Section 6

Principal Component and Maximum
Covariance Analyses

Maximum Covariance Analysis (MCA)

Purpose

- **Purpose**
  - Find patterns in 2 datasets which are highly correlated (i.e. are frequently met simultaneously). E.g. patterns of SST that go along with patterns of SLP.

- **Applications**
  - Study the coupling between parameters to understand physical mechanisms of climate variations. E.g. how do SST and SLP mutually influence each other?
  - Statistical downscaling. (Translate GCM derived climate change scenarios to the local/regional scale.)
  - Reconstruction / Forecasting

Paired Multivariate Datasets

- **Two paired multivariate datasets**

$$\mathbf{x}(i) = \left(x_1(i),..,x_n(i)\right) \qquad \mathbf{y}(i) = \left(y_1(i),..,y_m(i)\right)$$

  - i=1,.., N, joint observations of $\mathbf{x}$, $\mathbf{y}$
  - Two data clouds in two phase spaces
  - Centered (means subtracted!), (just for more simple notation!)
  - E.g. $\mathbf{x}$: anomalies of SLP Europe/Atlantic (n gridpoints),
    $\mathbf{y}$: Temperature anomalies in CH (m stations)
    all Januaries of the 20th century (N=100)

- **Data Matrices**

$$\mathbf{X} = \begin{bmatrix} x_1(1) & \cdots & x_n(1) \\ \vdots & \ddots & \vdots \\ x_1(N) & \cdots & x_n(N) \end{bmatrix} \qquad \mathbf{Y} = ...$$

Cross-Covariance

Call me
cross-covariofant

- **Cross-covariance matrix**

$$\mathbf{S}_{xy} := \begin{bmatrix} \mathrm{cov}(x_1, y_1) & \mathrm{cov}(x_1, y_j) & \mathrm{cov}(x_1, y_m) \\ \mathrm{cov}(x_i, y_1) & \mathrm{cov}(x_i, y_j) & \vdots \\ \mathrm{cov}(x_n, y_1) & \ldots & \mathrm{cov}(x_n, y_m) \end{bmatrix} = \frac{1}{(N-1)} \mathbf{X}^T \cdot \mathbf{Y}$$

  - Univariate cross-covariances between all pairs of components
  - n x m matrix
  - In general not square, not symmetric
  - diagonal elements do not have special meaning

Cross-Correlation

- **Cross-Correlation matrix**

$$\mathbf{C}_{xy} := \begin{bmatrix} \mathrm{cor}(x_1, y_1) & \mathrm{cor}(x_1, y_j) & \mathrm{cor}(x_1, y_m) \\ \mathrm{cor}(x_i, y_1) & \mathrm{cor}(x_i, y_j) & \vdots \\ \mathrm{cor}(x_n, y_1) & \dots & \mathrm{cor}(x_n, y_m) \end{bmatrix}$$

  - Analogous to $\mathbf{S}_{\mathrm{xy}}$ but divided by standard deviations
  - All matrix elements in {–1,+1}

$$\mathbf{C}_{xy} = \frac{1}{N-1}\mathbf{D}_x^{-\frac{1}{2}} \cdot \mathbf{X}^T \cdot \mathbf{Y} \cdot \mathbf{D}_y^{-\frac{1}{2}} = \mathbf{D}_x^{-\frac{1}{2}} \cdot \mathbf{S}_{xy} \cdot \mathbf{D}_y^{-\frac{1}{2}}$$

$$\mathbf{D}_x^{-\frac{1}{2}} = \left[\, {}^{1}\!\big/_{\sigma_{x_k}} \right] \qquad \mathbf{D}_y^{-\frac{1}{2}} = \left[\, {}^{1}\!\big/_{\sigma_{y_k}} \right]$$

MCA Mathematical Procedure

- **Singular Value Decomposition (SVD)**

>> Appendix A

$\mathbf{S}_{xy} = \mathrm{cov}(\mathbf{X},\mathbf{Y})$ the cross-covariance matrix (n x m)

➨ There are r real numbrs $\{\omega_1, \omega_2,.., \omega_r\}$, $\omega_k>0$, *singular values*
and r vectors $\{\mathbf{u_1}, \mathbf{u_2}, .., \mathbf{u}_r\}$, n-dim., unit-length, orthogonal,
and r vectors $\{\mathbf{v_1}, \mathbf{v_2}, .., \mathbf{v}_r\}$, m-dim., unit-length, orthogonal,
called *left* ($\mathbf{u}_k$) and *right* ($\mathbf{v}_k$) *singular vectors,*
such that:

$$\mathbf{S}_{xy} = \mathbf{U}^T \cdot \Omega \cdot \mathbf{V} \qquad \Omega = [\omega_k], \quad \text{diagonal, } r \times r$$

$$\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_r], \quad \mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r], \quad \text{vectors in columns}$$

`R: svd`

MCA Interpretation

- **Singular Vectors, Coefficients**
  - Consider singular vectors anchored in center of data clouds, spanning sub-spaces of the phasespaces of each dataset (new variables). Not necessarily a basis system ($r \le n$, $r \le m$!)
  - Projections of data $\mathbf{x}(i)$ onto singular vectors are new data coordinates:

$$a_j(i) = \mathbf{x}(i)^T \cdot \mathbf{u}_j, \quad b_j(i) = \underbrace{\mathbf{y}(i)^T \cdot \mathbf{v}_j}_{\text{projection of data vector on singular vector}}, \quad j = 1,..,r$$

  - New coordinates are linear combinations of original variables.
  - $a_j(i)$: *left coefficients*, (also *left SVD scores*)
    $b_j(i)$: *right coefficients*, (also *right SVD scores*)

MCA Interpretation

- **Singular values, cross-covariance**
  - Ω is the cross-covariance matrix of the new coordinates $\{a_k\}$, $\{b_k\}$:

>> Appendix B

$$\mathrm{cov}(a_i(.), b_j(.)) = 0 \quad \text{for } i \neq j, \quad \mathrm{cov}(a_i(.), b_i(.)) = \omega_i$$

➨ Coordinates corresponding to different indices of singular vectors are mutually uncorrelated.

➨ The first pair of singular vectors $\{\mathbf{u}_1, \mathbf{v}_1\}$ are the phase-space directions, for which the projections have the largest possible cross-covariance. *First coupled mode*.

Subsequent vector pairs $\{\mathbf{u}_k, \mathbf{v}_k\}$ maximise cross-covariance subject to orthoganility on previous pairs. *k^{th} coupled mode.*

MCA Interpretation

- **Within space variance**
  - Singular vectors do not maximize variance in individual spaces.
  - Singular vectors are not necessarily aligned along directions of large data spread or cloud symmetry.
  - Left and right coefficients are in general not uncorrelated between themselves:

$$\mathrm{var}\left(a_i, a_j\right) \neq 0 \quad \mathrm{var}\left(b_i, b_j\right) \neq 0 \quad \text{for } i \neq j$$

  - *Cumulative Explained Variance Fraction* of first l modes

$$CEVF_x^l = \sum_k^l \mathrm{var}(a_k) \Big/ \mathrm{tr}\left(\mathbf{S}_{xx}\right) \qquad CEVF_y^l = \sum_k^l \mathrm{var}(b_k) \Big/ \mathrm{tr}\left(\mathbf{S}_{yy}\right)$$

MCA Interpretation

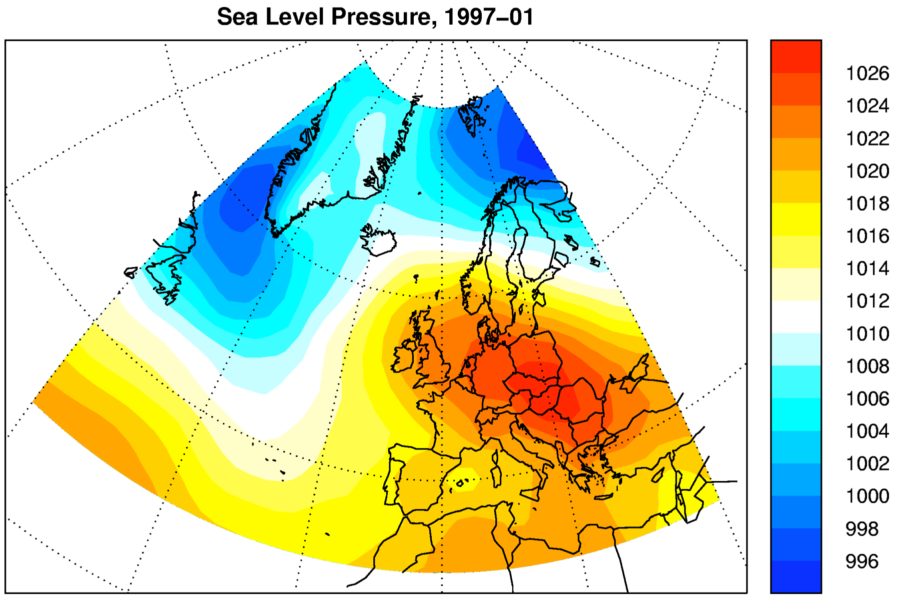
- **Singular vectors describe patterns of anomalies in each dataset that tend to occur simultaneously (are linearly correlated). *Modes of co-variability*, *coupled modes, canonical pairs*.**

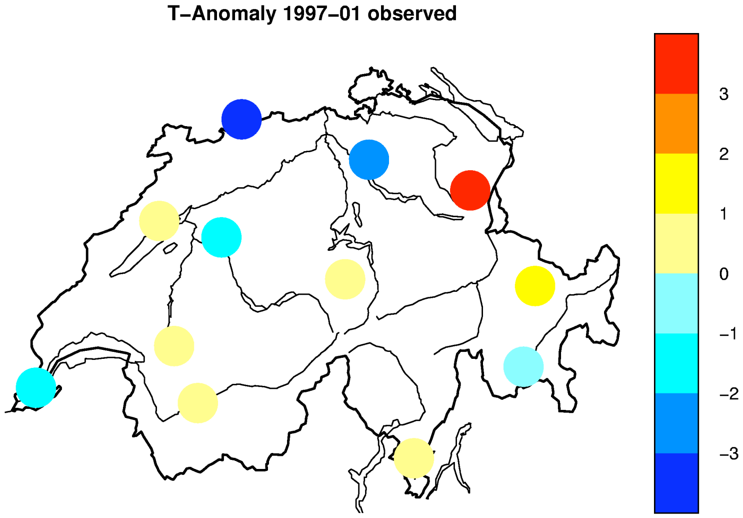
- **Coefficients represent amplitude (emphasis) of the respective patterns in each sample.**

- **The first few coefficient pairs have large co-variance and often show high correlations. *Dominant modes*.**

Example: Swiss T <> SLP

How does sea level pressure influence winter temperature-distributions in Switzerland?



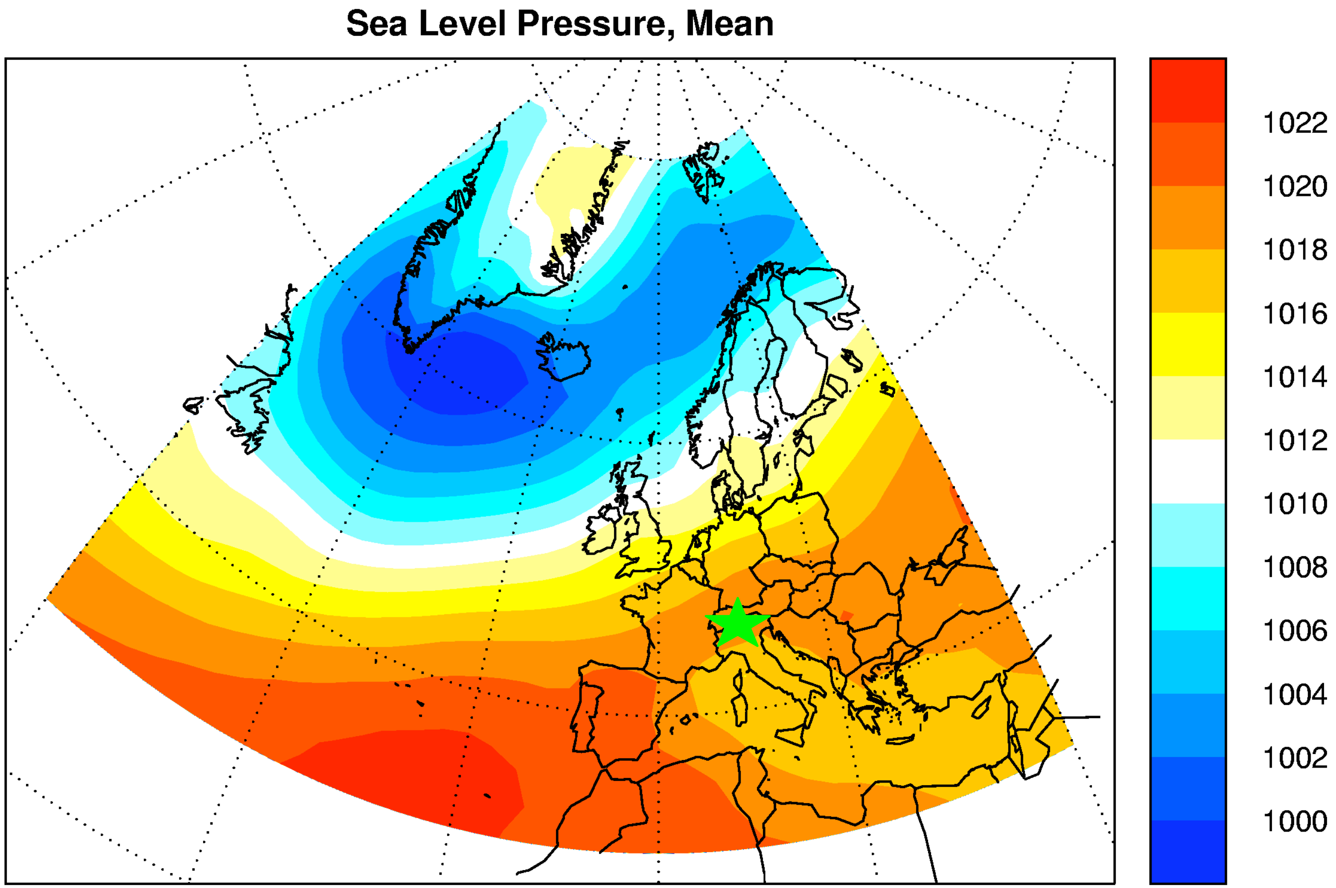


SLP: DJF, 1957 –1994
(60°W–40°E, 30°N–80°N)
41 x 21 grid points, 2.5 degrees
ECMWF, ERA40

T: DJF, 1957 –1994
12 stations
MeteoSwiss

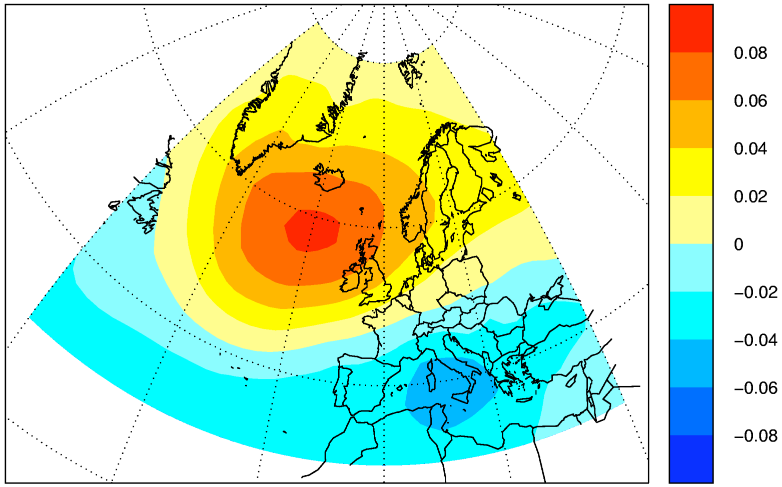
Begert et al. 2005, Simmons&Gibbson 2000

Example: Mean winter SLP

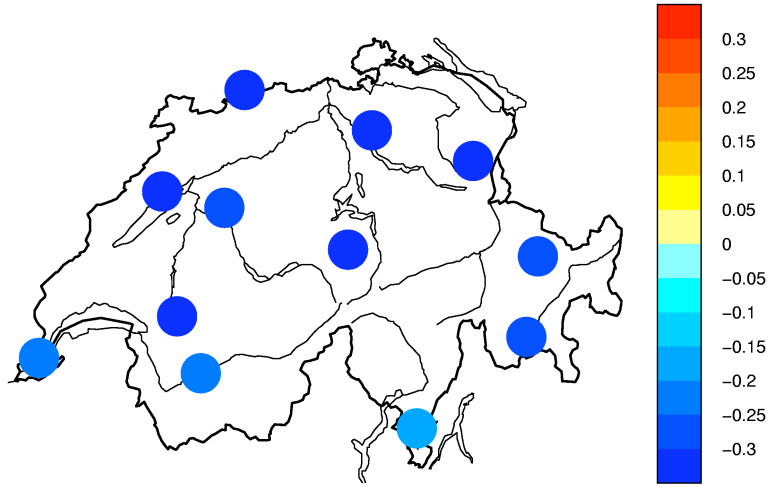


Example: Singular Vector Pair 1

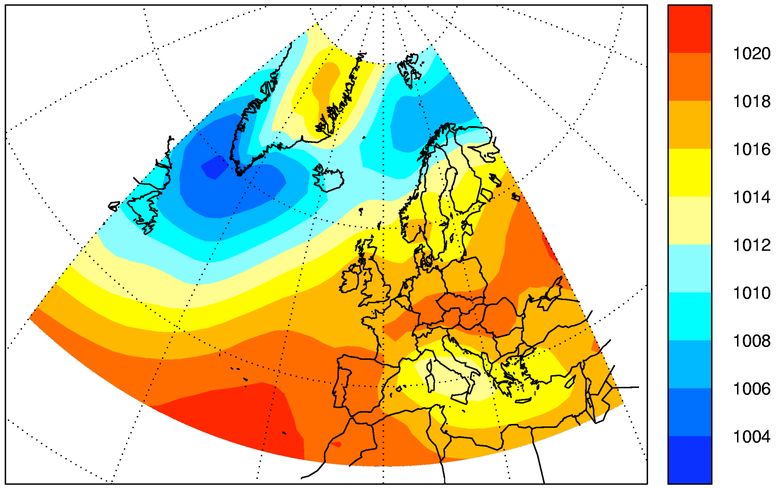
Sea Level Pressure, Singular Vector 1



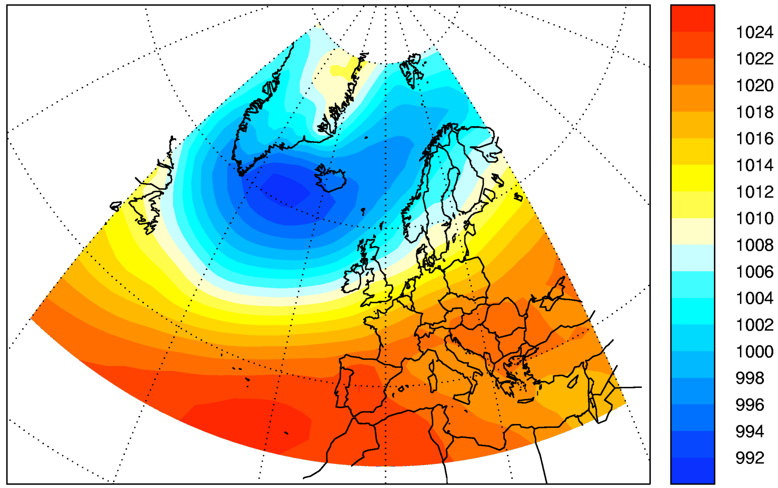
Temperature, Singular Vector 1



Sea Level Pressure, Mean+Singular Vector 1

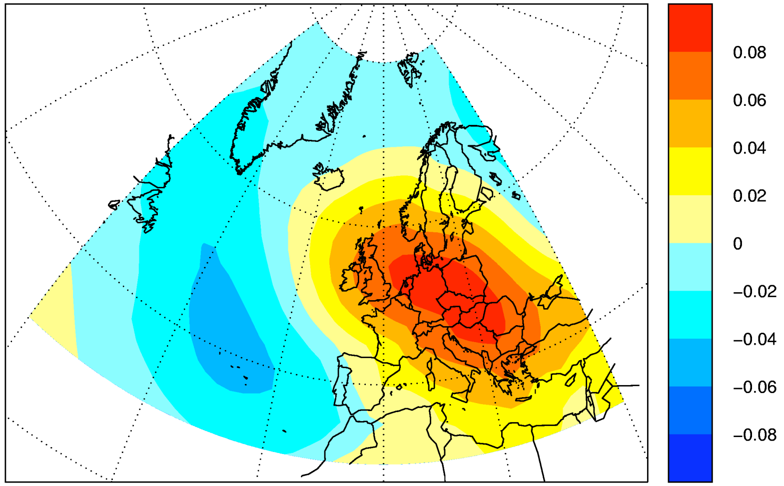


Sea Level Pressure, Mean−Singular Vector 1

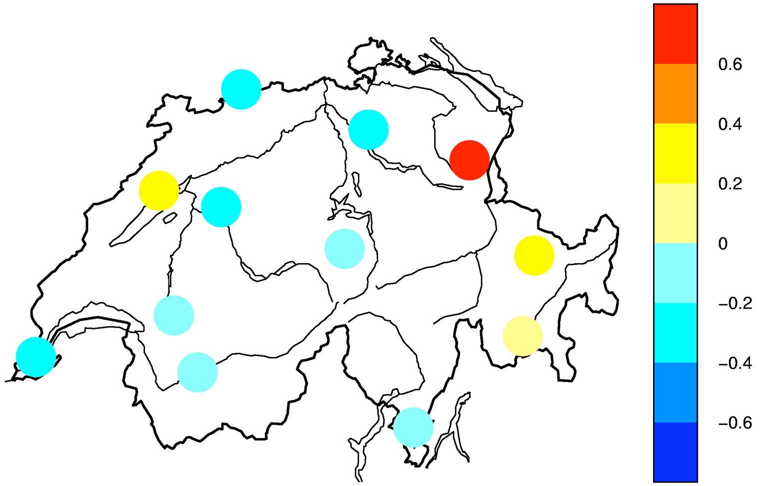


Example: Singular Vector Pair 2

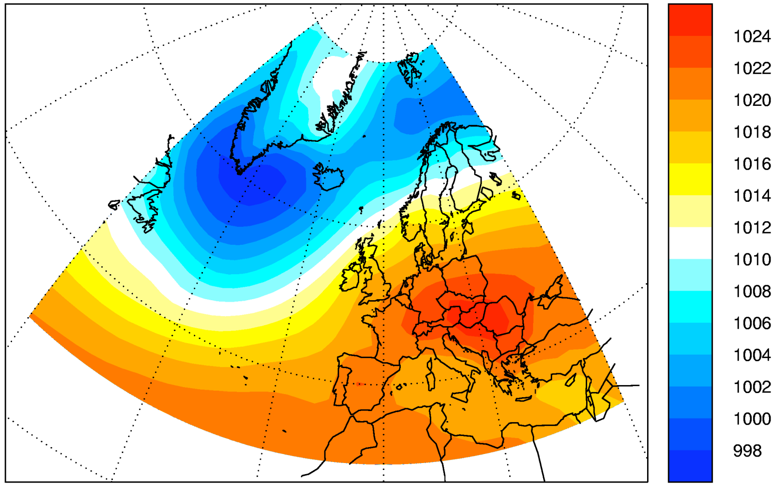
Sea Level Pressure, Singular Vector 2



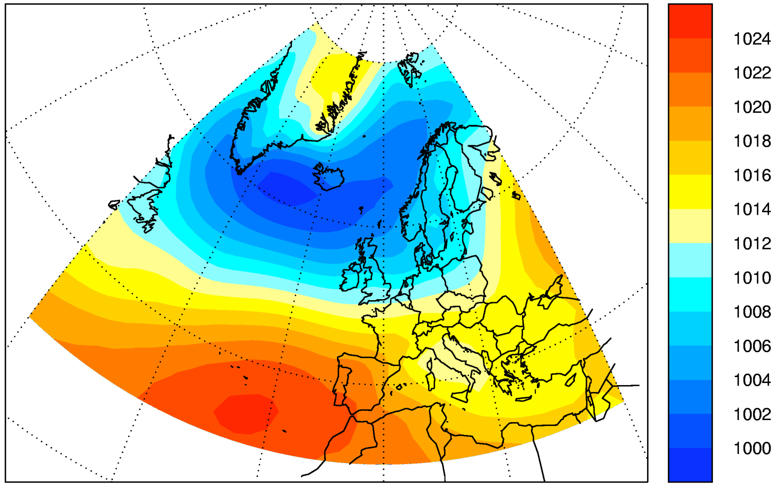
Temperature, Singular Vector 2



Sea Level Pressure, Mean+Singular Vector 2

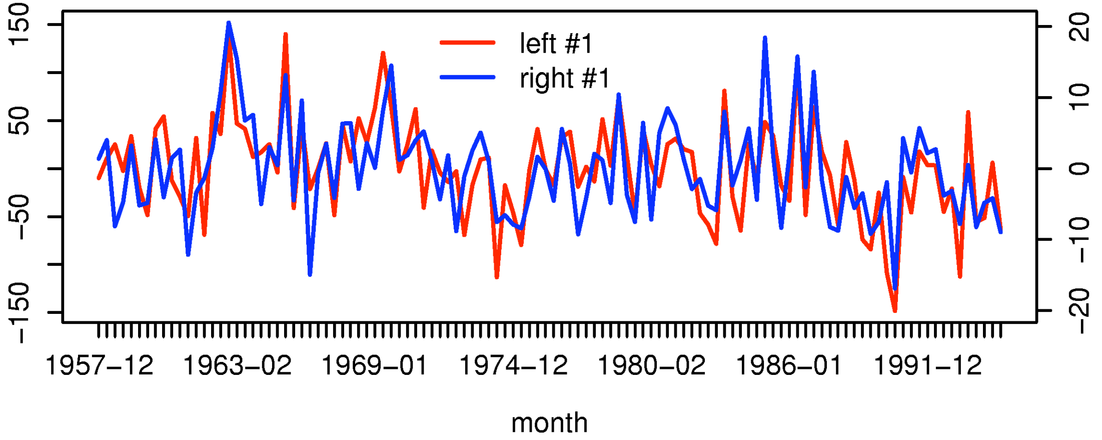


Sea Level Pressure, Mean−Singular Vector 2

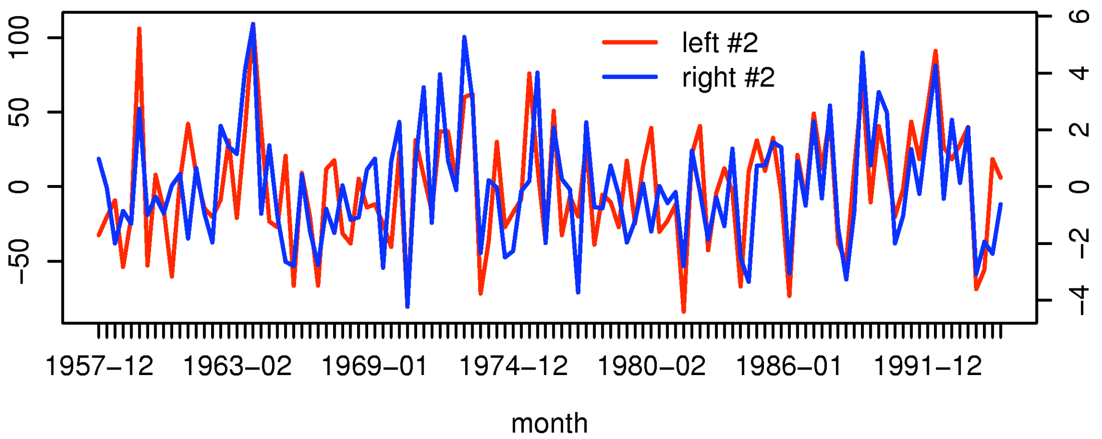


Example: Coefficients

Coefficients 1st mode



Coefficients 2nd mode



Mode	Correlation
1	0.70
2	0.67
3	0.60
4	0.39
5	0.28
6	0.37
7	0.29
8	0.22
9	0.33
10	0.20
11	0.34
12	0.34

Measuring Cross-Covariance

- ***Total squared covariance*:**

$$\left\| S_{xy} \right\|_F := \sum_i^n \sum_j^m s_{ij}^2 = \sum_i^r \omega_i^2$$

Frobenius Norm

- ***Squared covariance fraction:***
  - of singular vect. pair k:

$$SCF_k = \omega_k^2 \Big/ \sum_i^r \omega_i^2$$

  - Note *squared* quantities compared to PCA!

Truncation

- **Retain only first l coupled modes**
- **Projection onto first l modes**
  - yields an approximation of original data

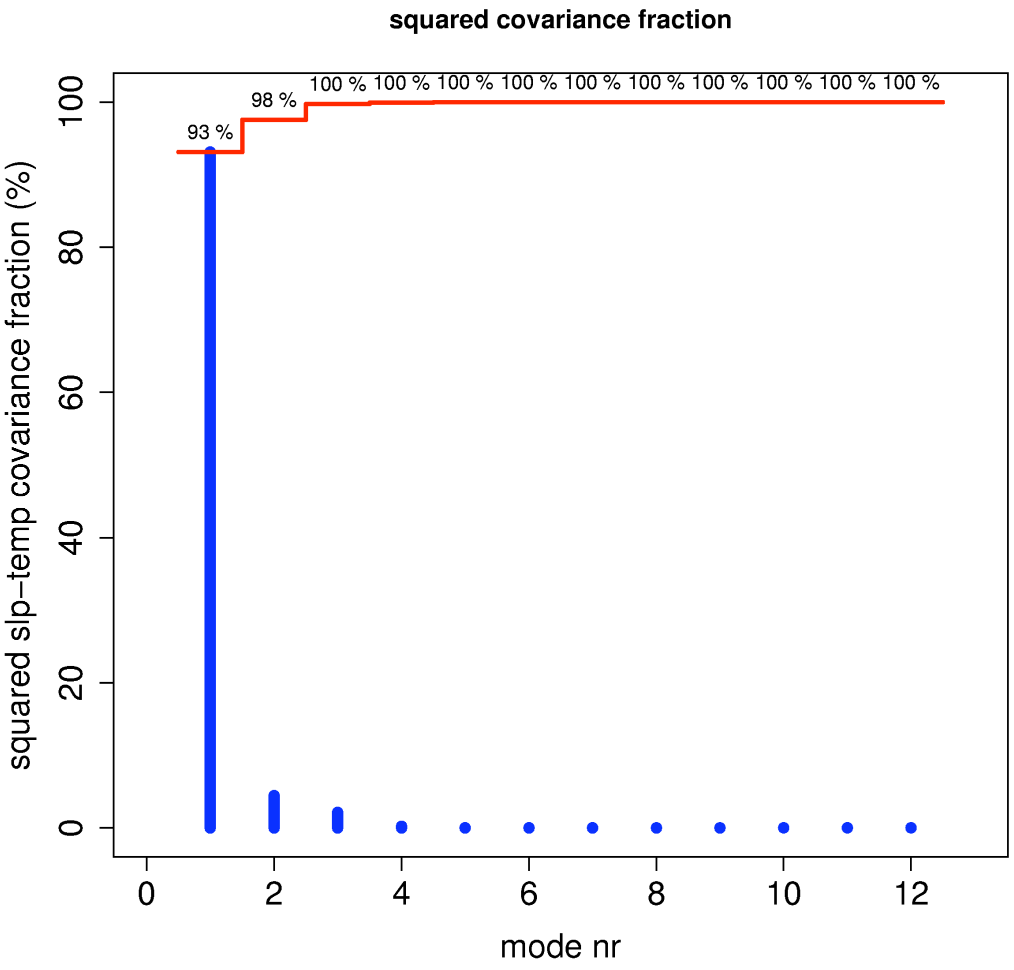
$$ {}_l\tilde{\mathbf{x}}(i) = \sum_k^l a_k(i)\cdot \mathbf{u}_k \qquad {}_l\tilde{\mathbf{y}}(i) = \sum_k^l b_k(i)\cdot \mathbf{v}_k \qquad l \le r $$

- ***Cumulative squared covariance fraction*:**

Residual cross-covariance,
I.e. not reproduces by approximation

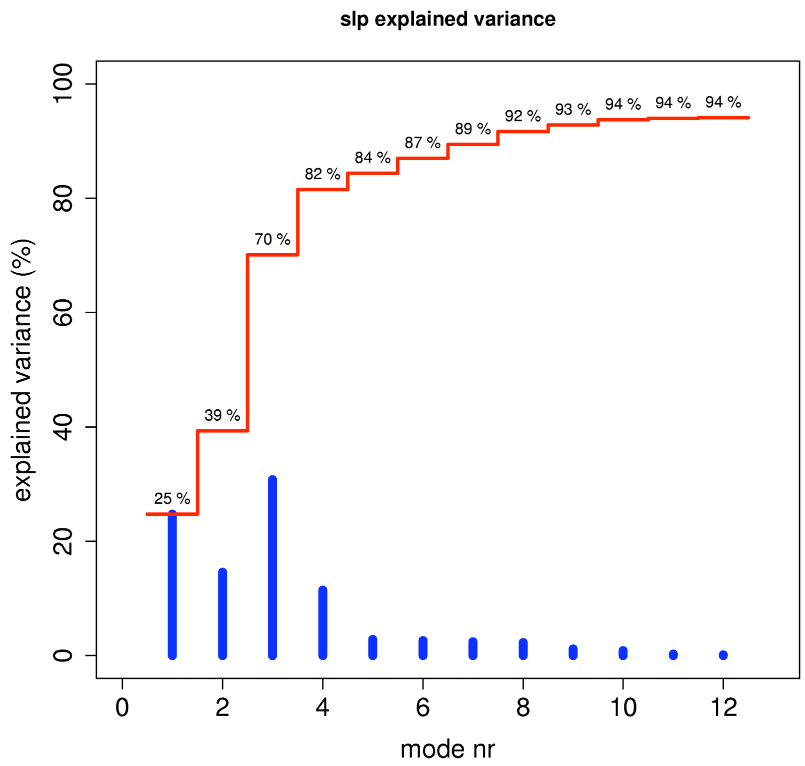
$$ CSCF_l = 1 - \frac{\left\| \overbrace{\mathbf{S}_{xy} - \tilde{\mathbf{S}}_{xy}^l} \right\|_F}{\left\| \mathbf{S}_{xy} \right\|_F} = \sum_k^l \omega_k^2 \bigg/ \sum_k^r \omega_k^2 = \sum_k^l SCF_k $$

Example: Squared Covariance Fraction

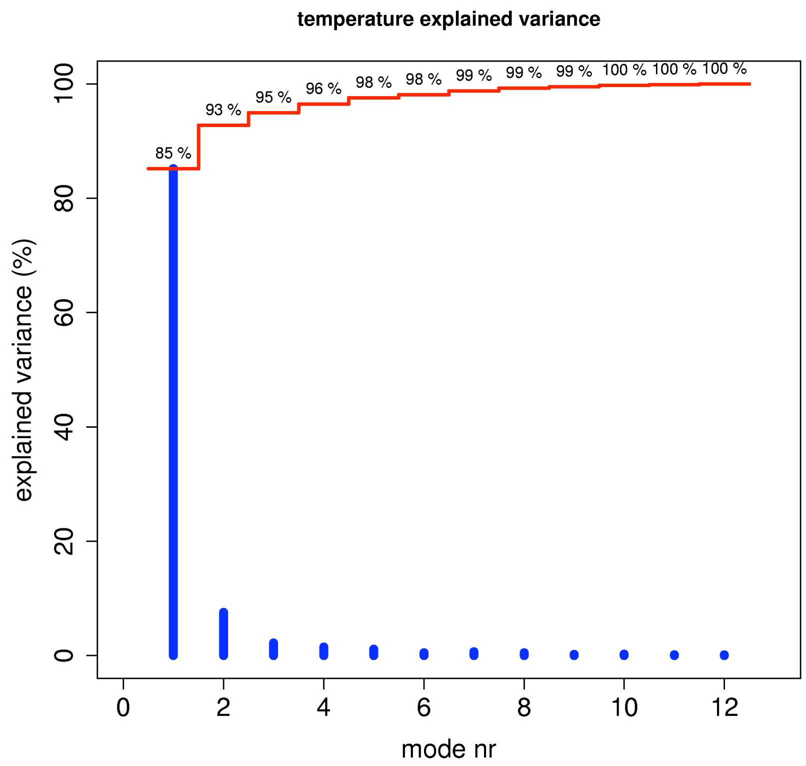


Example: Explained Variance

Sea Level Pressure



Temperature



Reconstruction / Prediction

- **Purpose**
  - Exploit cross-covariance to reconstruct/predict a right data from a left data (or vice versa).
- **Linear model between left and right coefficients**

right coefficients left coefficients (predictors)

$$b_k(i) = \beta_{k0} + \sum_{l}^{r} \beta_{kl} \cdot a_l(i) + \varepsilon_k(i)$$

random noise

regression coefficients

  - Simplifies to:

$$b_k(i) = \beta_{kk} \cdot a_k(i), \quad \beta_{kk} = \frac{\omega_k}{\mathrm{var}(a_k)}$$

$$a_k, b_l \quad centered \Rightarrow \beta_{k0} = 0$$

$$\mathrm{var}(a_k, b_l) = 0 \Rightarrow \beta_{kl} = 0, \; k \neq l$$

Reconstruction / Prediction

$$\mathbf{x}$$

a “predictor” data vector
in the left phase space
(centered)

$$\mathbf{a} = \mathbf{U}^T \cdot \mathbf{x}$$

$$\mathbf{b} = \mathbf{G} \cdot \mathbf{a} \quad \mathbf{G} = \left[\omega_k / \mathrm{var}(a_k)\right]$$

$$\mathbf{y} = \mathbf{V} \cdot \mathbf{b}$$

Operations using singular vectors
and singular values from a prior MCA

the “predictand” data vector
in the right phase space
(centered)

$$\mathbf{y}$$

- **MCA prediction / reconstruction equation:**

$$\mathbf{y} = \mathbf{V} \cdot \mathbf{G} \cdot \mathbf{U}^{\mathbf{T}} \cdot \mathbf{x}$$

  - Possibly only using a few leading coupled modes. I.e. all matrices truncated to the number of desired modes.

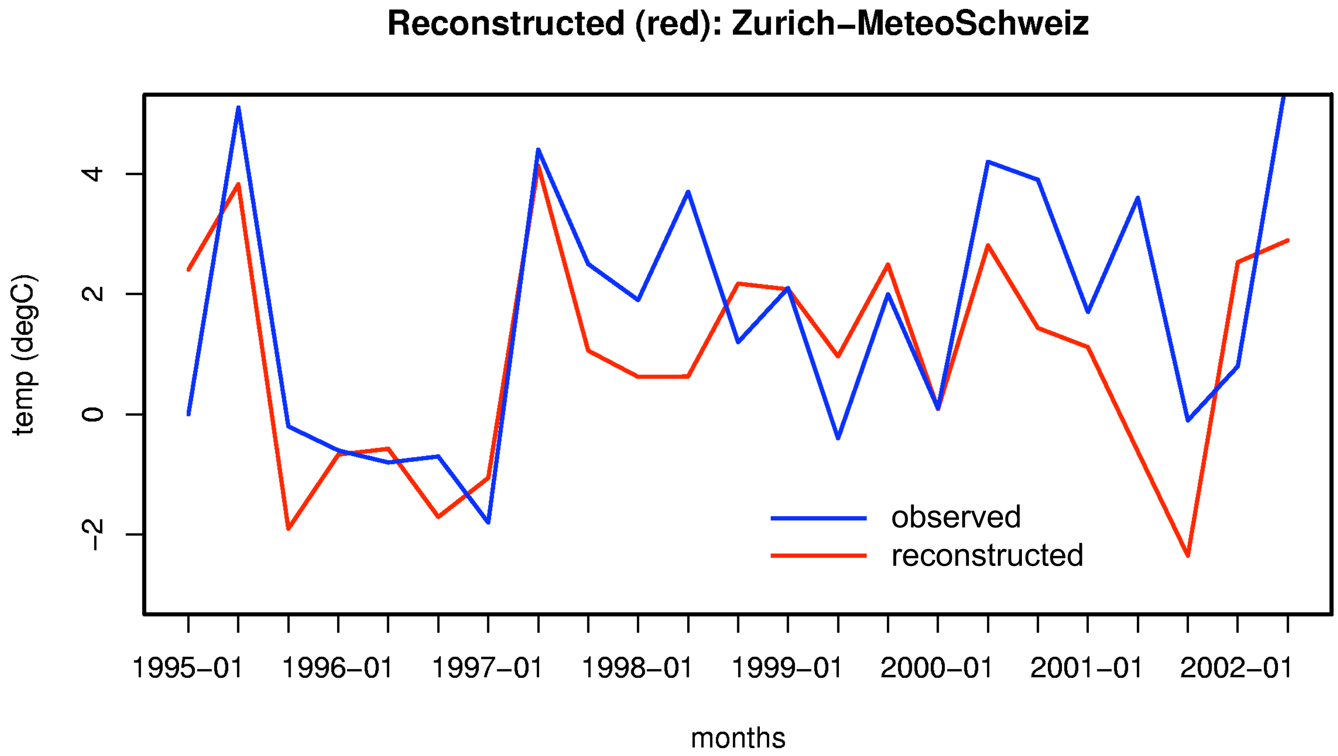
Reconstruction / Prediction

- **Accuracy of reconstruction depends on**
  - The degree of covariance between the two fields.
  - The cumulative squared covariance fraction represented by the leading modes. The more modes the better the reconstruction.
  - The variance explained by the right singular vectors in the right phase space. I.e. the component of the right space that is related to the left space.

Example: Reconstruction

Reconstruction for 1995-2002 (winter months) using:

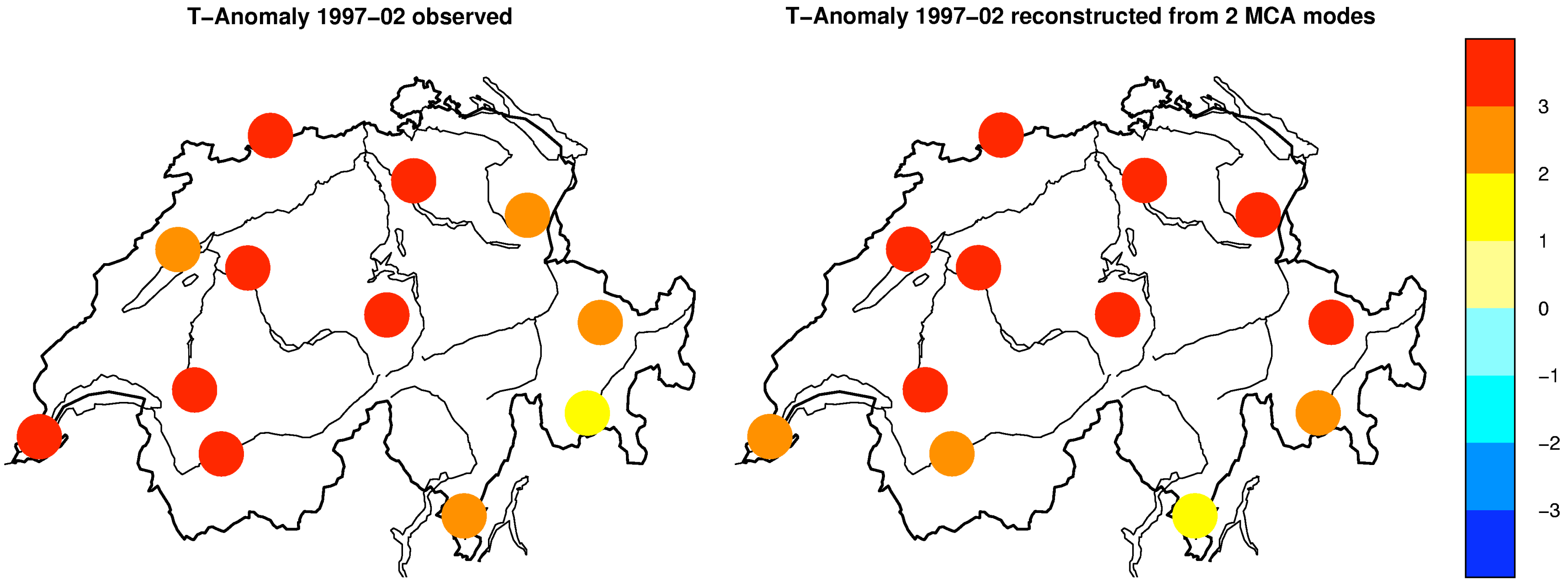
SLP (left field) as predictor
MCA calibrated for 1957-1994
2 leading coupled modes



Example: Reconstruction

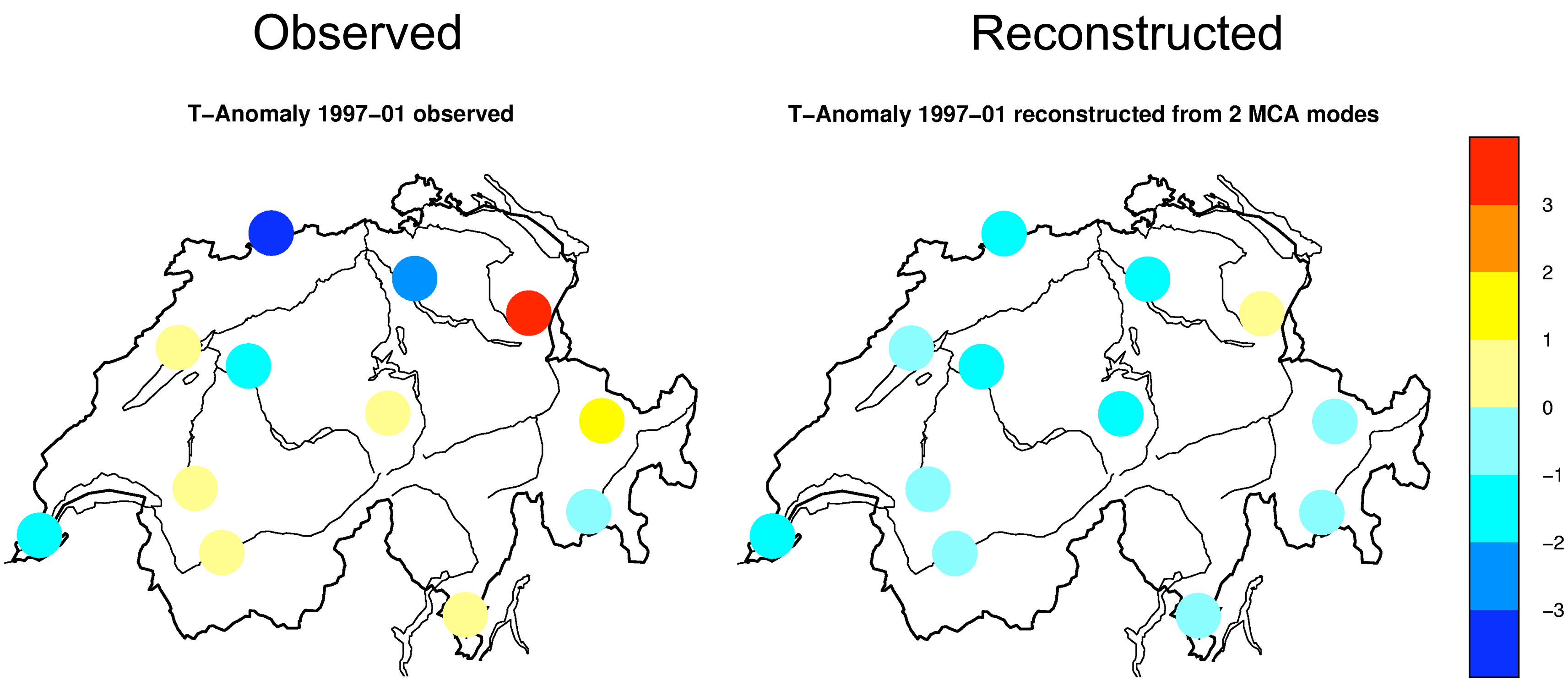
February 1997

Observed | Reconstructed



Example: Reconstruction

January 1997



Canonical Correlation Analysis

- **Similar to MCA but searching for max. correlation**
- **Procedure**
  - Conduct SVD with cross-correlation matrix
- **Interpretation (like for MCA)**
  - Canonical pairs describe coupled modes, emphasis on correlation
- **CCA or MCA?**
  - MCA focues on *covariance*: Modes tend to be large where the variance is large. Danger that physical modes are confounded by large variance.
  - CCA focuses on *correlation*: Coupling is identified also if associated variance is low. Danger that physical modes are confounded by small (insignificant) variations (sampling problems).

(see Wilks 2005, Chap 12 for details)

Summary

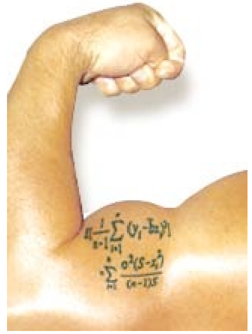
- **PCA & MCA are interesting techniques for characterising the (co-)variability in spatial datasets.**
- **Results require careful interpretation, corroboration in sensitivity experiments.**
- **Well established instruments in climate science:**
    - Technique of data reduction
    - Parsimonous reconstruction models (e.g. in historical climatology, paleo climatology).
    - Evaluation of physical mechanisms in climate models.
    - Statistical (empirical) forecasting, climate change downscaling.
    - Source of hypothesis building for later modelling exercises.

Section 6

Principal Component and Maximum
Covariance Analyses

Appendix MCA

Appendix A



- **Singular Value Decomposition (SVD)**

$\mathbf{Q}$ a real-valued nxm matrix (e.g. a cross-covariance matrix)

$r = \text{rank}(\mathbf{Q}) \leq \min(n,m)$

➨ There exist real-valued matrices $\mathbf{U}$, $\mathbf{V}$, Ω such that:

$$\mathbf{Q} = \mathbf{U} \cdot \Omega \cdot \mathbf{V}^T$$

Ω: a diagonal rxr matrix:

$\omega_1 \geq \omega_2 \geq .. \geq \omega_r > 0$ the *singular values*

$\mathbf{U}$: a n x r matrix $\mathbf{U} =[\mathbf{u}_1,\mathbf{u}_2,..,\mathbf{u}_r]$, orthonormal: $\mathbf{U}^T \cdot \mathbf{U} = \mathbf{1}$

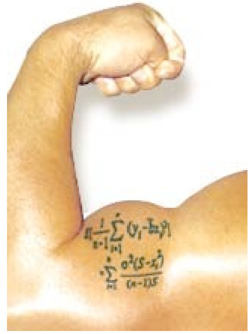
$\mathbf{u}_k$ $k=1$ .. r, the *left singular vectors*.

$\mathbf{V}$: a m x r matrix $\mathbf{V} =[\mathbf{v}_1,\mathbf{v}_2,..,\mathbf{v}_r]$, orthonormal: $\mathbf{V}^T \cdot \mathbf{V} = \mathbf{1}$

$\mathbf{v}_k$ $k=1$ .. r, the *right singular vectors*.

See Wilks 2005, Sec. 9.3.5

Appendix A



- Singular vectors $\{\mathbf{u}_k\}$ and $\{\mathbf{v}_k\}$ (k=1..r) constitue orthonormal coordinate system in r-dim subspace of n-dim and m-dim phase-spaces (i.e. not necessarily complete)

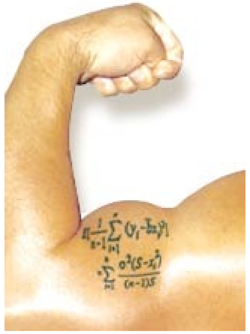
- Sum of squared singular values

$$\|Q\|_F^2 := \sum_i^n \sum_j^m q_{ij}^2 = \sum_i^r \omega_i^2$$

Frobenius norm = sum of squared matrix elements
= sum of squared singular values

see e.g. Bretherton et al. 1992

Appendix A



- **Calculation of SVD**
  - Singular vectors $\{\mathbf{u}_k\}$ are eigenvectors of QQ^T (symm, nxn) and $\{\mathbf{v}_k\}$ of Q^TQ (symm, mxm)

$$\left(\mathbf{Q}\cdot\mathbf{Q}^T\right)\cdot\mathbf{L} = \mathbf{U}\Omega\mathbf{V}^T\cdot\mathbf{V}\Omega^T\mathbf{U}^T\mathbf{U} = \mathbf{L}\Omega^2, \qquad \text{dito for } \mathbf{V}$$

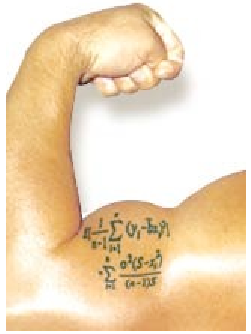
  - Singular left vectors can be obtained directly from singular right vectors (and vice versa):

$$\mathbf{QV} = \mathbf{U}\Omega, \quad \mathbf{Q}^T\mathbf{U} = \mathbf{V}\Omega^T$$

$$Q\mathbf{v}_k = \omega_k\mathbf{u}_k, \quad Q^T\mathbf{u}_k = \omega_k\mathbf{v}_k, \quad k = 1,..,r$$

  - In practice: dedicated software to svd

Appendix B



- **Transformation of Cross-Covariance Matrix**
  - If $\mathbf{X}, \mathbf{Y}$ are centered data matrices, the cross-covariance matrix is:

$$\mathbf{S}_{xy} = \frac{1}{N-1}\mathbf{X}^T \cdot \mathbf{Y}$$

  - Let $\mathbf{U}$, $\mathbf{V}$ be transformation matrices (with projection vectors in columns), the data matrices in transformed coordinates are:

$$\mathbf{A} = \mathbf{X} \cdot \mathbf{U}, \qquad \mathbf{B} = \mathbf{Y} \cdot \mathbf{V}$$

  - The cross-covariance matrix ov transformed variables is:

$$\mathbf{S}_{ab} = \frac{1}{N-1}\mathbf{A}^T \cdot \mathbf{B} = \frac{1}{N-1}\mathbf{U}^T \mathbf{X}^T \cdot \mathbf{Y}\mathbf{V} = \mathbf{U}^T \mathbf{S}_{xy} \mathbf{V}$$

  - When $\mathbf{U}, \mathbf{V}$ are singular vectors systems of $\mathbf{S}_{xy}$ then:

$$\begin{array}{c} \mathbf{S}_{xy} = \mathbf{U}\Omega\mathbf{V}^T \\ \mathbf{U}^T \cdot \mathbf{U} = \mathbf{1}, \mathbf{V}^T \cdot \mathbf{V} = \mathbf{1} \end{array} \quad \Rightarrow \quad \mathbf{S}_{ab} = \mathbf{U}^T \mathbf{S}_{xy} \mathbf{V} = \mathbf{U}^T \mathbf{U}\Omega\mathbf{V}^T \mathbf{V} = \Omega$$