

Pattern Analyses (EOF Analysis)

- **Introduction**
- **Definition of EOFs**
- **Estimation of EOFs**
- **Inference**
- **Rotated EOFs**

2.2 Pattern Analyses

Introduction: What is it about?

- Pattern analyses are techniques used to identify patterns of the simultaneous temporal variations
- Given a m -dimensional time series \vec{x}_t , the anomalies \vec{x}'_t defined as the deviations from the sample mean can be expanded into a finite series

$$\vec{x}'_t = \sum_{i=1}^k \hat{\alpha}_{i,t} \hat{p}^i$$

with time coefficients $\hat{\alpha}_{i,t}$ and fixed patterns \hat{p}^i . Equality is usually only possible when $k=m$

- The patterns are specified using different minimizations

EOFs: \vec{x}'_t is optimally described by $\sum_{i=1}^k \hat{\alpha}_{i,t} \hat{p}^i \Rightarrow \sum \left(\vec{x}'_t - \sum_{i=1}^k \hat{\alpha}_{i,t} \hat{p}^i \right)^2 = \min!$

POPs: \vec{x}'_t is optimally described by $\mathbf{A}\vec{x}'_t \Rightarrow \sum (\vec{x}'_t - \mathbf{A}\vec{x}'_{t-1})^2 = \min!$

- The patterns can be orthogonal

2.2 Pattern Analyses

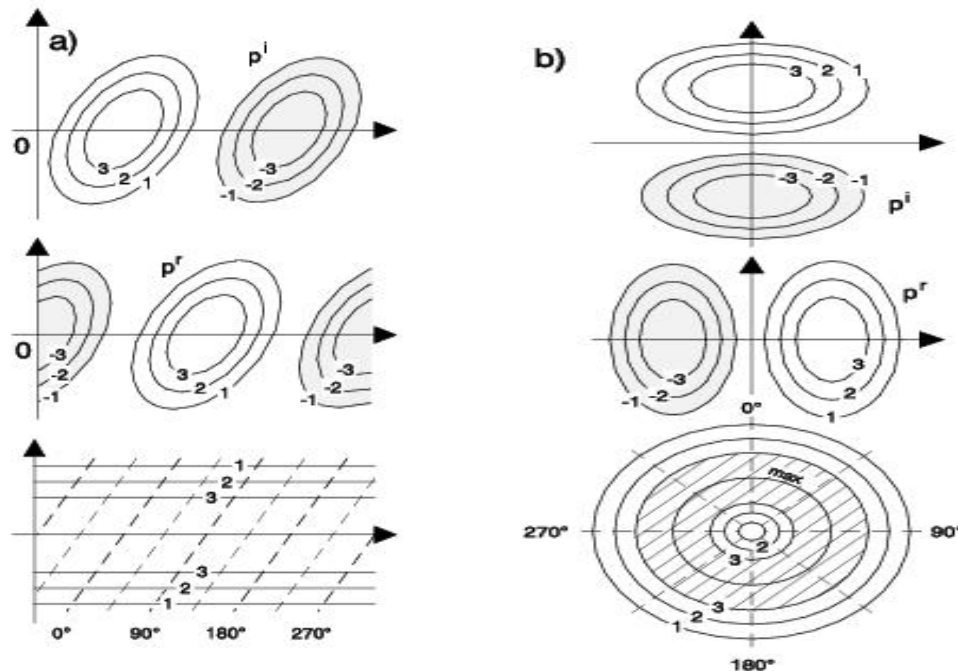
Introduction: What can patterns and their coefficients describe?

- Standing Signals

A fixed spatial structure whose strength varies with time

- Propagating Signals

A structure propagating in space. It has to be described by two patterns such that the coefficient of one pattern lags (or leads) the coefficient of the other one by a fixed time lag (often 90°)



Schematic representation of a linearly propagating (left) and clockwise rotating (right) wave using two patterns: p^i and p^r . If the initial state of the wave is p^i , then its state a quarter of period later will be p^r .

2.2 Pattern Analyses

Example: Daily Profile of Geopotential Height over Berlin

Data: 20-year data set containing 120 winter days times 9 vertical levels between 950 and 300 hPa, i.e. $20 \times 120 \times 9 = 21600$ observations

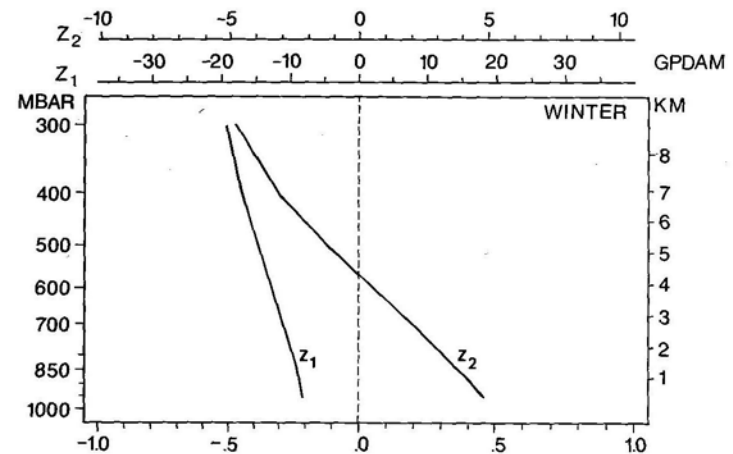
How should we describe the spatial variability?

One way is to compute the variance at each level. This however does not tell us how the variations are correlated in the vertical

Solution: describing spatial correlations using a few EOFs

Usefulness:

- To identify a small subspace that contains most of the dynamics of the observed system
- To identify modes of variability



The first two EOFs, labeled z_1 and z_2 , of the daily geopotential height over Berlin in winter. The first EOF represents 91.2% and the second 8.2% of the variance. They may be identified with the equivalent barotropic mode and the first baroclinic mode of the tropospheric circulation.

2.2 Pattern Analyses

Introduction: Elements of Linear Analysis

Eigenvalues and eigenvectors of a real square matrix

Let \mathbf{A} be an $m \times m$ matrix. A real or complex number λ is said to be an eigenvalue of \mathbf{A} , if there is a nonzero m -dimensional vector \vec{e} such that

$$\mathbf{A}\vec{e} = \lambda\vec{e}$$

Vector \vec{e} is said to be an eigenvector of \mathbf{A}

- Eigenvectors are not uniquely determined
- A real matrix \mathbf{A} can have complex eigenvalues. The corresponding eigenvectors are also complex. The complex eigenvalues and eigenvectors occur in complex conjugate pairs

Hermitian matrices

A square matrix \mathbf{A} is Hermitian if

$$\mathbf{A}^{T_c} = \mathbf{A}$$

where \mathbf{A}^{T_c} is the conjugate transpose of \mathbf{A} . Hermitian matrices have real eigenvalues only. Real Hermitian matrices are symmetric. Eigenvalues of a symmetric matrix are non-negative and eigenvectors are orthogonal

2.2 Pattern Analyses

Introduction: Elements of Linear Analysis

Bases

A collection of vectors $\{\vec{e}^1, \dots, \vec{e}^m\}$ is said to be a linear basis for an m -dimensional vector space \mathbf{V} if for any vector $\vec{a} \in \mathbf{V}$ there exist coefficients α_i , $i=1, \dots, m$, such that

$$\vec{a} = \sum_i \alpha_i \vec{e}^i$$

The basis is orthogonal, when $\langle \vec{e}^i, \vec{e}^j \rangle = 0$ if $i \neq j$

or orthonormal when $\langle \vec{e}^i, \vec{e}^j \rangle = 0$ if $i \neq j$ and $\|\vec{e}^i\| = 1$ for all $i = 1, \dots, m$

where $\langle \cdot, \cdot \rangle$ denotes the inner product which defines a vector norm $\|\cdot\|$. One has

$$\langle \vec{x}, \vec{y} \rangle = \vec{x}^T \vec{y} \text{ and } \|\vec{x}\|^2 = \langle \vec{x}, \vec{x} \rangle$$

Transformations

If $\{\vec{e}^1, \dots, \vec{e}^m\}$ is a linear basis and $\vec{y} = \sum_i \alpha_i \vec{e}^i$, then

$$\alpha_i = \langle \vec{y}, \vec{e}_a^i \rangle$$

where \vec{e}_a^i is the adjoint of \vec{e}^i satisfying $\langle \vec{e}^i, \vec{e}_a^j \rangle = 0$ for $i \neq j$ and $\langle \vec{e}^i, \vec{e}_a^i \rangle = 1$ for $i = j$

2.2 Pattern Analyses

Definition of Empirical Orthogonal Functions: The First EOF

EOFs are defined as parameters of the distribution of an m -dimensional random vector \vec{X} . The first EOF \vec{e}^1 is the most powerful single pattern representing the variance of \vec{X} defined as the sum of variances of the elements of \vec{X} . It is obtained by minimizing, subjected to $\|\vec{e}^1\|^2 = 1$,

$$\varepsilon_1 = E\left(\|\vec{X} - \langle \vec{X}, \vec{e}^1 \rangle \vec{e}^1\|^2\right) = \text{Var}(\vec{X}) - \text{Var}(\langle \vec{X}, \vec{e}^1 \rangle)$$

which results in

$$\Sigma \vec{e}^1 - \lambda \vec{e}^1 = 0$$

where λ is the Lagrange multiplier associated with the constraint $\|\vec{e}^1\|^2 = 1$

Note:

$$\begin{aligned} \text{Var}(\langle \vec{X}, \vec{e}^1 \rangle) &= E\left(\left(\vec{X}^T \vec{e}^1\right)^T \vec{X}^T \vec{e}^1\right) \\ &= \vec{e}^{1T} \Sigma \vec{e}^1 = \vec{e}^{1T} \lambda \vec{e}^1 = \lambda \end{aligned}$$

\vec{e}^1 is an eigenvector of covariance matrix Σ with a corresponding eigenvalue λ !

Minimizing ε_1 is equivalent to maximizing the variance of \vec{X} contained in the 1-dimensional subspace spanned by \vec{e}^1 , $\text{Var}(\langle \vec{X}, \vec{e}^1 \rangle)$.

ε_1 is minimized when \vec{e}^1 is an eigenvector of Σ associated with its largest eigenvalue λ

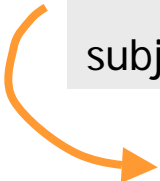
2.2 Pattern Analyses

More EOFs

Having found the first EOF, the second is obtained by minimizing

$$\varepsilon_1 = E\left(\left\|\vec{X} - \langle \vec{X}, \vec{e}^1 \rangle \vec{e}^1 - \langle \vec{X}, \vec{e}^2 \rangle \vec{e}^2\right\|^2\right)$$

subjected to the constraint $\|\vec{e}^2\|^2 = 1$



\vec{e}^2 is an eigenvector of covariance matrix Σ that corresponds to its second largest eigenvalue λ_2 . \vec{e}^2 is orthogonal to \vec{e}^1 because the eigenvectors of a Hermitian matrix are orthogonal to each other

EOF Coefficients or Principle Components

The EOF coefficients are given by

$$\alpha_i = \langle \vec{X}, \vec{e}^i \rangle = \vec{X}^T \vec{e}^i = \vec{e}^{iT} \vec{X}$$

2.2 Pattern Analyses

Theorem

Let \vec{X} be an m -dimensional real random vector with mean $\vec{\mu}$ and covariance matrix Σ . Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ be the eigenvalues of Σ and let $\vec{e}^1, \dots, \vec{e}^m$ be the corresponding eigenvectors of unit length. Since Σ is symmetric, the eigenvalues are non-negative and the eigenvectors are orthogonal.

- The k eigenvectors that correspond to $\lambda_1, \dots, \lambda_k$ minimize

$$\varepsilon_k = E\left(\left\|(\vec{X} - \mu) - \sum_{i=1}^k \langle \vec{X} - \mu, \vec{e}^i \rangle \vec{e}^i\right\|^2\right)$$

- $\varepsilon_k = \text{Var}(\vec{X}) - \sum_{i=1}^k \lambda_i$

- $\text{Var}(\vec{X}) = \sum_{i=1}^m \lambda_i$

use of any other k -dimensional subspace will lead to mean squared errors at least as large as ε_k

gives the mean squared error incurred when approximating \vec{X} in a k -dimensional subspace

broken up the total variance into m components

2.2 Pattern Analyses

Interpretation

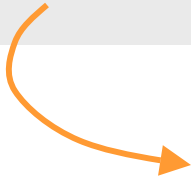
- The bulk of the variance of \bar{x} can often be represented by a first few EOFs
- The physical interpretation is limited by the fundamental constraint that EOFs are orthogonal. Real world processes do not need to be described by orthogonal patterns or uncorrelated indices

2.2 Pattern Analyses

Properties of the EOF Coefficients

The covariances of EOF coefficients α_i are given by

$$\begin{aligned} \text{Cov}(\alpha_i, \alpha_j) &= E(\langle \bar{X}, \bar{e}^i \rangle \langle \bar{X}, \bar{e}^j \rangle) \\ &= \bar{e}^{iT} E(\bar{X}\bar{X}^T) \bar{e}^j \\ &= \bar{e}^{iT} \Sigma \bar{e}^j = \bar{e}^{iT} \lambda_j \bar{e}^j \\ &= \begin{cases} 0, & i \neq j \\ \lambda_j, & i = j \end{cases} \end{aligned}$$



The EOF coefficients are uncorrelated

2.2 Pattern Analyses

Vector Notation

The random vector \vec{X} can be written as

$$\vec{X} = \mathbf{P}\vec{\alpha}, \quad \text{or} \quad \vec{\alpha} = \mathbf{P}^T \vec{X}$$

with $\mathbf{P} = (\vec{e}^1 | \vec{e}^2 | \dots | \vec{e}^m)$, $\vec{\alpha} = (\alpha_1, \dots, \alpha_m)^T$, which leads to

$$\begin{aligned} \Sigma &= E(\vec{X}\vec{X}^T) = \mathbf{P}E(\vec{\alpha}\vec{\alpha}^T)\mathbf{P}^T \\ &= \mathbf{P}\Lambda\mathbf{P}^T \end{aligned}$$

where Λ is the diagonal $m \times m$ matrix composed of the eigenvalues of Σ .

2.2 Pattern Analyses

Degeneracy

It can be shown that the eigenvalues are the m roots of the m -th degree polynomial

$$p(\lambda) = \det(\Sigma - \lambda \mathbf{I})$$

where \mathbf{I} is the $m \times m$ identity matrix.

- If λ_o is a root of multiplicity 1 and \vec{e} is the corresponding eigenvector, then \vec{e} is unique up to sign
- If λ_o is a root of multiplicity k , the solution space

$$\Sigma \vec{e} = \lambda_o \vec{e}$$

is uniquely determined in the sense that it is orthogonal to the space spanned by the $m-k$ eigenvectors of Σ with $\lambda_i \neq \lambda_o$. But any orthogonal basis for the solution space can be used as EOFs. In this case the EOFs are said to be degenerated.

Bad: patterns which may represent independent processes cannot be disentangled

Good: for $k=2$ the pair of EOFs and their coefficients could represent a propagating signal. As the two patterns representing a propagating signal are not uniquely determined, degeneracy is a necessary condition for the description of such signals

2.2 Pattern Analyses

Coordinate Transformations

Consider two m -dimensional random vectors \vec{x} and \vec{z} related through $\vec{z} = \mathbf{L}\vec{x}$ where \mathbf{L} is an invertible matrix. If the transformation is orthogonal (i.e. $\mathbf{L}^{-1} = \mathbf{L}^T$), the eigenvalue of the covariance matrix of \vec{x} , Σ_{xx} , is also the eigenvalue of the covariance matrix of \vec{z} , Σ_{zz} , and the EOFs of \vec{x} , \vec{e}^x , are related to those of \vec{z} , \vec{e}^z , via

$$\vec{e}^z = \mathbf{L}\vec{e}^x$$

Proof:

Since $\Sigma_{zz} = \mathbf{L}\Sigma_{xx}\mathbf{L}^T$, $\Sigma_{xx}\vec{e}^x = \lambda\vec{e}^x$

$$\Sigma_{zz}\mathbf{L}\vec{e}^x = \mathbf{L}\Sigma_{xx}\mathbf{L}^T\mathbf{L}\vec{e}^x = \mathbf{L}\Sigma_{xx}\vec{e}^x = \lambda\mathbf{L}\vec{e}^x$$

Consequence of using an orthogonal transformation:

The EOF coefficients are invariant, since

$$\vec{\alpha}_x = \mathbf{P}_x^T \vec{X} = \mathbf{P}_x^T \mathbf{L}^T \vec{Z} = (\mathbf{L}\mathbf{P}_x)^T \vec{Z} = \mathbf{P}_z^T \vec{Z} = \vec{\alpha}_z$$

2.2 Pattern Analyses

Estimation of Empirical Orthogonal Functions

Approach I

Estimate the covariance matrix and use the eigenvectors and the eigenvalues of the estimated covariance matrix as estimators of the EOFs and the corresponding eigenvalues

Approach II

Use a set of orthogonal vectors that represent as much as the sample variance as possible as estimators of EOFs

The two approaches are equivalent and lead to the following theorem

2.2 Pattern Analyses

Theorem

Let $\hat{\Sigma}$ be the estimated covariance matrix derived from a sample $\{\bar{x}_1, \dots, \bar{x}_n\}$ representing n realization of \bar{X} . Let $\hat{\lambda}_1, \dots, \hat{\lambda}_m$ be the eigenvalues of $\hat{\Sigma}$ and $\hat{e}_1, \dots, \hat{e}_m$ the corresponding eigenvectors of unit length. Since $\hat{\Sigma}$ is symmetric, the eigenvalues are non-negative and the eigenvectors are orthogonal

- The k eigenvectors corresponding to $\hat{\lambda}_1, \dots, \hat{\lambda}_m$ minimize

$$\hat{\varepsilon}_k = \sum_{j=1}^n \left| \bar{x}_j - \sum_{i=1}^k \langle \bar{x}_j, \hat{e}^i \rangle \hat{e}^i \right|^2$$

- $\hat{\varepsilon}_k = \text{Var}(\bar{X}) - \sum_{i=1}^k \hat{\lambda}_i$
- $\text{Var}(\bar{X}) = \sum_{i=1}^m \hat{\lambda}_i$

The EOF estimates represent the sample variance in the same way as the EOFs do with the random variable

2.2 Pattern Analyses

Properties of the Coefficients of the Estimated EOFs

- As with the true EOFs, the estimated EOFs span the full m -dimensional vector space. The random vector \vec{X} can be written as

$$\vec{X} = \sum_{j=1}^m \hat{\alpha}_j \vec{e}^j \quad \text{with } \hat{\alpha}_j = \langle \vec{X}, \hat{e}^j \rangle$$

- When \vec{X} is multivariate normal, the distribution of the m -dimensional vector of EOF coefficients, conditional upon the sample used, is multivariate normal with mean and covariance matrix

$$E(\hat{\alpha} | \vec{x}_1, \dots, \vec{x}_m) = \hat{\mathbf{P}}^T \vec{\mu}, \quad \text{Cov}(\hat{\alpha}, \hat{\alpha} | \vec{x}_1, \dots, \vec{x}_m) = \hat{\mathbf{P}}^T \Sigma \hat{\mathbf{P}}$$

where $\hat{\mathbf{P}}$ has \hat{e}^j in j -th column

- The variance of the EOF coefficients computed from the sample is

$$\frac{1}{n} \sum_{i=1}^n |\hat{\alpha}_{ji} - \bar{\alpha}_j|^2 = \hat{\lambda}_j$$

- The sample covariance of a pair of EOF coefficients computed from the sample is zero

Two interpretations of $\hat{\lambda}_j$

- as an estimate of the variance of the true α
- as an estimate of the variance of $\hat{\alpha}_j$

2.2 Pattern Analyses

The Variance of EOF Coefficients of a Given Set of Estimated EOFs

Given a set of eigenvalues and EOFs derived from a finite sample, any random vector \vec{X} can be represented in the space spanned by these estimated EOFs using the transformation

$$\vec{X} = \hat{\mathbf{P}} \hat{\alpha}, \quad \hat{\alpha} = \hat{\mathbf{P}}^T \vec{X}$$

Question: is the variance of the transformed random variables $\hat{\alpha}_i = \langle \vec{X}, \hat{e}^i \rangle$ equal the true EOF coefficient (i.e. is the eigenvalue of the estimated covariance matrix equal to the true eigenvalue)?

The answer is no

Since the first EOF minimizes

$$\varepsilon_1 = E\left(\left\|\vec{X} - \langle \vec{X}, \vec{e} \rangle \vec{e}\right\|^2\right)$$

one has

$$Var(\vec{X}) - Var(\alpha_1) = E\left(\left\|\vec{X} - \langle \vec{X}, \vec{e}_1 \rangle \vec{e}_1\right\|^2\right)$$

$$< E\left(\left\|\vec{X} - \langle \vec{X}, \hat{e}_1 \rangle \hat{e}_1\right\|^2\right) = Var(\vec{X}) - Var(\hat{\alpha}_1)$$



- $Var(\alpha_i) > Var(\hat{\alpha}_i)$ for the first few EOFs

- Since the total variance is estimated with nearly zero bias by $Var(\vec{X}) = \sum_{j=1}^m \hat{\lambda}_j$, it follows that

$$Var(\alpha_i) < Var(\hat{\alpha}_i)$$

for the last few EOFs

2.2 Pattern Analyses

The Bias in Estimating Eigenvalues

The bias can be assessed using the following asymptotic formulae that apply to eigenvalue estimates computed from samples that can be represented by n iid normal random vectors (Lawley)

$$E(\hat{\lambda}_i) = \lambda_i \left(1 + \frac{1}{n} \sum_{\substack{j=1 \\ j \neq i}}^m \frac{\lambda_j}{\lambda_i - \lambda_j} \right) + O(n^{-2})$$

$$\text{Var}(\hat{\lambda}_i) = \frac{2\lambda_i^2}{n} \left(1 - \frac{1}{n} \sum_{\substack{j=1 \\ j \neq i}}^m \left(\frac{\lambda_j}{\lambda_i - \lambda_j} \right)^2 \right) + O(n^{-3})$$



- The eigenvalue estimators are consistent:

$$\lim_{n \rightarrow \infty} E\left((\hat{\lambda}_i - \lambda_i)^2 \right) = 0$$

- The estimations of the largest and the smallest eigenvalues are biased

$$E(\hat{\lambda}_i) \begin{cases} > \lambda_i & \text{for the largest } \lambda_i \\ < \lambda_i & \text{for the smallest } \lambda_i \end{cases}$$

- $E(\hat{\lambda}_i) > \lambda_i = \text{Var}(\alpha_i) > \text{Var}(\hat{\alpha}_i)$ for the largest λ_i
 $E(\hat{\lambda}_i) < \lambda_i = \text{Var}(\alpha_i) < \text{Var}(\hat{\alpha}_i)$ for the smallest λ_i

2.2 Pattern Analyses

Reliability of EOF estimates I

The reliability is often assessed using so-called selection rules. The basic supposition is

$$\text{full space} = \text{signal-subspace (EOFs)} + \text{noise-subspace (degenerated)}$$

Thus, the idea is to identify the signal-subspace as the space spanned by the EOFs that are associated with large, well-separated eigenvalues. This is done by considering the eigenspectrum

Problems

- The determination of signal- and noise-subspace is vague. Generally, the shape of the eigenspectrum is not necessarily connected to the presence or absence of dynamical signal
- No consideration of the reliability of the estimated patterns, since the selection rules are focused on the eigenvalues

2.2 Pattern Analyses

Reliability of EOF estimates II: North's Rule-of-Thumb

Using a scale argument, North et al. obtained an approximation for 'typical' error of the estimated EOFs, which in combination with a simplified version of Lawley's formula, reads

$$\Delta \hat{e}^i \approx \sqrt{\frac{2}{n}} \sum_{\substack{j=1 \\ j \neq i}}^m \frac{c}{\lambda_j - \lambda_i} \vec{e}^j$$
$$\approx \frac{c' \Delta \lambda_i}{\lambda_{\text{closest}} - \lambda_i} \vec{e}^j$$

where c and c' are constants, n is the number of independent samples, $\Delta \lambda \sim (2/n)^{1/2} \lambda_i$ the 'typical error' in $\hat{\lambda}_i$, λ_{closest} the closest eigenvalue to λ_i

- The first-order error is of the order of $(1/n)^{1/2}$. Thus convergence to zero is slow
- The first-order error is orthogonal to the true i -th EOF
- The estimate of the i -th EOF is most strongly contaminated by the patterns of those other EOFs that correspond to the eigenvalues λ_j closest to λ_i . The smaller the difference between λ_j and λ_i , the more severe the contamination

North's 'Rule-of-Thumb'

If the sampling error of a particular eigenvalue is comparable to or larger than the spacing between λ and a neighboring eigenvalue, then the sampling error of the i -th EOF will be comparable to the size of the neighboring EOF

→ EOFs are mixed

2.2 Pattern Analyses

North et al.'s Example

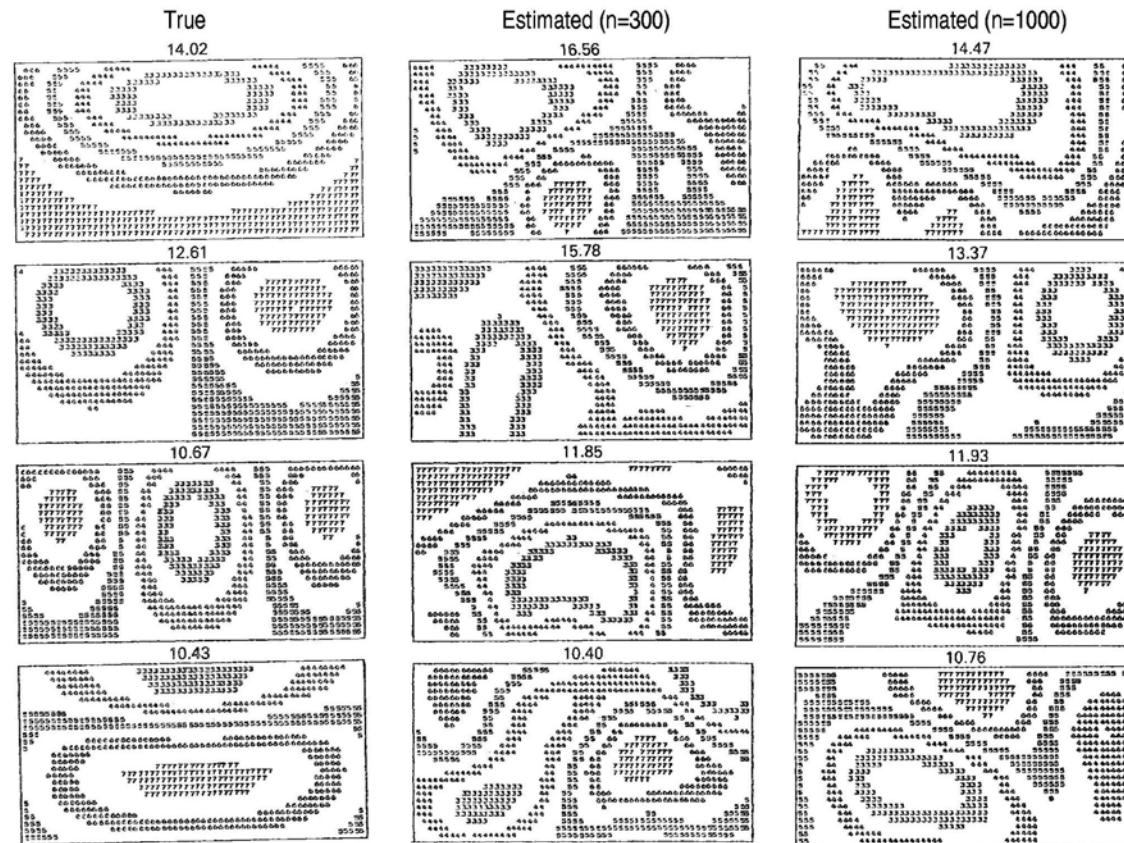
North et al. constructed a synthetic example in which the first four eigenvalues and the typical errors for the estimated eigenvalues are

$$\lambda_1=14, 12.6, 10.7, 10.4, \lambda_1-\lambda_2=1.4, \lambda_2-\lambda_3=2, \lambda_3-\lambda_4=0.3$$

$$|\Delta\lambda_i|=1, \text{ for } n=300, |\Delta\lambda_i|=0.6 \text{ for } n=1000$$

The first two EOFs are mixed when $n=300$.

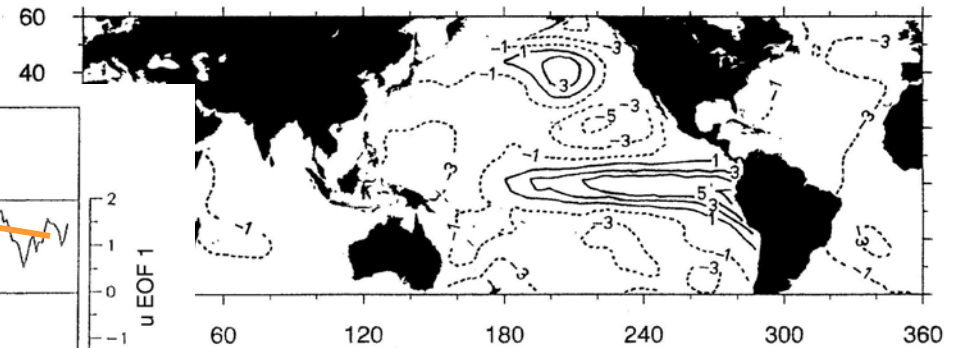
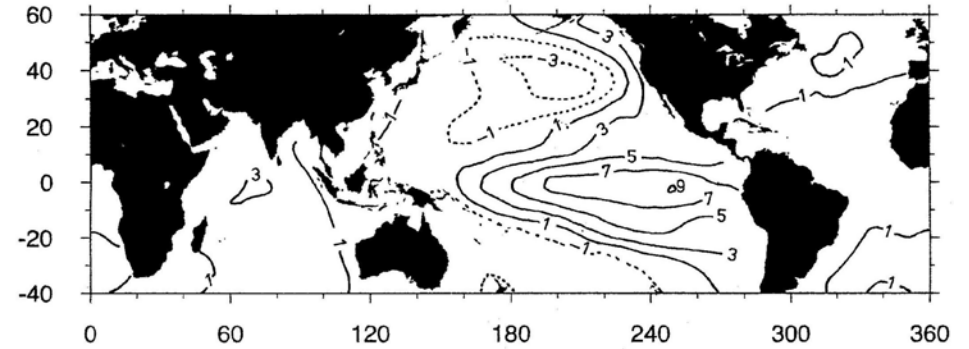
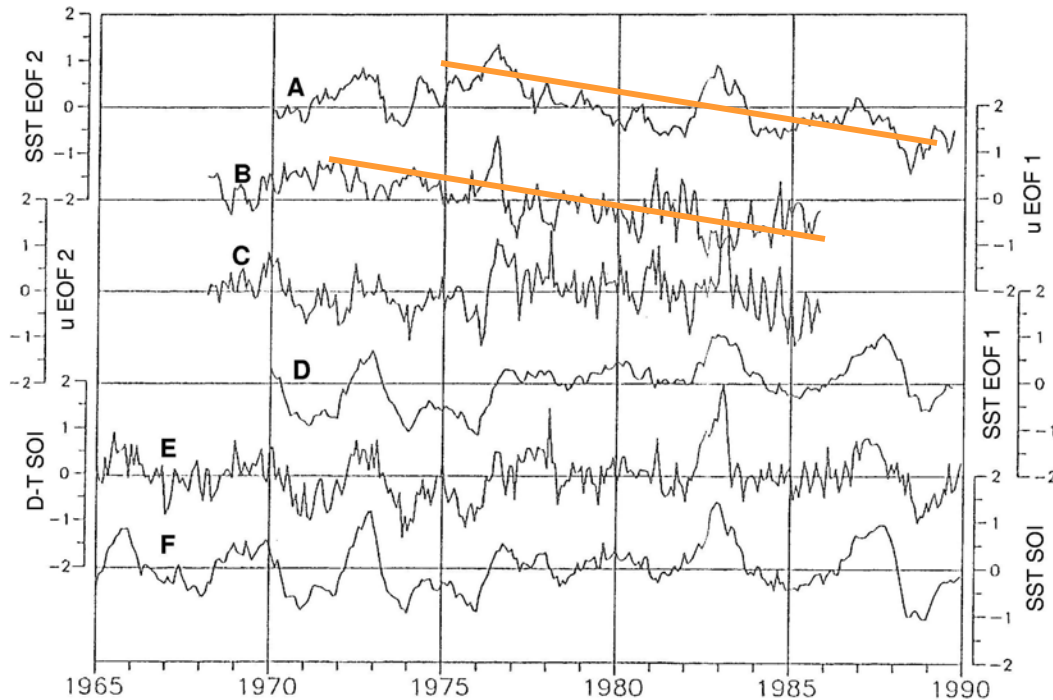
The third and fourth EOFs are mixed for both $n=300$ and $n=1000$



2.2 Pattern Analyses

Examples

- The first EOF represents ENSO, whose coefficient is shown as curve D
- The second EOF may represent trend, as suggested by its coefficient shown as curve A.



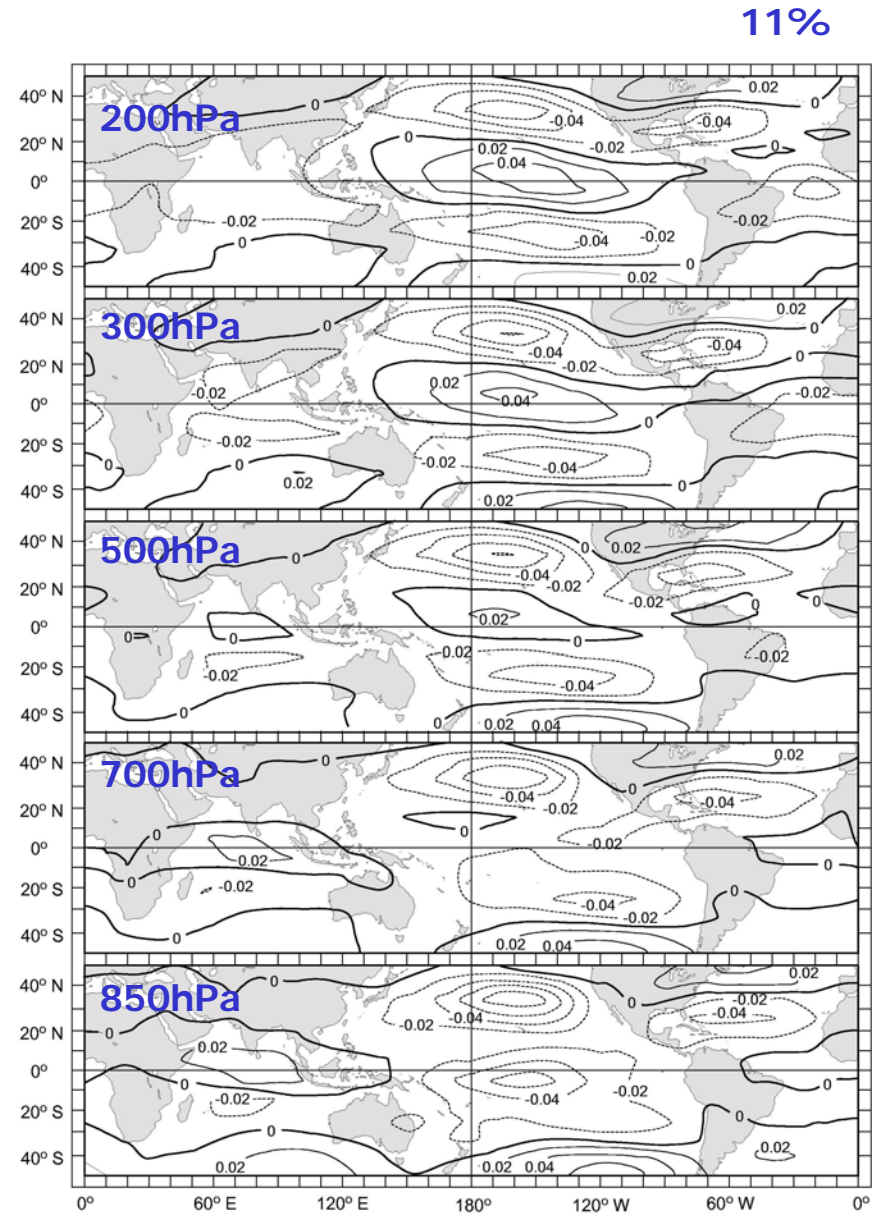
The first two EOFs of the monthly mean sea surface temperature of the global ocean between 40S and 60N

2.2 Pattern Analyses

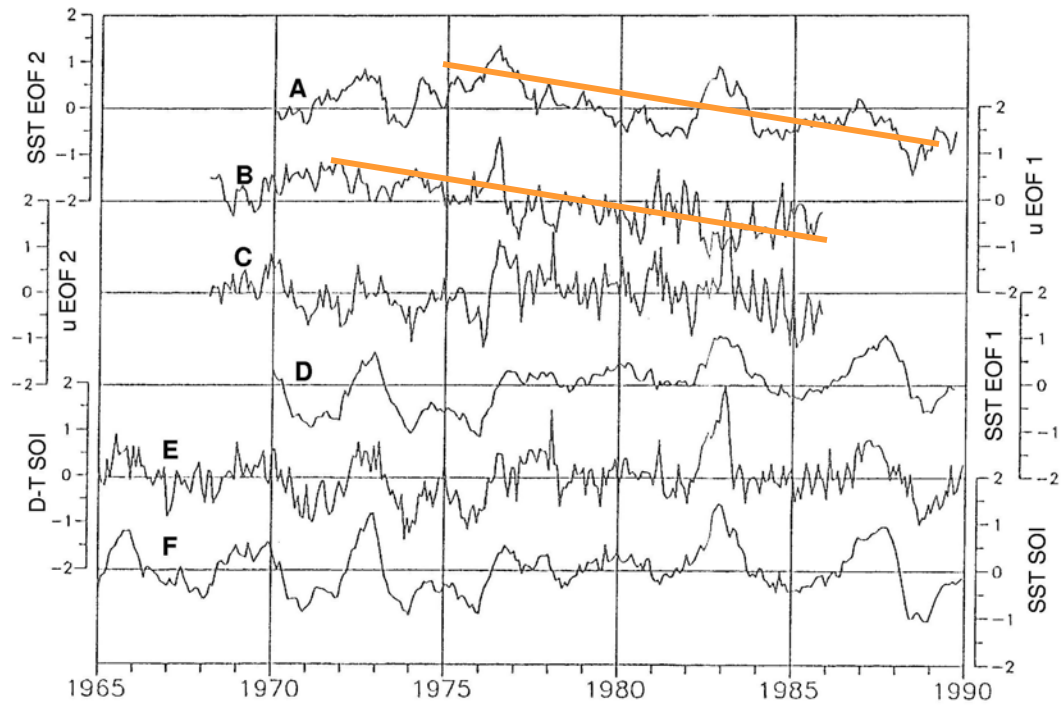
Examples

The first EOF of the tropospheric zonal wind between 45S and 45N at 850, 700, 500, 300 and 200 hPa

- The analysis is performed in two steps by first estimating EOF at each level and retaining coefficients representing 90% of the variance and secondly performing EOF analysis with a vector composing EOF coefficients selected for five levels
- The coefficient time series ([curve B](#)) exhibits a trend parallel to that found in the coefficient of the second SST EOF
- Does this trend originate from a natural low-frequency variation or from some other cause?



2.2 Pattern Analyses



2.2 Pattern Analyses

Rotation of EOFs

Why rotated EOFs?

One hopes that the rotated EOFs can be more easily interpreted than the EOFs themselves

The idea of 'rotation'

Given a subspace that contains a substantial fraction of the total variance, it is sometimes interesting to look for a linear basis of the subspace with specified properties, such as

- Basis vectors that contain *simple geometrical patterns*, e.g. patterns which are regionally confined or have two regions, one with large positive and the other with negative values
- Basis vectors that have *time coefficients with specific types of behavior*, such as having nonzero values only during some compact time episodes

The result depends on the number or the length of the input vectors, and on the measure of simplicity

Pro: a means for diagnosing physically meaningful and statistically stable patterns

2.2 Pattern Analyses

The Mathematics of the 'Rotation'

'Rotation' consists of a transformation and a constant

The transformation

A set of 'input' vectors $\mathbf{P} = (\vec{p}^1 | \dots | \vec{p}^K)$ is transformed into another set of vectors $\mathbf{Q} = (\vec{q}^1 | \dots | \vec{q}^K)$ by means of an invertible $K \times K$ matrix $\mathbf{R} = (r_{ij})$:

$$\mathbf{Q} = \mathbf{P}\mathbf{R}$$

or for each vector \vec{q}^i :

$$\vec{q}^i = \sum_{j=1}^K r_{ij} \vec{p}^j$$

The constraint

The matrix \mathbf{R} is chosen from a class of matrices, such as orthogonal ($\mathbf{R}^{-1} = \mathbf{R}^T$), subjected to the constraint that a functional $V(\mathbf{R})$ is minimized

2.2 Pattern Analyses

Consequence of a Orthogonal Transformation

A random vector which is represented by the K input vectors can be written, because of the rotation, as

$$\begin{aligned}\vec{X} &= \mathbf{P}\vec{\alpha} \\ &= (\mathbf{P}\mathbf{R})(\mathbf{R}^{-1}\vec{\alpha}) = \mathbf{Q}\vec{\beta}\end{aligned}$$

where $\vec{\alpha}$ and $\vec{\beta} = \mathbf{R}^{-1}\vec{\alpha}$ are K-dimensional vector of random expansion coefficients for the input and the rotated patterns, respectively.

If \mathbf{R} is orthonormal

$$\mathbf{Q}^T\mathbf{Q} = \mathbf{R}^T\mathbf{P}^T\mathbf{P}\mathbf{R} = \mathbf{R}^T\mathbf{D}\mathbf{R}$$

Thus, given orthogonal input vectors, the rotated vectors will be orthogonal only if $\mathbf{D} = \mathbf{I}$, or, if the input vectors are normalized to unit length


$$\Sigma_{\beta\beta} = \text{Cov}(\mathbf{R}^T\vec{\alpha}, \mathbf{R}^T\vec{\alpha}) = \mathbf{R}^T\Sigma_{\alpha\alpha}\mathbf{R}$$

Thus, given uncorrelated expansion coefficients of the input vectors, the coefficients of the rotated patterns are also pair wise uncorrelated only if coefficients α_j have unit variance

2.2 Pattern Analyses

Consequence of a Orthogonal Transformation

- The rotated EOFs derived from normalized EOFs are also orthogonal, but their time coefficients are not uncorrelated
- The rotated EOFs derived from non-normalized EOFs (i.e. the variance of EOF coefficients equal one) are no longer orthogonal, but the coefficients are pairwise uncorrelated

- 
- The result of the rotation depends on the lengths of the input vectors. Differently scaled but directionally identical sets of input vectors lead to sets of rotated patterns that are directionally different from one another

The rotated vectors are a function of the input vectors rather than the space spanned by the input vectors

- The rotated EOFs and their coefficients are not orthogonal and uncorrelated at the same time. Consequently, the percentage of variance represented by the individual patterns is no longer additive

2.2 Pattern Analyses

An Example of the Simplicity Functional: The 'Varimax' Method

'Varimax' is a widely used orthogonal rotation that minimizes the simplicity functional

$$V(\vec{q}^1, \dots, \vec{q}^K) = \sum_{i=1}^K f_V(\vec{q}^i)$$

with

$$\vec{q}^i = \sum_{j=1}^K r_{ij} \vec{p}^j,$$

$$f_V(\vec{q}) = \frac{1}{m} \sum_{i=1}^m \left(\frac{q_i}{s_i} \right)^4 - \frac{1}{m^2} \left(\sum_{i=1}^m \left(\frac{q_i}{s_i} \right)^2 \right)^2$$

- The functional f_V can be viewed as the spatial variance of the normalized squares $(q_i/s_i)^2$, i.e. f_V measures the 'weighted square amplitude' variance of the rotated EOF

- The constants s_i can be chosen freely. One deals with

a raw varimax rotation when $s_i=1$

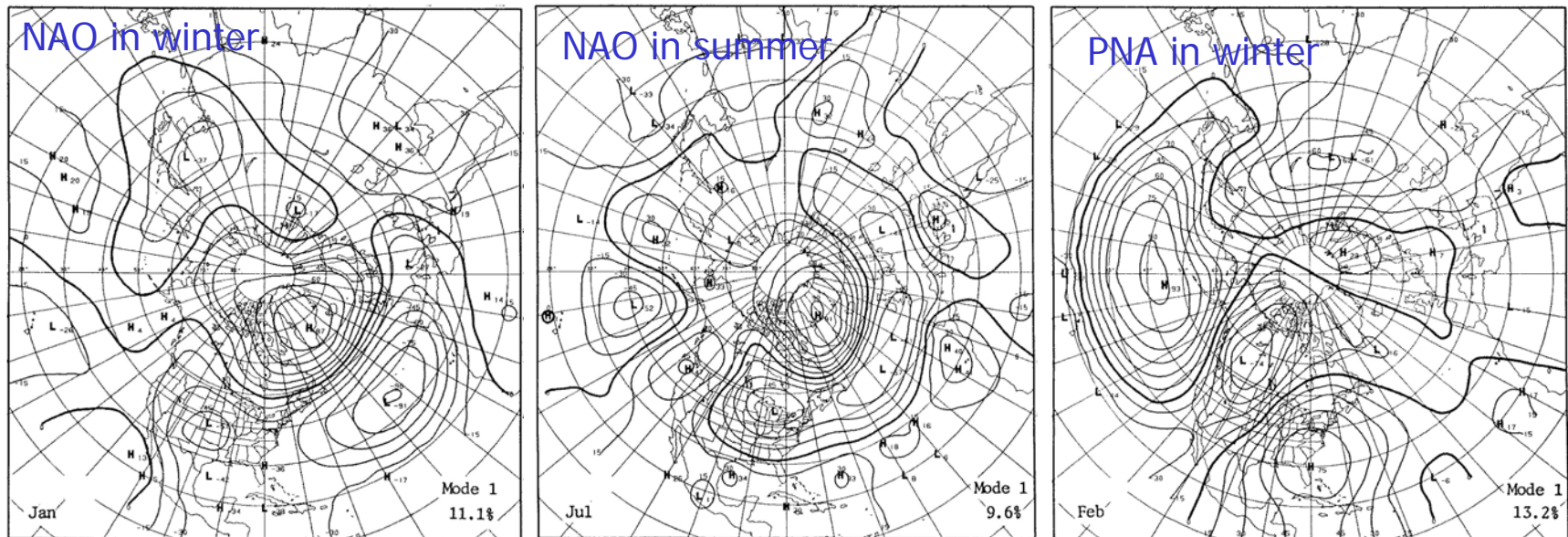
a normal varimax rotation when $s_i = \sum_{j=1}^K (p_i^j)^2$

2.2 Pattern Analyses

Example I: Reproducible Identification of Teleconnection Patterns

Barzon and Livzey used a varimax rotation of normalized EOFs to isolate the dominant circulation patterns in the Northern Hemisphere:

- EOFs are computed for each calendar month using a 35-year data set of monthly mean 700 hPa heights
- Rotation is performed on the first 10 EOFs representing 80% of the total variance in winter and 70% in summer

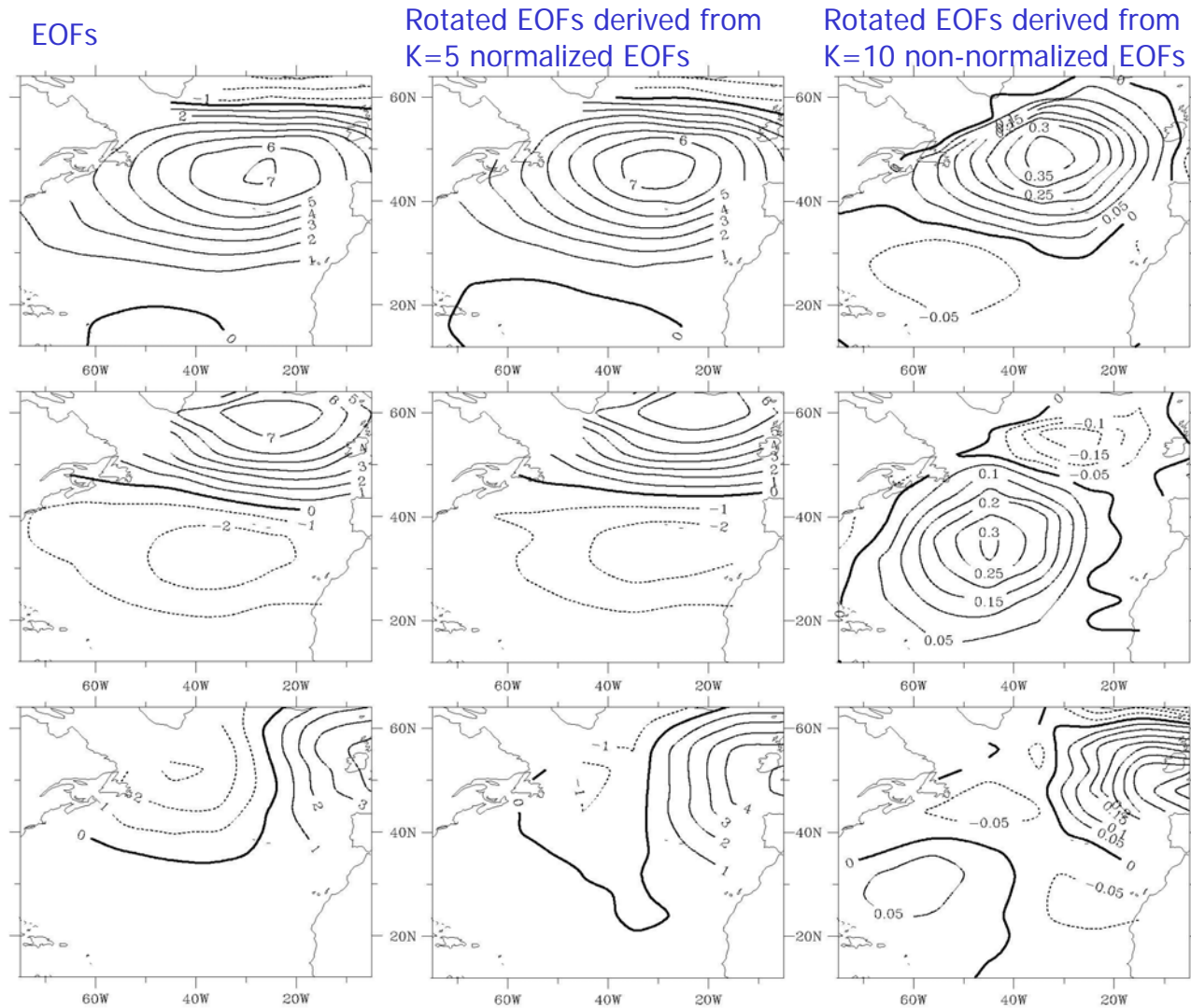


2.2 Pattern Analyses

Example II: Weak Effect of Rotation

EOFs and rotated EOFs of North Atlantic monthly mean SLP in winter:

- The difference between the unrotated and the rotated EOF is not large
- If the EOFs have simple structures, the effect of rotation is negligible



2.2 Pattern Analyses

Example III: Rotation could split features into different patterns even though they are part of the same physical pattern

EOFs and rotated EOFs of North Atlantic monthly mean SST in DJF:

- The rotated EOFs tend to represent the three action centers in the first EOF separately in different EOFs

EOFs

Rotated EOFs derived from K=5 normalized EOFs

Rotated EOFs derived from K=5 non-normalized EOFs

