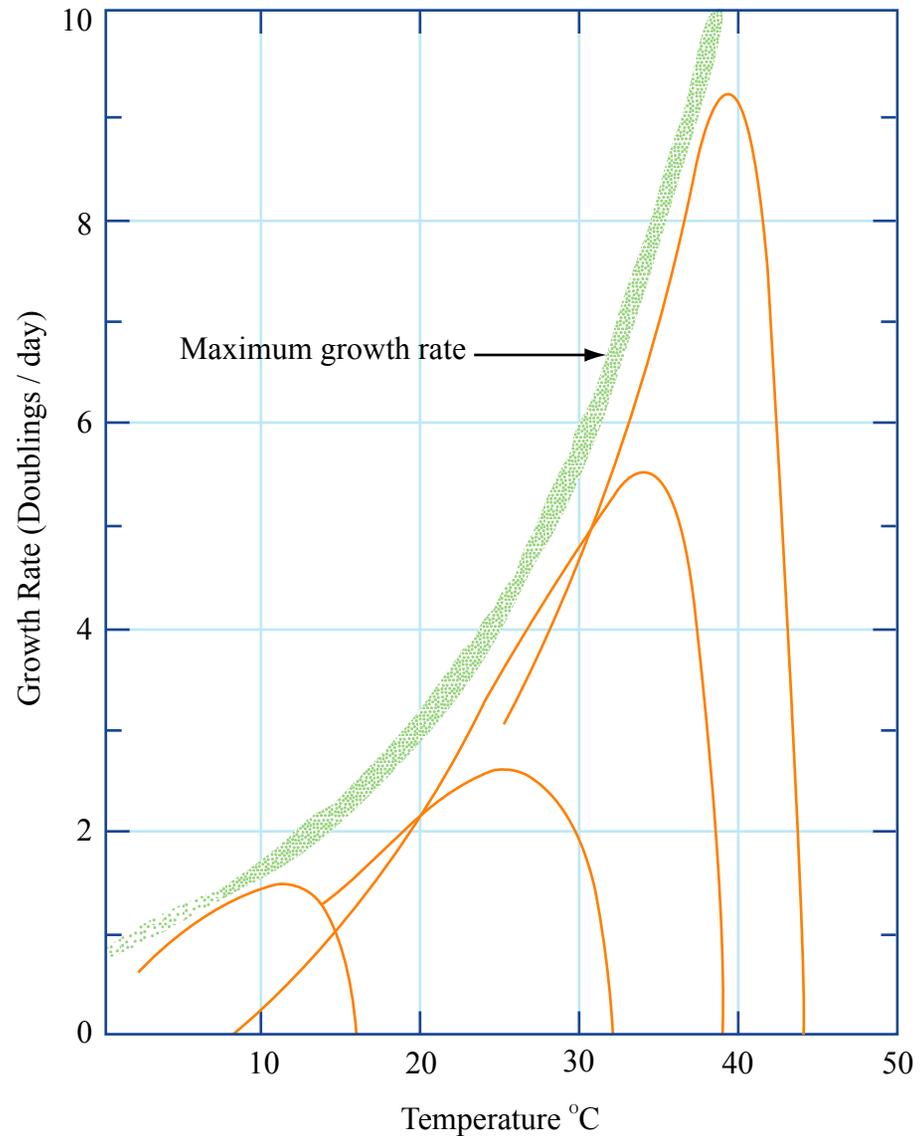
Growth rates of some species of phytoplankton as a function of temperature. At optimum conditions for growth, warm-adapted plankton have a greater rate of production than do cold-adapted ones. Curves based on responses of certain diatoms and chlorophytes (left to right *Detonula confervacea. Dit ylum brightwellii. Dunaliella tertiolecta. Chlorella pyrenoidosa*).

from Wolfgang Berger

Figure by MIT OpenCourseWare.

Problem: there are more than 30 planktonic foraminifera species

- You have to randomly split sediment samples down to 300-600 specimens and then examine them under a binocular microscope. Sometimes this requires rolling the specimen over. Even today, computer-aided recognition isn't up to the task.

- So the method of counting all forams is slow and tedious.

- To get around this problem, Erikson developed methods based on single indicator species,

  e.g. *N. pachyderma* (L) or *G. menardii.*

Erikson faunal vs. Emiliani isotopic records from the Caribbean Sea

Image removed due to copyright restrictions.
Please see : Imbrie and Imbrie (1979) *Ice Ages: Solving the Mystery*, Figure 33.

How does one deal with complete faunal counts?

- As you have just seen, plotting each species percentages vs depth results in a visually tedious presentation whose meaning is hard to grasp.

- John Imbrie's solution: statistical factor analysis

Variance and Covariance

- Variance (of x is denoted S_x^2; variance of y is denoted S_y^2) is a measure of the scatter of values of a variable about its mean:

$$S_x^2 = \sum_{i=1}^{n} (x_i - \overline{x})^2 / n$$

- Covariance (of x and y) expresses the relationship between two variables (a measure of the scatter of values of points in a plane relative to the centroid of the data set):

$$S_{xy}^2 = \sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y}) / n$$

Correlation Coefficient

$$r = \frac{1}{n-1}\sum_{i=1}^{n}(\frac{x_i - \bar{x}}{s_x})(\frac{y_i - \bar{y}}{s_y}) = \frac{1}{n-1}\frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{s_x s_y}$$

where s_x = std. dev. of x = $[\Sigma(x_i - x)^2 / (n-1)]^{1/2}$

and s_y = std. dev. of y = $[\Sigma(y_i - y)^2/ (n-1)]^{1/2}$

- s_x and s_y relate the deviations of points from the average relative to the "range" (actually std. dev.) of the observations.

- r^2 is "the variance of Y accounted for by its covariance with x" (usually expressed in % units).

- In other words, r^2 is the Covariance divided by the Variance.

Linear regression

1. Common linear regression of y on x: y = A + Bx

Let $S = \Sigma (y_i - A - B\, x_i)^2$

To minimize S, set $\partial S/\partial A = 0$; $\partial S/\partial B = 0$; solve for A and B.

2. Matrix math solution of Linear Regression:

for eq'n Ax = b (m eq'ns, n unknowns),
if columns of A are linearly independent,
then:

$$\underline{x} = (A^T A)^{-1} A^T \underline{b}$$

For example, for the simple linear regression
$$y = C + Dx,$$
where we want to fit pairs of data
$$x_i, y_i$$
we want to find
$$C, D$$
that mimimize
$$\Sigma \, [y_i - (C + Dx_i)]^2$$
In matrix form, we write the equation $y = C + Dx$ as:

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ & \cdot \cdot \\ & \cdot \cdot \\ 1 & x_n \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_n \end{bmatrix}$$

i.e.   A   x = b

Numerical example:

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 4.1 \\ 5.9 \\ 7.8 \\ 10.3 \\ 12.1 \end{bmatrix}$$
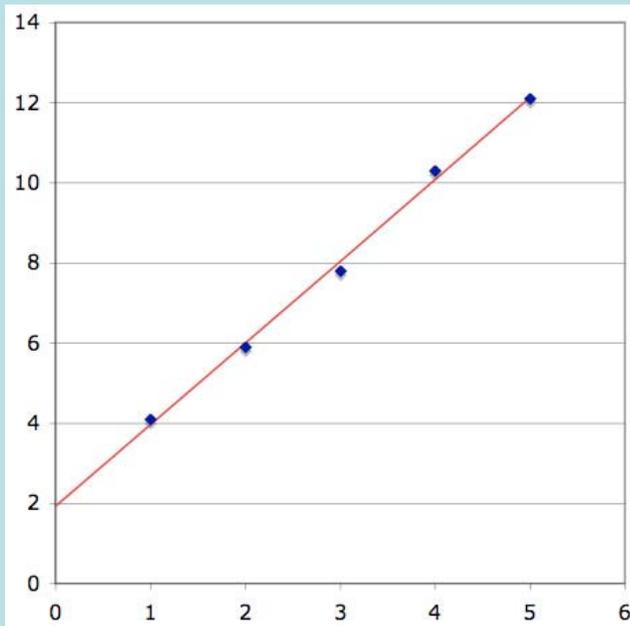
$$\underline{x} = (A^T A)^{-1} A^T \underline{b}$$

gives C=1.92, D=2.04

Similarly, to solve the equation $y = A + Bx + Cx^2$:

$$\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ & . & . \\ & . & . \\ 1 & x_n & x_n^2 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ y_n \end{bmatrix}$$

Simple matrix formulas also allow you to compute the estimated uncertainties of the regression coefficients and the correlation coefficients.

Correlation in n dimensions

- Multiple linear regression

    e.g. $y = A + Bx_1 + Cx_2$

    (the equation for a plane in 3D space)

$$\begin{bmatrix} 1 & x_{1\text{-}1} & x_{2\text{-}1} \\ 1 & x_{1\text{-}2} & x_{2\text{-}2} \\ & . & . \\ & . & . \\ 1 & x_{1\text{-}n} & x_{2\text{-}n} \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ y_n \end{bmatrix}$$

- r-matrix (later we will sometimes refer to this as the matrix Σ)

Property	1	2	3	4	5
1	1.00	0.86	0.45	0.83	0.45
2	0.86	1.00	0.74	0.23	0.64
3	0.45	0.74	1.00	0.78	0.57
4	0.83	0.23	0.78	1.00	0.39
5	0.45	0.64	0.57	0.39	1.00

(ρ_{ij})

Ellipsoids and eigenvectors

Bivariate scatter diagram with equal-density contour lines. The inner ellipse encloses 66% of the data points, The outer ellipse 95% of the data points.
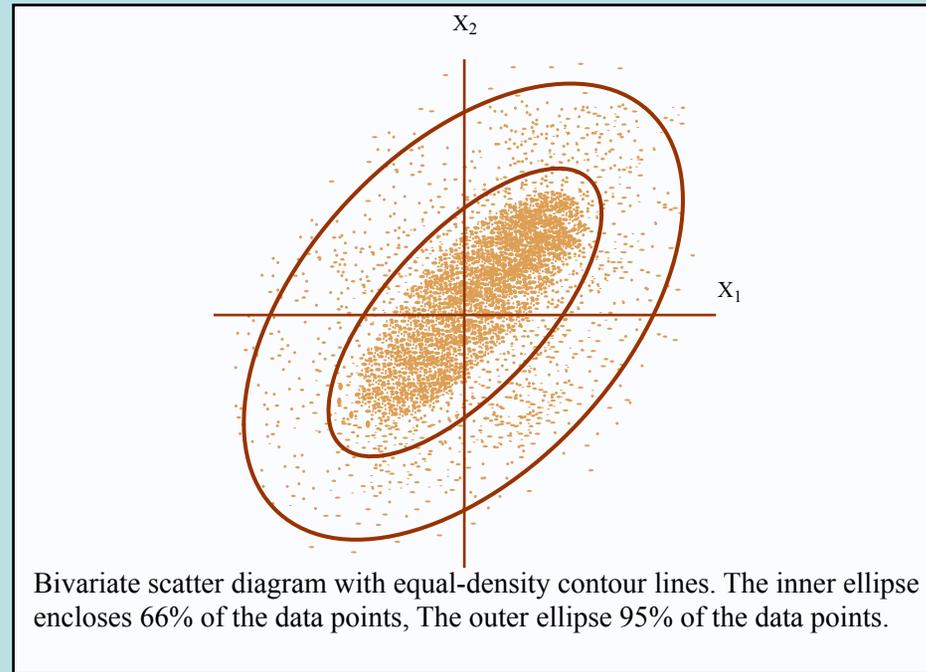
Figure by MIT OpenCourseWare. Adapted from source: Joreskog et al. *Geological Factor Analysis* (1976).

$$\mathbf{A} \quad \underline{\mathbf{x}} \quad = \quad \lambda \quad \underline{\mathbf{x}}$$

square matrix	eigen-vector		eigen-value	eigen-vector

One way to find eigenvectors:

1st eigenvector = major axis of ellipsoid
2nd eigenvector = largest minor axis of ellipsoid
   … etc.

This works more or less as if we did a regression to get a line that "explains" most of the variance (the dominant linear trend of the data in n-dimensional space), subtracted that regression from the data, then perform another regression to find the next most important contribution to the variance, and so forth.

This procedure works because we can rewrite the equation as:

$$(A - \lambda I)\,x = 0$$

where I is the identity matrix:

$$\begin{bmatrix} 1 & 0 & 0... \\ 0 & 1 & 0... \\ 0 & 0 & 1... \\ ...... \end{bmatrix}$$

In other words, the unknown vector x is orthogonal to all row vectors of $(A-\lambda I)$.

The expression for the determinant of $A-\lambda I$ is a polynomial of degree equal to the number of rows and columns of the square matrix A. The roots of the polynomial are the eigenvalues. If A is a real, square, and symmetric matrix, the roots are always real. However, these p eigenvalues may not always be different and some may equal zero. If two or more eigenvalues are equal, we say that they are multiple eigenvalues; otherwise we say the eigenvalue is distinct.

Example: consider a 2x2 symmetrical singular (i.e., determinant = 0) matrix; the equation is then:

$$\begin{matrix} r_{11}-\lambda & r_{12} \\ r_{21} & r_{22}-\lambda \end{matrix} \quad = \quad 0$$

$$(r_{11}-\lambda)(r_{22}-\lambda) - r_{12}r_{21} = 0$$

$$\lambda^2 - \lambda(r_{11}+r_{22}) + (r_{11}r_{22} - r_{12}r_{21}) = 0$$

i.e., a simple quadratic equation (also: remember, by assumption of symmetry, $r_{12}=r_{21}$)

Things get out of hand quickly as the matrix gets bigger (that's what computers are for!).

Matrix collection of eigenvectors

- Once the eigenvalues are known, the eigenvectors can be calculated from (a.)

  A unique solution cannot be obtained for an eigenvector. If x is a solution, so is cx, where c is a scalar. By convention, eigenvectors are therefore always normalized (unit length).

  Eigenvectors associated with different eigenvalues are orthogonal. Multiple solutions are possible for multiple eigenvalues, but it is always possible to choose an eigenvector which is orthogonal to all other eigenvectors.

- If the eigenvalues λ are placed as the elements of a diagonal matrix Λ, and the eigenvectors are collected as columns into the matrix U, then the equation becomes:

$$A\,U = U\,\Lambda$$

$$
\begin{array}{c}
\\1\\2\\3\\\\\\\\n
\end{array}
\overset{1\,2\,3\ldots\ n}{
\begin{bmatrix}
 & & & \\
 & & & \\
n\times n & . & . & \\
\text{Square} & . & . & \\
\text{Matrix} & . & . & \\
 & . & . & \\
 & . & . &
\end{bmatrix}}
\overset{1\quad 2\ldots\quad n}{
\begin{bmatrix}
EV_1\ EV_2\ \ldots & \\
\downarrow\quad\downarrow & \\
. & . & . \\
. & . & \\
. & . & . \\
. & . &
\end{bmatrix}}
=
\overset{1\quad 2\ldots\quad n}{
\begin{bmatrix}
EV_1\ EV_2\ \ldots & \\
\downarrow\quad\downarrow & \\
 & & \\
 & & \\
 & & \\
 & &
\end{bmatrix}}
\overset{1\ 2\ 3\ldots\ n}{
\begin{bmatrix}
\lambda_1\ 0\ 0\ldots & \\
0\ \lambda_2\ 0\ldots & \\
 & & \\
 & & \\
 & & \\
 & & \lambda_n
\end{bmatrix}}
$$

The matrix U (the eigenvector matrix) is square orthonormal (the matrix is nxn and the eigenvectors are of unit length), so $U U^T = U^T U = I$, therefore:

$$A = U \Lambda U^T$$

Therefore any symmetric matrix such as the correlation coefficient table can be expressed in this form:

$$\Sigma \quad = \quad U \quad \Phi \quad U^T$$

m x m   m x N   N x N   N x m
r-table

          eigenvectors

              eigenvalues

                    eigenvectors(transpose)

Principle Components Analysis and Factor Analysis

- PCA is basically the eigenvectors. It is inherently variance-oriented; it accounts for maximum variance of all the observed variables. In other words, PCA accepts that large part of the total variance of a variable is important and common to other observed variables.

- Factor analysis (below) throws away the least significant eigenvectors and then rotates the remaining ones to maximize the variance explained. It is correlation-oriented; it accounts for the maximum intercorrelation of variables. In other words, factor analysis allows for a considerable "amount of uniqueness" to be present in the data and utilizes only that part of a variable that takes part in correlation with other variables; i.e., to account for covariance rather than variance.

- Example: imagine a lake with three stream systems draining into it. Suppose that each of the drainage basins has a distinctive (average) sediment composition, and that the sediments within the lake can be described as linear combinations of the sediments derived from each stream system. What is interesting in this case is the distribution of the three stream system component sediments, not the mineralogy or chemical composition per se (i.e., we don't care if Al correlates with Si). So we adopt a framework that lets us describe the raw data (chemical or mineralogical composition) in terms of sums of source components, and map the distribution of the components.


Source of example: Joreskog et al.
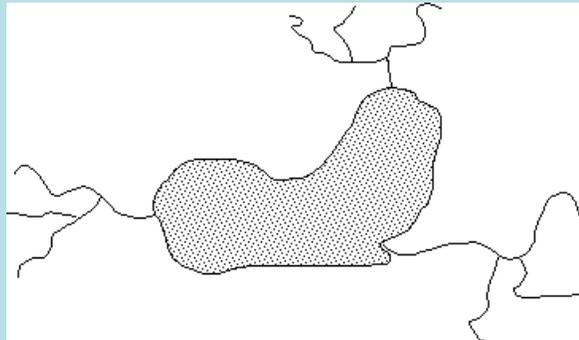
Species Space and Sample Space

- Samples in species space:

PROPERTY 2

Sample 1

PROPERTY 1

sample 2

PROPERTY 3

- Species in sample space:

sample 2

PROPERTY 1

sample 1

PROPERTY 2
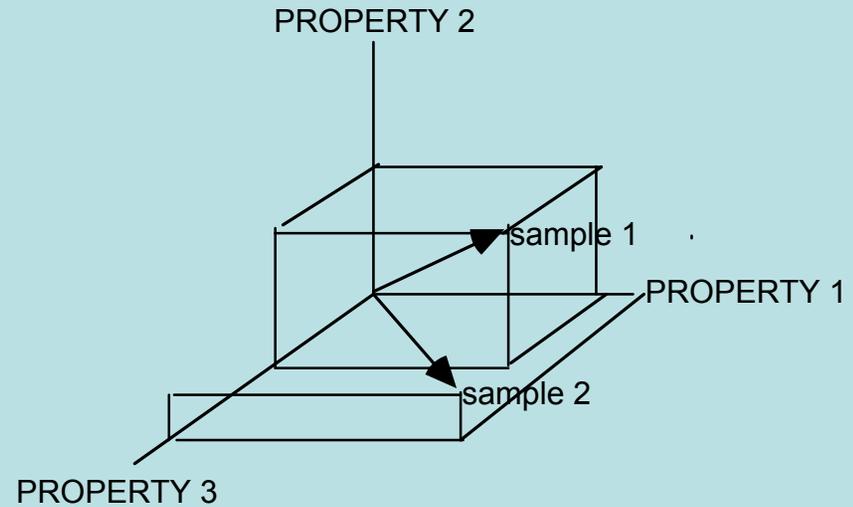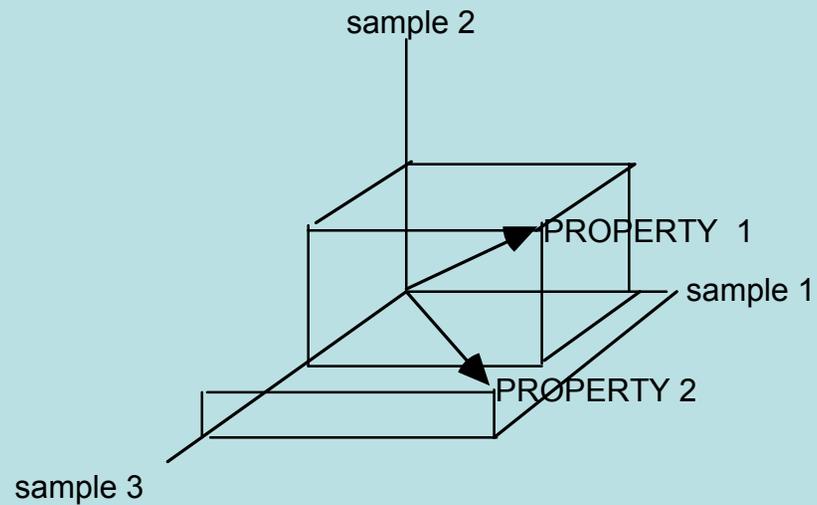
sample 3

What is sample space?

1. Consider the species percentage data:

	PaL	PaR	Gbu	Ndu
Sample 1	17.4	17.4	39.1	26.1
Sample 2	36.4	22.7	27.3	13.6
Average	26.9	20.1	33.2	19.9

2. Express each species percentage in terms of its deviation from the average for the samples:

	PaL	PaR	Gbu	Ndu
Sample 1	-9.5	-2.7	+5.9	+6.2
Sample 2	+9.5	+2.7	-5.9	-6.2

3. Transpose:

	Sample 1	Sample 2	length
PaL	-9.5	+9.5	13.42
PaR	-2.7	+2.7	3.77
Gbu	+5.9	-5.9	8.38
Ndu	+6.2	-6.2	8.80

4. Row-normalize (to unit vector length):

	Sample 1	Sample 2
PaL	-0.71	+0.71
PaR	-0.71	+0.71
Gbu	+0.71	-0.71
Ndu	+0.71	-0.71

Species as vectors in sample space:
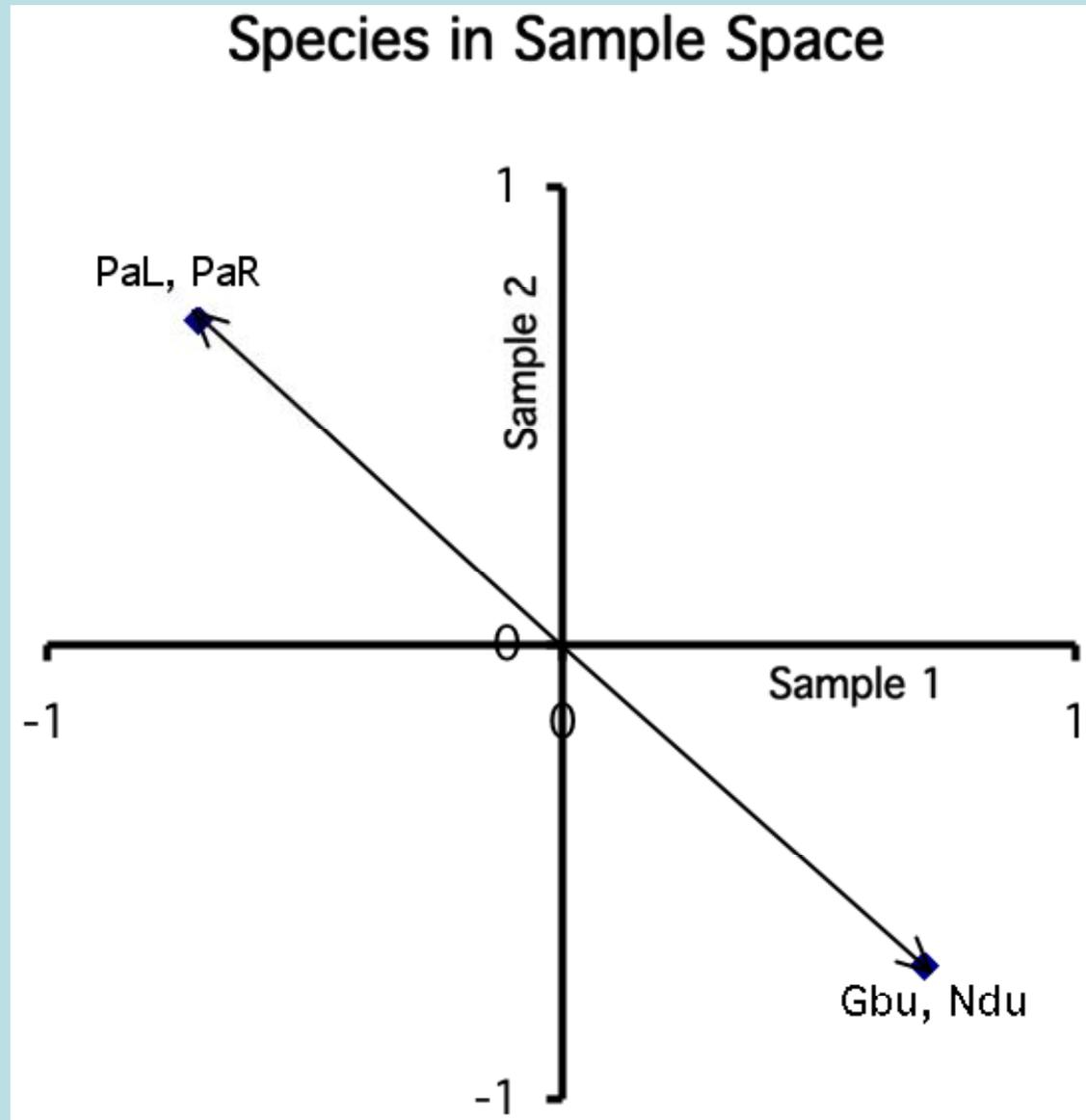
Note that the PaL and PaR vectors coincide, as do the Gbu and Ndu vectors.

In sample space, species that are highly correlated fall in the same vector region,
i.e., there is a very small <u>angle</u> between highly correlated species.

The correlation coefficient is the cosine of the angle between the two PROPERTY
vectors in sample space: Let a and b any two PROPERTY vectors (with a common origin).
Then:

$$\rho = \cos \theta = \frac{\mathbf{a^T b}}{\mathbf{a} \quad \mathbf{b}}$$

- What we are going to do next is reduce the number of "species" by combining them into "factors". In other words, we are going to reduce our system from having (say) 33 species counts for each sample to having (say) 6 "factor loadings" for each sample.

- Continuing on from equation II.C.4., a math theorem says that a symmetric matrix like the correlation matrix S can be expressed as:

- $$S = A\,F\,A^T + Y$$

  - where "S is the p by p population covariance matrix of the observed variables" (the symmetrical correlation coefficient matrix), "A is the p by k matrix of factor loadings, F is the k by k covariance matrix for the factors (if the factors are in standardized form, this is a correlation matrix with ones in the diagonal), and Y is the p by p residual covariance matrix."

    Source: Joreskog et al.

- If we force the solution to be orthogonal (factors are uncorrelated), then $F = I$ and hence:

- $$S = A\,A^T + Y$$

- Y is now a diagonal matrix (i.e. there is covariance between the factors)

A math theorem tells us how to calculate the **Singular Value Decomposition (SVD)** of a matrix:

Let us assume that we have an N x p (N>p) data matrix X. The product moments are defined as:

major product moment (MPM) $V = X X^T$
minor product moment (mPM) $U = X^T X$

Then we can decompose the data matrix as:

$$X = V G U$$

where

V is an N by r matrix with orthonormal columns
U is a p by r matrix with orthonormal columns, and
G is a diagonal matrix of order r by r with positive diagonal elements g_1, g_2, ...g_r called
singular values of X.

The major product moment XX^T, which is square, symmetric, and of order N by N has r positive eigenvalues and (N-r) zero eigenvalues. The positive eigenvalues are g_{12}, g_{22}, ..., g_{r2} and the corresponding eigenvectors are u_1, u_2, ..., u_r.

The analysis of the minor product moment X^TX is referred to as R-mode analysis and that of the major product moment XX^T as Q-mode analysis.

The minor product moment $X^T X$, which is square, symmetric, and of order p by p has r positive eigenvalues and (p-r) zero eigenvalues. The positive eigenvalues are g_{12}, g_{22}, ..., g_{r2} and the corresponding eigenvectors are v_1, v_2, ..., v_r.

The positive eigenvalues of $X X^T$ and $X^T X$ are the same, namely g_{12}, g_{22}, ..., g_{r2} . Furthermore, if vm is an eigenvector of $X X^T$ and u_m and eigenvector of $X^T X$ corresponding to one of these eigenvalues g_{m2}, then the following relationships hold between u_m and v_m :

$v_m = (1/g_m) X u_m$ and $u_m = (1/g_m) X^T v_m$

These relationships make it possible to compute v_m from u_m and vice-versa, i.e.

$$V = X U G^{-1} \quad and \quad U = X^T V G^{-1}.$$

Major problem with this analysis: there are an infinite number of solutions! If we find a solution to the problem (a set of orthogonal vectors that describe the data), then any <u>rotation</u> of that set of vectors is also a solution! How do we choose any one of these?

Rotation of a set of row vectors in the matrix X can simply be done through the operation:

$$\mathbf{Y} = \mathbf{X} \mathbf{R}$$

where R is a rotation (transformation) matrix conforming to the requirement that $\mathbf{R} \mathbf{R}^T = \mathbf{I}$ for a rigid rotation.

In the end, the choice is <u>arbitrary</u> (this is one of the problems with factor analysis). However, certain more-or-less reasonable choices can be made:

> As much as possible, try to make the factors simple, i.e., have a few high loadings and many zero or near-zero loadings.

> Rotate the axes so as to put as much variance as possible into the factors (VARIMAX criterion). i.e., we try to "explain" as much of the variance as possible with the fewest possible factors. This solution is often favored because it is an objective solution (i.e., it arises untouched by human hands).
> That doesn't necessarily make it a <u>better</u> solution, however.

Imbrie-Kipp (1971) method: factor analysis and transfer functions

This method assumes that there are certain groupings of species (assemblages) which can be identified by linear multivariate statistics: orthogonal Q-mode factor analysis, and which remain coherent over time. It also assumes that correlations (of at least some of these groupings) maintain consistent relationships with environmental properties (e.g. surface temperature).

In an oversimplified form, the assemblages can be viewed of as grouping of species. Suppose you have three species, and three groupings:

	Species 1	Species 2	Species 3
Group 1	7	2	1
Group 2	4	3	3
Group 3	2	2	6

Then each sample would be decomposed into linear combinations of groups 1, 2, and 3. This obviously has little advantage unless the three groupings follow the environmental factors better than the individual species. However, if 33 species of foraminifera can be reduced to 6 groupings, a significant simplification has been achieved. But how should you group the species data? Imbrie's answer was to use linear statistical analysis.

The first step is a linear factor analysis of core-top (CT) data:

$$\mathbf{U_{CT}} = \mathbf{B_{CT}}\ \mathbf{F} + \mathbf{E}$$

Error matrix
*(N x n,
to allow for
imperfect fit)*

Row-normalized core top
data matrix
(N samples, n species)

Varimax matrix:
*(N samples described as
linear combinations of
the loadings of
m assemblages)*

Assemblage description matrix
*(m assemblages described in
terms of n species)*

each sample is
composed of linear
comb. of each factor

each species
contributes a certain
amount to each factor

Species # 1 2 3...n		*Factor #* 1 2 3…m		*Species #* 1 2 3...n
1 2 3 · · · N [N x n]	=	1 2 3 · · · N [N x m]		1 2 3 · · m [m x n]

N samples, n species *N samples, m factors* *m factors, n species assemblages*

i.e.

$$U_{57,3} = B_{57,1}F_{1,3} + B_{57,2}F_{2,3} + B_{57,3}F_{3,3} + \ldots\ldots$$

% species 3
in sample 57

loading of factor
2 for sample 57

contribution of
species 3 to factor 3

where
   N is of the order of 50-1000.
   n is of the order of 30-100.
   m is of the order of 5-10.

   This equation describes each of the very large number of samples for which a large number of species has been counted as the linear combination of a small number of a small number of assemblages.

Then, to interpret paleo-data, it is assumed that the assemblages remain stable, so that we can then interpret downcore data as linear combinations of those same assemblages:

$$\mathbf{B} = \mathbf{U} \ \mathbf{F^T}$$

transpose of F
determined from
core-top study

	Factor			*Species*			*Factor*
	1 2 3..m			1 2 3…n			1 2 3…m
1			1			1	
2			2			2	
sample# 3	N x m	=	*sample#* 3	N x n	*species#* 3		n x m
:			:		:		
N			N		n		

SAMPLE
DESCRIPTION

DATA

FACTOR
DESCRIPTION

i.e.:

$$B_{57,4} = U_{57,1}F^T_{1,4} + U_{57,2}F^T_{2,4} + U_{57,3}F^T_{3,4} + ...$$

loading of factor
4 on sample 57

% species 1
in sample 57

contribution of species 2
to factor 4

e.g. (Kipp, 1976, p. 23) factor description:

	Trop	Sub-trop	Transit.	Subpolar	Polar	Gyre Margin
N. pachyderma (left)	-0.018	0.016	-0.029	-0.027	0.987	0.011
G. ruber (white)	0.922	0.112	-0.013	-0.054	0.030	-0.098

Imbrie and Kipp (1971) foraminiferal factors vs. temperature
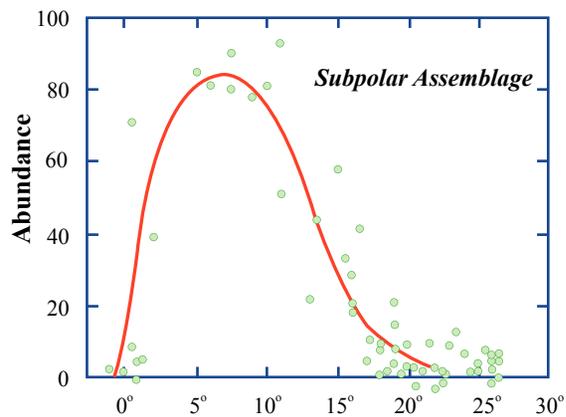
Abundance of the subpolar assemblage versus winter surface temperature for 61 core top samples. Curve fitted by eye.
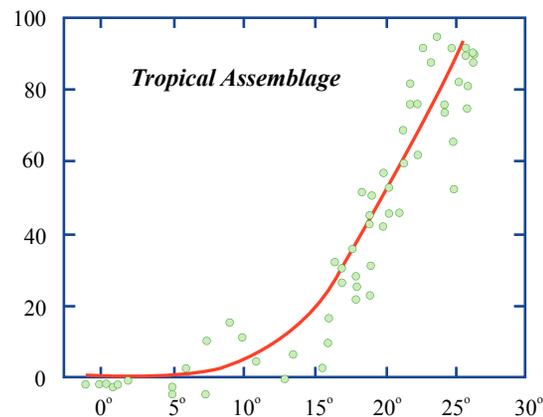
Abundance of the tropical assemblage versus winter surface temperature for 61 core top samples. Curve fitted by eye.
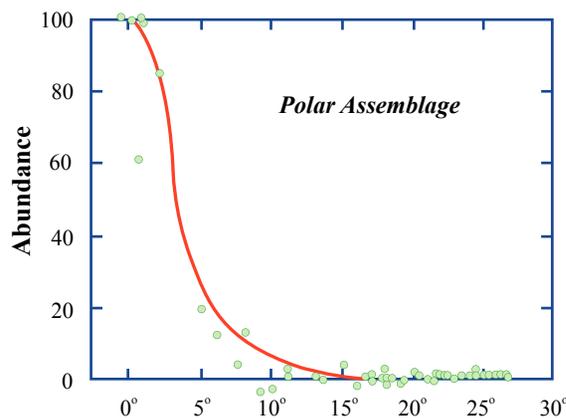
Abundance of the polar assemblage versus winter surface temperature for 61 core top samples. Curve fitted by eye.
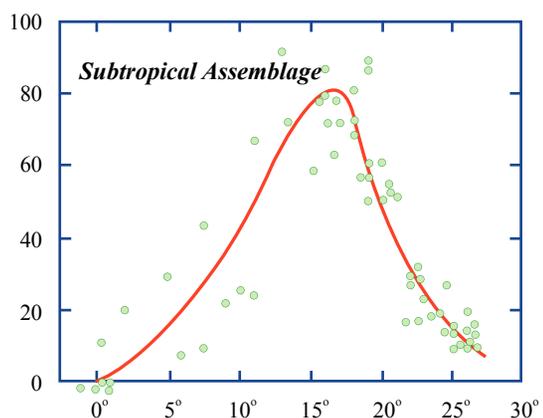
Abundance of the subtropical assemblage versus winter surface temperature for 61 core top samples. Curve fitted by eye.

Figure by MIT OpenCourseWare.

Imbrie and Kipp (1971) factors vs. T, regression fit
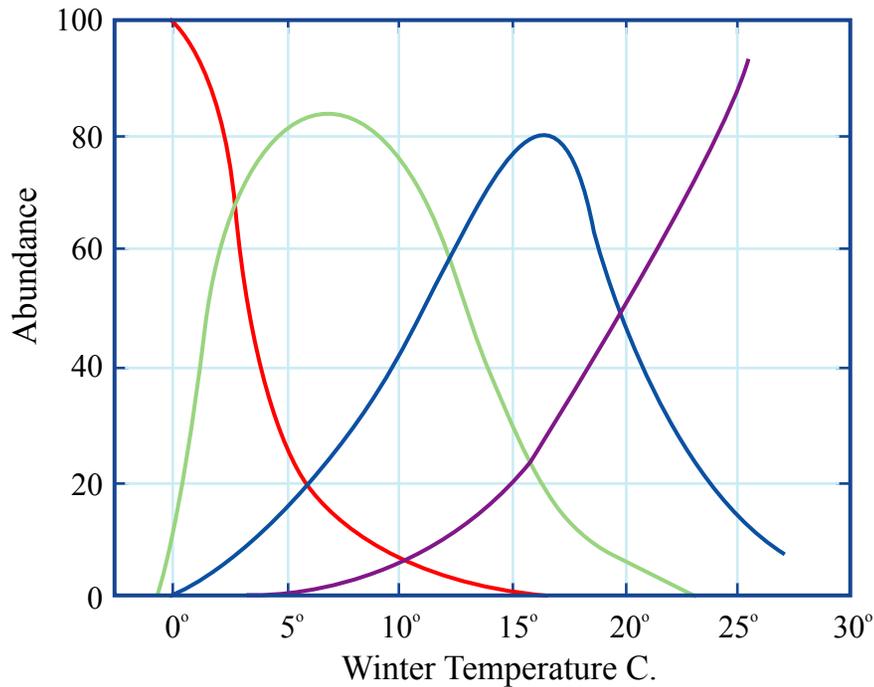
General abundance trends for four of the varimax assemblages related to winter surface temperature.

Legend:
- Polar
- SubPolar
- Tropical
- SubTropical
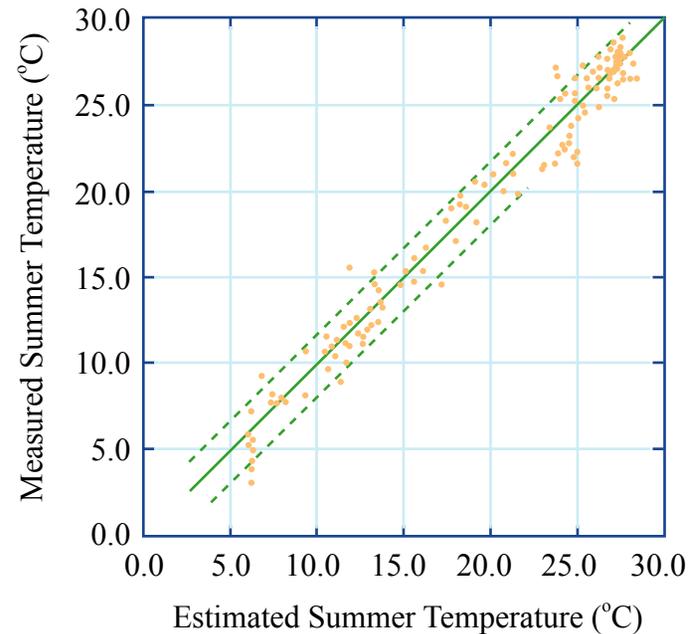
Observed values of summer sea-surface temperature versus estimates calculated by transfer function F13 for 191 North Atlantic sea-bed samples. Scatter diagram with 80% confidence intervals indicated by dashed lines.

Algebraically, there are an infinite number of ways to decompose the core top data so that it satisfies these equations. Varimax optimization is an objective method for finding one of these solutions, given a choice of the number of factors to be used. The number of factors is arbitrary, however!

Geologists' approach to picking number of assemblages is to map them to find out how many factors are reasonable (i.e., factors that are 'unmappable' are considered to be in the noise).

1. Paleo-ecological assumption:
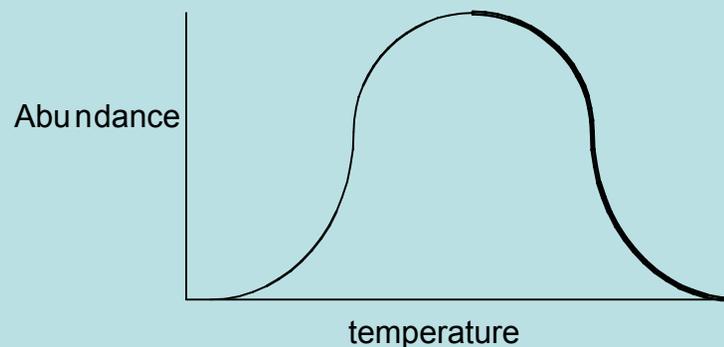
$$i.e. \quad T_i = f\,[B_{CT}(i,j)]$$

Imbrie-Kipp used least squares matrix techniques to find an equation that fit T vs. factor loadings and their cross-products:

$$\mathbf{A} \quad \mathbf{x} \quad = \quad \mathbf{b}$$

column vector
$m \times 1$

column vector
$n \times 1$

matrix:
$m \times n$

m rows n columns
(#samples)   (#factors and cross-products used)

factor function #
1 2 3...n

$$
\text{sample\# }
\begin{matrix} 1 \\ 2 \\ 3 \\ \vdots \\ \vdots \\ \vdots \\ m \end{matrix}
\begin{bmatrix} & & \\ & \mathbf{A} & \\ & & \end{bmatrix}
\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ \mathbf{x} \\ \vdots \\ a_n \end{bmatrix}
=
\begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ \mathbf{b} \\ \vdots \\ T_m \end{bmatrix}
$$

m equations
n unknowns

(factor loadings   reg'n   Core-top site
and cross-products)   coefficients   surface T_s or T_w

2. Least squares criterion: find x_i such that

$\quad\quad \Sigma\ (A_{ij}x_i - T_i)^2$ is minimized

3. SOLUTION:

$\quad\quad$ for eq'n $Ax = b$ (m eq'ns, n unknowns),
$\quad\quad$ if columns of A are linearly independent, then

$$\mathbf{x = (A^T A)^{-1} A^T b}$$

$\quad\quad$ (paleo-ecological solution)

$\quad$ THEN compute paleo T: $\mathbf{b = A\ x}$

How well does this work? Discussion of Imbrie and Kipp (1971) and Kipp (1976) papers: bio-ecological deductions

- 1. Abstract - discuss
- 2. Rejection of samples - objectivity
- 3. "Summer" "Winter" definitions (cold/warm season). What do summer/winter T maps mean? Is a "seasonality" really available in this data?
- 4. Effect of differential dissolution
- 5. Effect of transport by currents, etc.
- 6. What is the "Gyre margin" assemblage?
- 7. *G. sacculifer*, *N. dutertrei*, *G. menardii*: Gulf Stream indicators?
- 8. Comparison with plankton- tow observations
- 9. Problem of very cold temperatures: below about 8°C, foraminiferal assemblages are monospecific.
- 10. Species maps - how good is the T-correlation?
- 11. Factor distributions
- 12. Differences map

Initial results from the Imbrie-Kipp technique: CLIMAP project

A. Global Ocean Temperature Reconstructions during the last Glacial Maximum

      1. Largest change is in the temperatures of the North Atlantic Ocean polewards of 40°N: >10°C cooling over a large area.

      2. Tropical sea surface temperatures are surprisingly stable: change is less than 2°C over large areas of the tropics. This result fits in well with the Shackleton/Dansgaard revision of the interpretation of foraminiferal $\delta^{18}O$.

      3. Other than the North Atlantic, polar waters expanded about 5° (latitude) equatorward.

      4. Some evidence of cooler temperatures in wind-driven upwelling environments (especially off Northwest Africa, but perhaps also on other eastern boundaries and on the equator.

B. Downcore records of paleotemperature variability during the last 150,000 years. Northern North Atlantic temperature variability: looks similar to $\delta^{18}O$ variability.

Critique of the Imbrie-Kipp Method 1

A. The Imbrie-Kipp method (and all other methods that require completely empirical calibration) explicitly assumes that the temperature responses of the assemblages remains constant through time. But what happens during genetic evolutionary change? Can species evolve their temperature tolerance in the face of environment pressure?

     1. This kind of problem is most serious in continental environments where species may not be very mobile; e.g., even if the environment changes, in the short term there may not be any "seed corn" to allow the optimum assemblage to develop, so a non-optimum species has some time in which to respond to evolutionary pressures before more efficient competitors arrive.

     2. In marine environments, this problem is mitigated by the great mobility of the species; it is generally very easy for species to migrate (but remember that in the case of foraminifera, "migration" occurs via ocean currents and eddies) as the temperature bands move.

     a. This is not always true however. How can cold north polar and south polar species intermix? (Perhaps by subsurface exchange, although this may be difficult.)

     b. One counter-example of this rule is the disappearance of pink-pigmented G. ruber 120,000 years ago in the Pacific and Indian Oceans. This species existed in all of the oceans before that time, and is still quite abundant in the Atlantic Ocean. There is nothing obvious about the environmental differences between the oceans that could account for the non-occurrence of this species, so it appears that simply there is no "seed corn" for this species in the Pacific and Indian Oceans. Migration of this species from the Atlantic is prevented by the hostile environment (e.g. cold) of the circumpolar connection between the oceans. This might even be an example of a time-transgressive extinction in the making....

     3. Even in marine environments, this problem is very serious for very old environments because of the extinction of species; samples that are (say) about 50 million years old have none of the same species that exist today. How can you calibrate a species that doesn't exist in the modern ocean?

     a. For foraminifera, this problem becomes significant at about 1-3 million years ago.

Critique of the Imbrie-Kipp Method 2

B. The Imbrie method has some significant advantages: it is objective and the orthogonality of the factors allows one (at least in principle, if not always in practice) to "throw away" information that is not relevant to reconstructions of the particular property that we are interested in (e.g., if some factors respond to salinity or some other variable rather than temperature, we can leave them out of the paleotemperature regression (e.g., the "gyre margin" assemblage). But it also has some disadvantages: the regression equations have completely arbitrary form: they have no basis in theory. Since the equations are chosen to fit the data, they are OK for samples that lie within the range of the calibration samples. But for samples that lie outside of the calibration range, the equations are unconstrained and can blow up: the "no-analogue" problem. One of the solutions to this problem is to avoid using the technique for samples that fall outside the calibration range of factor compositions. Unfortunately, this includes some interesting areas, and it also has a somewhat arbitrary nature: how does one assess exactly when the no-analogue problem arises?

C. The method also has some problems with dissolution of calcium carbonate on the seafloor: some species of foraminifera are more sensitive to dissolution than others, so dissolution can alter the species composition and hence the factor composition. One way to miminize the problem is to include "dissolved" samples in the set of calibration samples. This was done by Kipp (1976) in her Atlantic transfer function. It does not work as well in the Pacific however, where the extent of dissolution is greater; much of the micropaleontological information has been lost.

Critique of the Imbrie-Kipp Method 3

D. A more troublesome aspect of the regression procedure is the arbitrary nature of the calibration data (e.g., sea surface temperature, SST). One can also do a calibration to "temperature at 100m depth" or "temperature at 200m depth"; these work almost as well (as judged by the standard deviation of the fits), mainly because SST is highly correlated with subsurface temperatures. Also, since temperature has a seasonal cycle, one can get equations for "winter temperature", "summer temperature", or "annual average temperature"; again, these all work just about as well. One could also regress against "sea-surface salinity" (SSS) or "productivity"; these don't work as well, but they do give calibrations that are better than random. But SST is highly correlated with SSS; what is it that foraminifera are really measuring?

1. I recommend that you look at these regressions in this way: in effect we are trying to put a number that everybody can understand (e.g. SST) onto data that only micropaleontologists or mathematicians understand (e.g. species composition of a fossil sample, or factor loadings). The idea is valid only to the extent that the variable you are assigning to the data is in reality a master variable controlling the formation of a fossil assemblage. There is good reason (i.e., maps of global foraminiferal distributions) to believe that foraminiferal species composition is highly correlated with SST (and subsurface temperatures which must be highly correlated with SST). So it is reasonable to derive temperatures from foram species data. It is not so clear whether we should make the correlation to winter SST, summer SST, or any closely-correlated variant of this theme. Since paleoenvironmental models (e.g. General Circulation Models (GCM) of the atmosphere) need to have seasonal SST information to operate, it is reasonable to make the regressions for each season so as to provide the necessary information to make the model operate properly.

2. It is probably not reasonable to take the winter/summer estimates seriously. Relenting a littlefrom this skeptical attitude, we should acknowledge that there are many species and several factor analysis assemblages in that sample, and these may contain information on the separate seasons. But it is not clear that a simple assemblage regression against winter or summer SST is a valid way to get at this information.

3. It is certainly not reasonable to take the paleosalinity estimates seriously. SST and SSS are too highly correlated, so if T is the dominant variable, then the regressions simply tend to give us the modern T-S correlation. The mean salinity of the ocean was 1‰ higher during the last ice ice; did this have any effect on the species composition of warm tropical waters? However, it is not entirely unreasonable that salinity variations play some role in the growth of foraminifera. For example, few foraminifera grow in the most saline waters of the Red Sea (S>40‰). Some species of foraminifera probably can't tolerate very low salinities. So there is probably some paleosalinity information when near the tolerance limits of foraminifera. How do we use this information however, and what is its margin of error?

An Alternative to the Imbrie-Kipp method:
Modern Analogue Technique (MAT)

As suggested previously, in trying to assign a paleotemperature to a fossil assemblage from a sediment core, in effect we are trying to find a modern sample of approximately the same species composition and saying that the temperature above that modern site is the paleotemperature of the ancient site at the time the fossil assemblage formed. MAT proposes to do this comparison explicitly; e.g. it takes the species composition of the fossil sample, and assigns the it the temperature associated with the most similar modern samples.

1. There are many ways this might be done - how do we assess the level of similarity between two samples? The form of MAT adopted for oceanic samples uses a simple and easy-to-understand method. Consider the n-dimensional space represented by the percentage of each species as an independent dimension. Samples are considered the most similar when the the distance between them is the smallest. So the paleotemperature estimate adopted is the average of the temperatures of the core-top samples closest to the paleo-sample in species space.

2. One of the advantages of this approach is that it provides a quantitative measure of the degree to which you are able to find a suitable analogue. You are always able to find the most similar modern samples, and hence derive a paleotemperature estimate, but you also get a dissimilarity coefficient that can tip you off when the modern and ancient samples are really not all that similar.

3. Work by Bob Thunell in western tropical Pacific suggests that MAT works better than Imbrie-Kipp in samples that are relatively heavily dissolved. MAT simply finds a modern heavily-dissolved analogue sample, whereas Imbrie-Kipp struggles to keep the species assemblage groupings constant. similar.

4. The LGM temperatures of the Atlantic Ocean recently were mapped out using a MAT variant by Pflaumann et al. (2003). Note: discuss distance-weighting of SST estimate.

A paleoclimate conundrum

Imbrie-Kipp and planktonic oxygen isotope evidence suggest that tropical sea-surface temperatures were not changed much during the last glacial maximum. But high-elevation evidence from continental environments and islands suggest that the tropics cooled considerably: snow lines on tropical mountains were lower, and the vegetation in the tropical mountains indicate a cooler climate. These lines of evidence (and others which have appeared more recently) appear to suggest tropical sea surface cooling where CLIMAP suggests little change. What's the problem? similar.

      A. One possible "clue" to this problem was provided by Prell et al. (1976), who did separate factor analyses on core tops and LGM samples and found that the LGM samples yielded one more factor than the core top samples. They dubbed this factor the "glacial cool equatorial" factor.

      B. Much later, after being forced to reassess the problem of tropical sea surface temperatures, Hostetler and Mix (1999) performed factor analyses on downcore samples from the tropics, on the assumption that the modern core top miss out on a significant factor. Using the factors derived from this downcore factor analysis, they devised a paleotemperature transfer function by regressing modern SST vs. these factors. Then they used this new transfer function to estimate tropical LGM SSTs – and they found much cooler LGM SST estimates.

Prell et al. (1976) : Factor analyzing tropical glacial age samples turns up a factor that doesn't appear in the core top samples:

Image removed due to copyright restrictions.
Please see: Prell, et al. (1976) *GSA Mem* 145.

CLIMAP redone after downcore factors added

Image removed due to copyright restrictions.
Please see: Hostetler and Mix, *Nature,* 17 June 1999, Vol 399, Page 674, Figure 1.

No-analogue factors?

- The SST of the LGM ocean is being remapped by the EPILOG (Mix et al., 2001) & GLAMAP (Pflaumann et al. 2003) programs based on this knowledge and other types of paleo-temperature tracers.

- Not everyone is entirely happy with this resolution of the problem; it's a sort of "closing the barn after the horse has left" solution (or perhaps a "tell us what the answer is and we will figure out a way to get the forams to get that answer" solution!

- We will return to this problem later.

Readings

- Read the * items to glean the essence of what they are trying to accomplish - they should not be considered the final word >Y2K!

- *1. Imbrie, J. and N.G. Kipp (1971) A new micropaleontological method for quantitative paleoclimatology, in Late Cenozoic Glacial Ages (ed. K.K. Turekian), Yale Univ. Press, New Haven, pp. 71-182.

- *2. 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, in Investigation of Late Ouaternary Paleoceanography and Paleoclimatology (eds. R.M. Cline and J.D. Hays), Geol. Soc. Am. Bull. Memoir 145, pp. 3-42.

- 3. Overpeck, J.T., T. Webb II, and I.C. Prentice (1985) Quantitative interpretation of fossil pollen spectra: dissimilarity coefficients and the method of modern analogs, Quat Res. 12: 47-82.

- 4. Anderson D. M., Prell W. L. ,and Barratt N. J. (1989) Estimates of sea surface temperature in the Coral Sea at the last glacial maximum. Paleoceanogr. 4, 615-627.

- 5. Joreskog, Klovan, and Reyment (1976) Geological Factor Analysis, Elsevier, Amsterdam, 1976, 178 p.

- 6. Hostetler, S. W. and A. C. Mix (1999). "Reassessment of ice-age cooling of the tropical ocean and atmosphere." Nature 399: 673-676.

- 7. Lidz L. (1966) Deep-sea Pleistocene Biostratigraphy. Science. 154, 1448-1452.

- 8. Mix, A. C., E. Bard, et al. (2001). "Environmental Processes of the Ice Age: Land, Oceans, Glaciers (EPILOG)." Quat. Sci. Rev. 20: 1-34.

- 9. Mix, A. C., E. Bard, et al. (2001). "Environmental Processes of the Ice Age: Land, Oceans, Glaciers (EPILOG)." Quat. Sci. Rev. 20: 620-657.

- 10. Pflaumann, U., M. Sarnthein, et al. (2003). "Glacial North Atlantic: Sea-surface conditions reconstructed by GLAMAP 2000." Paleoceanogr. 18: 1065, doi:10.1029/2002PA000774.

- 11. Prell, W. L., J. V. Gardner, et al. (1976). Equatorial Atlantic and Caribbean foraminiferal assemblages, temperatures, and circulation: Interglacial and glacial comparisons. Investigation of Late Quaternary Paleoceanography and Paleoclimatology Geol. Soc. Am Memoir. R. M. Cline and J. D. Hays. Boulder, CO USA, Geo. Soc. Am. 145.