Coupled field analysis

1 Required packages and plot function

```r
library(sinkr) # https://github.com/marchtaylor/sinkr
library(vegan)
library(parallel)
library(calibrate)
library(akima)
library(maps)
library(pals)
library(CCP)
library(knitr)

coupledModes <- function(sx, sy, tx, ty,
  sx_zlim = max(abs(sx$z), na.rm = T)*c(-1,1),
  sy_zlim = max(abs(sy$z), na.rm = T)*c(-1,1),
  tlim = range(c(tx,ty)),
  scale_tx = F, scale_ty = F,
  sx_col = pals::ocean.balance(21), sy_col = pals::ocean.balance(21),
  tx_col = 3, ty_col = 4,
  widths = c(2,2), heights = c(1.5,1), respect = F,
  mar = c(3,3,2,1), mgp = c(2,0.5,0), ps = 10,
  sx_title = "sx", sy_title = "sy",
  t_title = ""){

  op <- par(mar = mar, mgp = mgp, ps = ps)
  layout(matrix(c(1,2,3,3), 2, 2, byrow = T),
    widths = widths, heights = heights, respect = respect)
  # sx
  image(sx, zlim = sx_zlim, col = pals::ocean.balance(21), xlab = "", ylab = "")
  contour(sx, add = T, col = 8)
  maps::map("world", add = T, fill = F, boundary = T, col = 1, lwd = 1)
  mtext(sx_title, side = 3, line = 0.5, col = tx_col)
  # sy
  image(sy, zlim = sy_zlim, col = pals::ocean.balance(21), xlab = "", ylab = "")
  contour(sy, add = T, col = 8)
  maps::map("world", add = T, fill = F, boundary = T, col = 1, lwd = 1)
  mtext(sy_title, side = 3, line = 0.5, col = ty_col)
  # tx & ty
  if(scale_tx){ tx$y <- scale(tx$y) }
  if(scale_ty){ ty$y <- scale(ty$y) }
  ylim <- range(c(tx$y, ty$y), na.rm = T)
  plot(tx, col = tx_col, t = "n", xlim = tlim, ylim = ylim, xlab = "", ylab = "")
  abline(h = 0, col = 1, lty = 3)
  lines(tx, col = tx_col, t = "l")
  lines(ty, col = ty_col, t = "l")
  mtext(t_title, side = 3, line = 0.5)
  par(op)
}
```

2 Introduction

The following methods are focused on the identification of coupled spatiotemporal patterns between two data fields. As an example, we will look at the patterns between monthly anomalies of sea level pressure (SLP) and sea surface temperature (SST) in the equatorial Pacific (180°W - 70°W, 30°S - 30°N). Both datasets are available in the *sinkr* package and include a matrix of values (`data$field`; rows = time, cols = space), with further information on the two dimensions (`data$date`, `data$grid$lon`, `data$grid$lat`).

Three multivariate methods are demonstrated: 1 Maximum Covariance Analysis (MCA), 2.Canonical Correlation Analysis (CCA), and 3. Redundancy Analysis (RDA). There are subtle differences between them, but the main distinction is that RDA tries to explain the variance in one field (*constrained data*; i.e. response) using the other (*constraining data*, i.e. explanatory) while CCA and MCA identifies coupled patterns without defining a response and explanatory field.

2.1 Data

```r
# load
data(slp) # Hadley SLP monthly mean dataset
data(sst) # Kaplan monthly mean anomaly dataset

# make anomaly for slp
slp$field <- fieldAnomaly(y = slp$field, x = slp$date, level = "month")

# data matrices
X <- slp$field
Y <- sst$field

# filter land (relevant for sst)
x_grd_incl <- which(!is.na(colSums(X)))
y_grd_incl <- which(!is.na(colSums(Y)))

# intersecting dates (relevant for coupled field methods)
date_intersect <- intersect(x = slp$date, y = sst$date)
x_date_incl <- which(slp$date %in% date_intersect)
y_date_incl <- which(sst$date %in% date_intersect)
```

# 3 Empirical Orthogonal Function (EOF) analysis

Empirical Orthogonal Function (EOF) analysis (or Principal Component Analysis, PCA) is a commonly used method for describing dominant modes of spatiotemporal variability. I will refer to these results as the EOFs (spatial patterns) and corresponding Principal Components (PCs) (temporal patterns). There is no guarantee that there is a true physical meaning to the patterns, so care should be taken in their interpretation. *Rotation* of EOFs is an approach that may help to focus on more physically-based patterns, but one loses the other properties of the EOF (i.e. orthogonality of the PCs).

EOF is used as a precursor (or analogue) in some of the following coupled field methods. It may be used as a type of filter to remove small-scale features from the field and focus the analysis more on the large-scale patterns. For this reason, we are less concerned with interpreting the individual EOFs, and will not be interpreting their specific physical meaning. Below in an example of an EOF applied to SST anomalies, and a plot of the resulting leading ($n = 3$) patterns (i.e. *standing oscillations*).

```
 # centering is necessary, scaling is optional
PY <- prcomp(Y[,y_grd_incl], center = T, scale. = F)
explVar <- PY$sdev^2/sum(PY$sdev^2)

layout(matrix(1:6, nrow = 3, ncol = 2, byrow = T),
  widths = c(1,2), heights = c(1,1), respect = F)
par(mar = c(2.5, 2.5, 2, 1), mgp = c(2, 0.5, 0))
for(i in seq(3)){
  modeNum <- i
  sy <- interp(x = sst$grid$lon[y_grd_incl], y = sst$grid$lat[y_grd_incl],
    z = PY$rotation[,modeNum])
  ty <- data.frame(x = sst$date, y = PY$x[,modeNum])
    sy_zlim = max(abs(sy$z), na.rm = T)*c(-1,1)
  sy_zlim = max(abs(sy$z), na.rm = T)*c(-1,1)
  sy_col = pals::ocean.balance(21)
  # sx
  image(sy, zlim = sy_zlim, col = pals::ocean.balance(21), xlab = "", ylab = "")
  contour(sy, add = T, col = 8)
  maps::map("world", add = T, fill = F, boundary = T, col = 1, lwd = 1)
  mtext(text = paste0("EOF", i), side = 3, line = 0.5)
  # tx
  plot(ty, t = "l", xlim = c(as.Date("1940-01-01"), as.Date("2000-01-01")),
    xlab = "", ylab = "")
  abline(h = 0, lty = 3, col = 8)
  mtext(text = paste0("PC", i, "; Expl.var. = ",
    round(explVar[i]*100, 1), "%"), side = 3, line = 0.5)
}
```
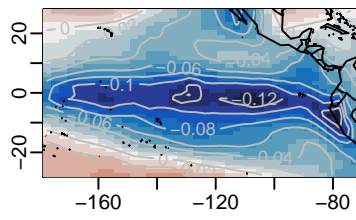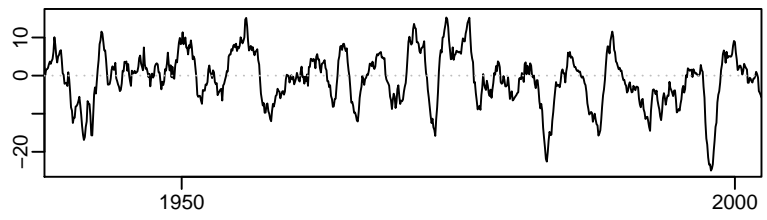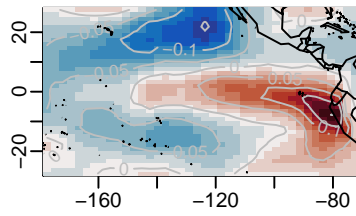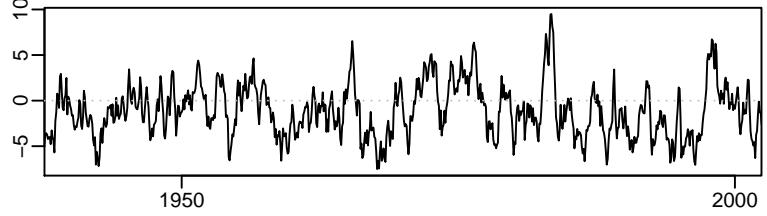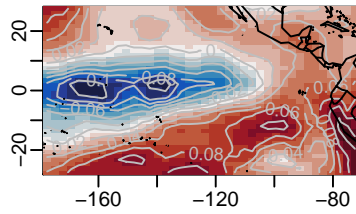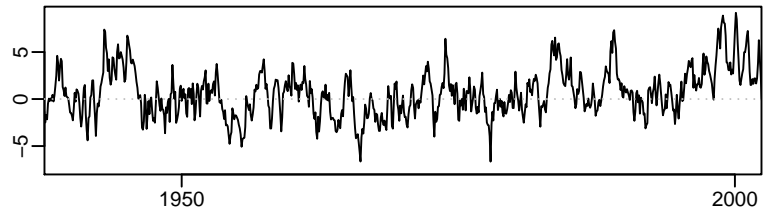
4   Maximum Covariance Analysis (MCA/SVD)

This is a very simple method which is quite related to EOF. Is is basically the decomposition of a covariance matrix calculated across the two fields (e.g. between all time series, *T-mode*). A covariance matrix of a single field will be square, and can be decomposed with an Eigen decomposition. In a coupled field covariance matrix, the number of grids can be different (but the length of the time series are the same), resulting an a non-square covariance matrix. To decompose this, we can use Single Value Decomposition (SVD), and this is another name sometimes used for MCA. This results in the spatial patterns (*modes*), and the corresponding temporal modes can be derived by projecting the original data onto these. For a more detailed description, see Bjornsson and Venegas (1997).

```
# Both fields are first centered
Xsc <- scale(X[x_date_incl, x_grd_incl], center = T, scale = F)
Ysc <- scale(Y[y_date_incl, y_grd_incl], center = T, scale = F)

# compute covariance matrix and decompose via Singular Value Decomposition (SVD)
COV <- cov(x = Xsc, y = Ysc)
PXY <- svd(x = COV)

# squared covariance fraction
scfXY <- (PXY$d)/sum(PXY$d)

# expansion of principal components (temporal modes)
A <- Xsc %*% PXY$u
B <- Ysc %*% PXY$v
```
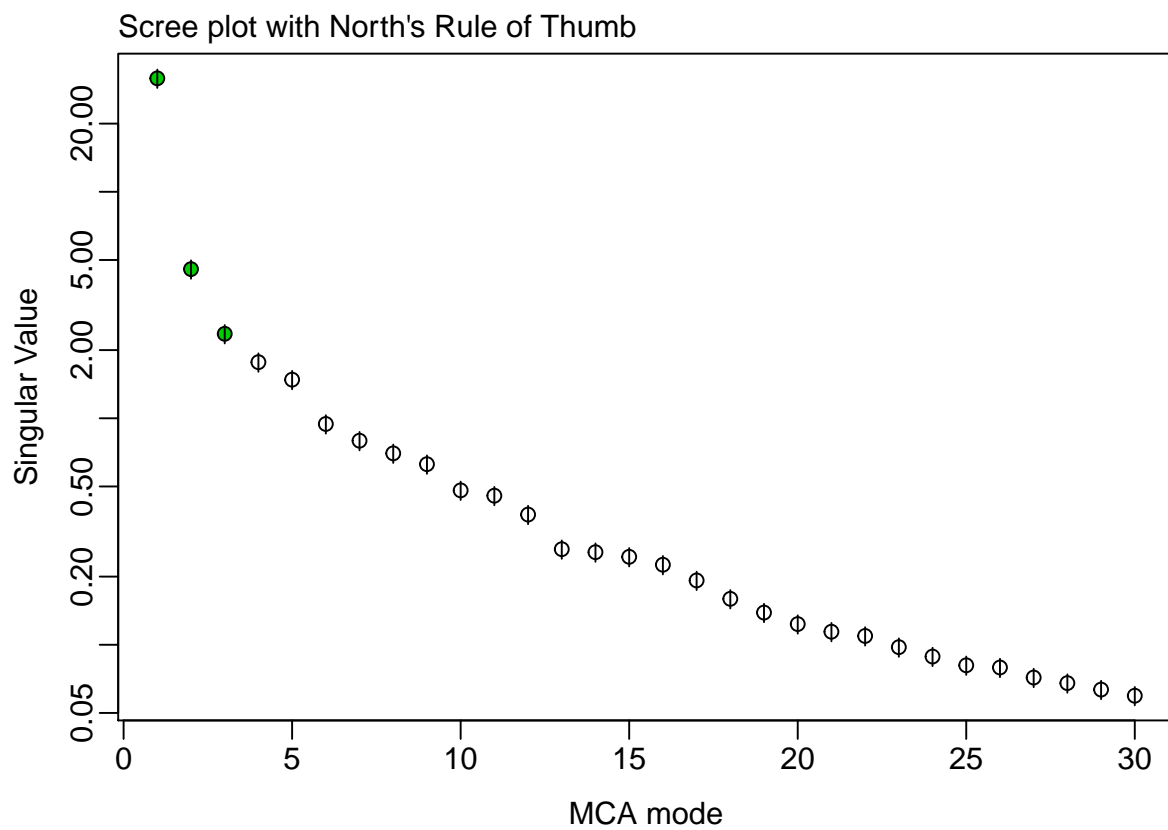
4.1   MCA mode significance

There are many approaches to determining *significant* EOF modes, and these are probably equally applicable to MCA. Bjornsson and Venegas (1997) outline several of these. One of the most straightforward approaches is a *Scree plot*, which examines the Eigen or Singular values in a plot, and tries to identify when their amplitude decreases more-or-less linearly (in log scale). Another approach is *North's Rule of Thumb*, which identifies where associated errors are no-longer overlapping neighboring values. These are not tests of statistical significance, but rather attempt to identify modes that describe large-scale spatiotemporal features above a given level of background noise.

```
Lambda_err <- sqrt(2/min(dim(COV)))*PXY$d
upper.lim <- PXY$d+Lambda_err
lower.lim <- PXY$d-Lambda_err
NORTHok <- 0*PXY$d
for(i in seq(PXY$d)){
  Lambdas <- PXY$d
  Lambdas[i] <- NaN
  nearest <- which.min(abs(PXY$d[i]-Lambdas))
  if(nearest > i){if(lower.lim[i] > upper.lim[nearest]) NORTHok[i] <- 1}
  if(nearest < i){if(upper.lim[i] < lower.lim[nearest]) NORTHok[i] <- 1}
}
n_sig <- min(which(NORTHok==0))-1

# Scree plot
COL <- rep(NaN, 30); COL[seq(n_sig)] <- 3
par(mar = c(3,3,1.5,1), mgp = c(2,0.5,0))
plot(PXY$d[1:30], log = "y", xlab = "MCA mode", ylab = "Singular Value",
  pch = 21, bg = COL)
segments(x0 = 1:30, x1 = 1:30, y0 = lower.lim[1:30], y1 = upper.lim[1:30])
mtext("Scree plot with North's Rule of Thumb", line = 0.25, side = 3, adj = 0)
```

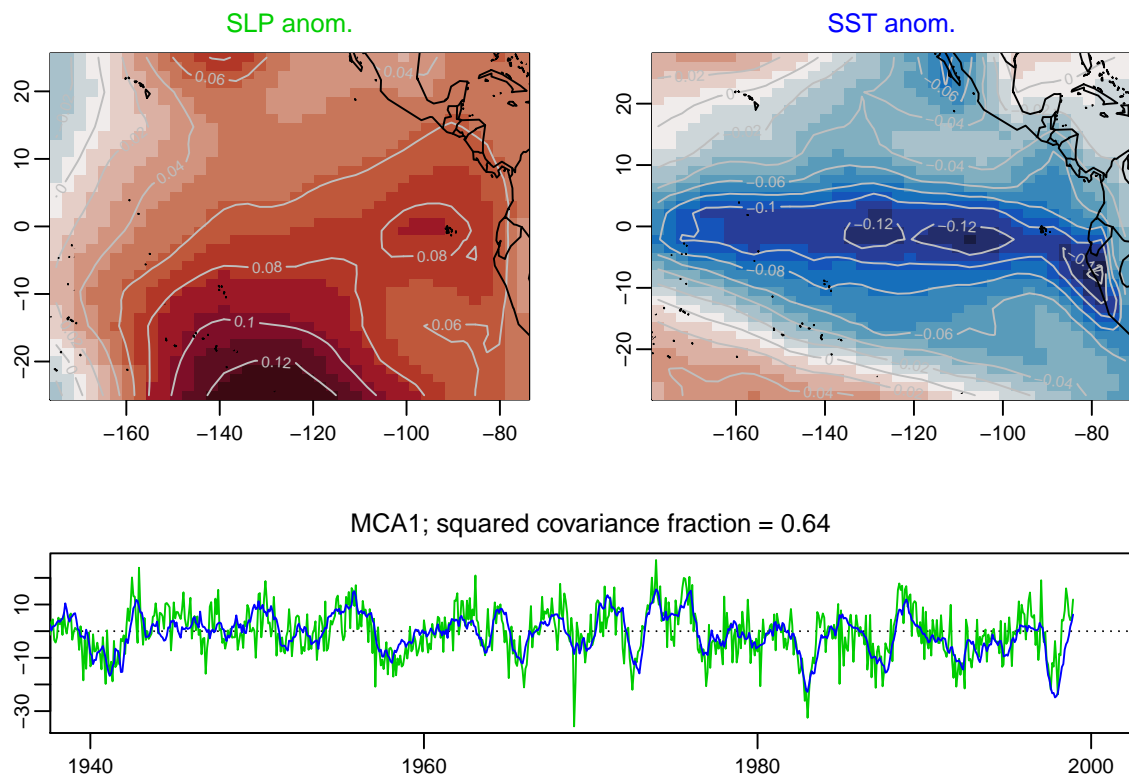Scree plot with North's Rule of Thumb

4.2 MCA mode 1

```
modeNum <- 1
sx <- interp(x = slp$grid$lon[x_grd_incl], y = slp$grid$lat[x_grd_incl],
  z = PXY$u[,modeNum])
sy <- interp(x = sst$grid$lon[y_grd_incl], y = sst$grid$lat[y_grd_incl],
  z = PXY$v[,modeNum])
tx <- data.frame(x = slp$date[x_date_incl], y = A[,modeNum])
ty <- data.frame(x = sst$date[y_date_incl], y = B[,modeNum])

coupledModes(sx, sy, tx, ty, sx_title = "SLP anom.", sy_title = "SST anom.",
  tlim = c(as.Date("1940-01-01"), as.Date("2000-01-01")),
  t_title = paste0("MCA", modeNum, "; ",
    "squared covariance fraction = ", round(scfXY[modeNum], 2)))
```


7

4.3 MCA mode 2

```
modeNum <- 2
sx <- interp(x = slp$grid$lon[x_grd_incl], y = slp$grid$lat[x_grd_incl],
  z = PXY$u[,modeNum])
sy <- interp(x = sst$grid$lon[y_grd_incl], y = sst$grid$lat[y_grd_incl],
  z = PXY$v[,modeNum])
tx <- data.frame(x = slp$date[x_date_incl], y = A[,modeNum])
ty <- data.frame(x = sst$date[y_date_incl], y = B[,modeNum])

coupledModes(sx, sy, tx, ty, sx_title = "SLP anom.", sy_title = "SST anom.",
  tlim = c(as.Date("1940-01-01"), as.Date("2000-01-01")),
  t_title = paste0("MCA", modeNum, "; ",
    "squared covariance fraction = ", round(scfXY[modeNum], 2)))
```
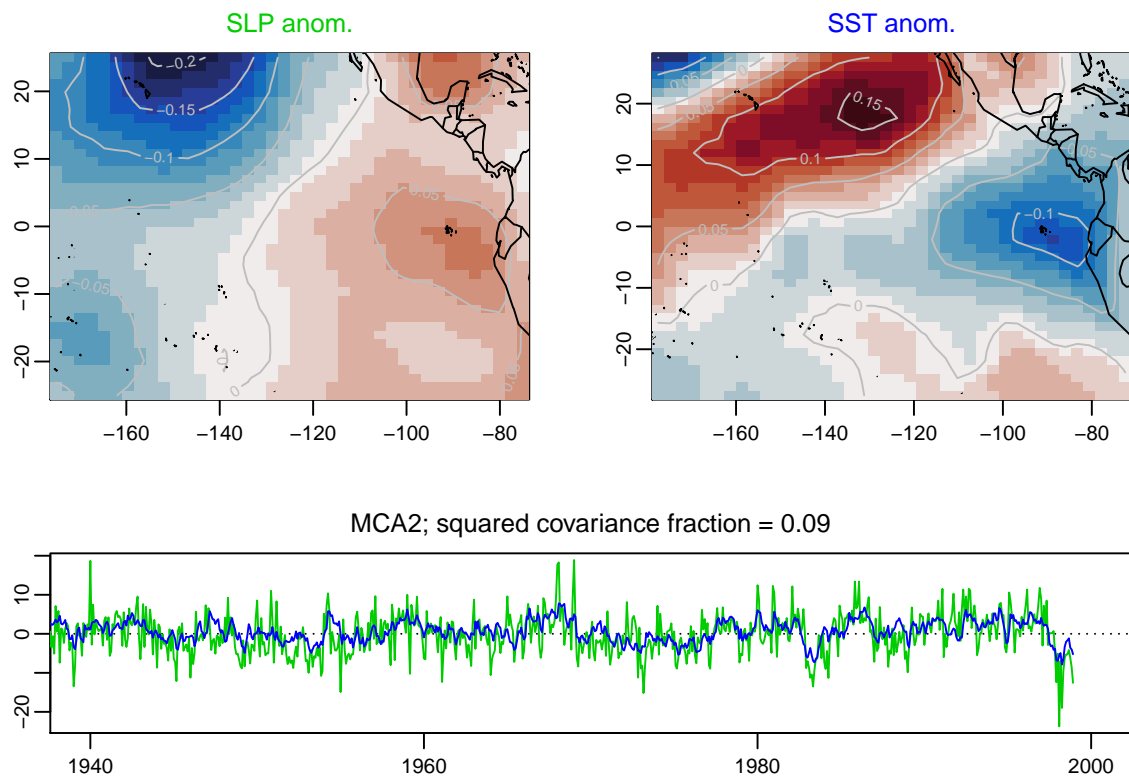
5 Canonical Correlation Analysis (CCA)

CCA is for identifying correlated spatiotemporal patterns (i.e. no predictor and predictand). As opposed to using the full data fields, a common approach is to do a pre-filtering of the data using EOF, often referred to as the Barnett and Preisendorfer (Barnett and Preisendorfer 1987) approach (i.e. *BPCCA*), and the temporal EOF modes (*Principle Components*) are supplied to the operation. This approach has the advantage of removing noise held by trailing EOFs in order to focus on the main patterns. Furthermore, the reduced datasets allow for a more computationally efficient solution. The level of truncation to use for the PCs is subjective, but one approach may be to only include those EOFs that explain up to a given cumulative explained variance, or a threshold of explained variance. Bretherton, Smith, and Wallace (1992) have shown BPCCA and MCA to be among the methods that produced the most robust results in identifying coupled modes.

```
# EOF
PX <- prcomp(x = X[, x_grd_incl], center = T, scale. = F)
PY <- prcomp(x = Y[, y_grd_incl], center = T, scale. = F)
explVarX <- (PX$sdev^2)/sum(PX$sdev^2)
cumExplVarX <- cumsum(explVarX)
explVarY <- (PY$sdev^2)/sum(PY$sdev^2)
cumExplVarY <- cumsum(explVarY)

# truncated PCs
thresh <- 0.01 # expl.var threshold for PC inclusion into RDA
x_pc_incl <- which(explVarX > thresh)
y_pc_incl <- which(explVarY > thresh)
Xx <- as.matrix(PX$x[x_date_incl, x_pc_incl])
Yx <- as.matrix(PY$x[y_date_incl, y_pc_incl])

# run CCA
CCA <- calibrate::canocor(X = Xx, Y = Yx)

# project spatial mode loadings
x_spatial <- PX$rotation[,x_pc_incl] %*% CCA$A
y_spatial <- PY$rotation[,y_pc_incl] %*% CCA$B
```

5.1 CCA mode significance

```
# tests of canonical dimensions
rho <- diag(CCA$ccor)

# Define number of observations, number of variables in first set, and number of variables in the se
n <- dim(CCA$U)[1]
p <- dim(CCA$U)[2]
q <- dim(CCA$V)[2]

# Calculate p-values using the F-approximations of different test statistics:
p.asym(rho, n, p, q, tstat = "Pillai")
```

```
 Pillai-Bartlett Trace, using F-approximation:
                stat      approx df1   df2       p.value
1 to 11:    1.434545639 20.777871 121 16764 0.000000e+00
2 to 11:    0.910878271 15.154939 100 16786 0.000000e+00
3 to 11:    0.641019037 12.840588  81 16808 0.000000e+00
4 to 11:    0.449338594 11.199488  64 16830 0.000000e+00
5 to 11:    0.285822499  9.174723  49 16852 0.000000e+00
```

```
 6 to 11:   0.184790350  8.008661  36 16874 0.000000e+00
 7 to 11:   0.119792473  7.441085  25 16896 0.000000e+00
 8 to 11:   0.057005820  5.508230  16 16918 5.985323e-12
 9 to 11:   0.029106474  4.993655   9 16940 9.658169e-07
10 to 11:   0.011177665  4.313373   4 16962 1.731569e-03
11 to 11:   0.002074256  3.203255   1 16984 7.350969e-02
```
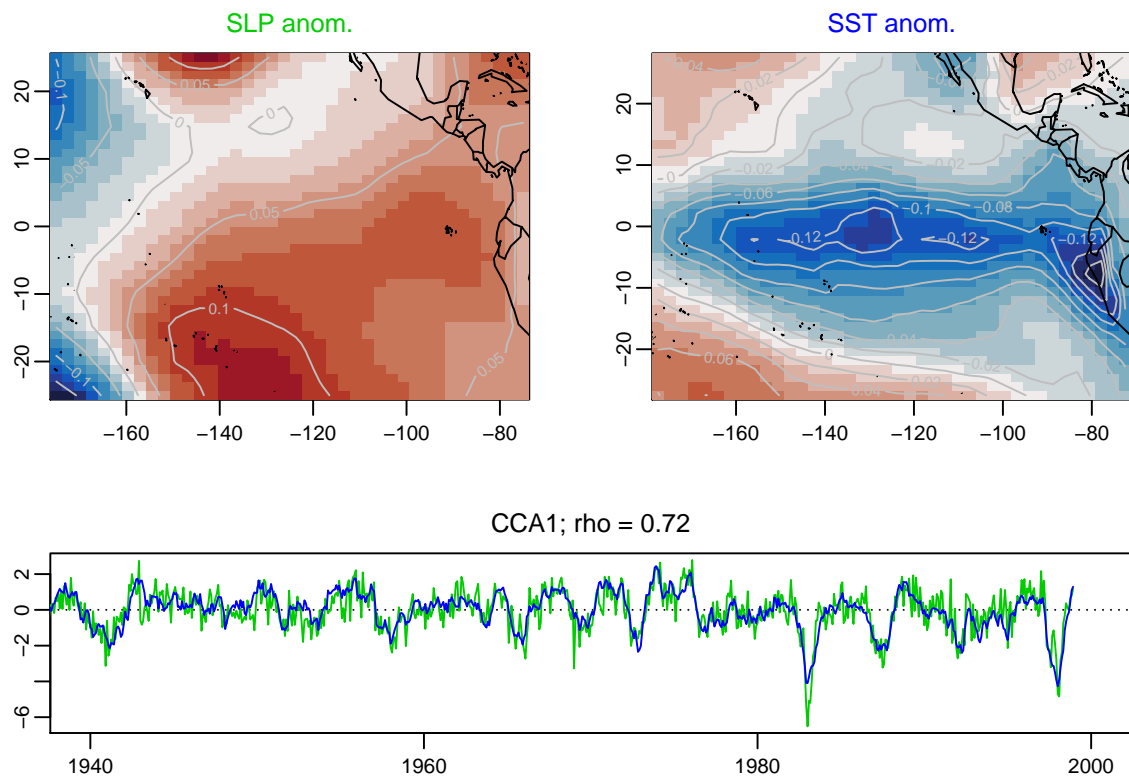
5.2 CCA mode 1

```
modeNum <- 1
sx <- interp(x = slp$grid$lon[x_grd_incl], y = slp$grid$lat[x_grd_incl],
  z = x_spatial[,modeNum])
sy <- interp(x = sst$grid$lon[y_grd_incl], y = sst$grid$lat[y_grd_incl],
  z = y_spatial[,modeNum])
tx <- data.frame(x = slp$date[x_date_incl], y = CCA$U[,modeNum])
ty <- data.frame(x = sst$date[y_date_incl], y = CCA$V[,modeNum])

coupledModes(sx, sy, tx, ty, sx_title = "SLP anom.", sy_title = "SST anom.",
  tlim = c(as.Date("1940-01-01"), as.Date("2000-01-01")),
  t_title = paste0("CCA", modeNum, "; ", "rho = ",
    round(diag(CCA$ccor)[modeNum], 2)))
```

5.3 CCA mode 2

```r
modeNum <- 2
sx <- interp(x = slp$grid$lon[x_grd_incl], y = slp$grid$lat[x_grd_incl],
  z = x_spatial[,modeNum])
sy <- interp(x = sst$grid$lon[y_grd_incl], y = sst$grid$lat[y_grd_incl],
  z = y_spatial[,modeNum])
tx <- data.frame(x = slp$date[x_date_incl], y = CCA$U[,modeNum])
ty <- data.frame(x = sst$date[y_date_incl], y = CCA$V[,modeNum])

coupledModes(sx, sy, tx, ty, sx_title = "SLP anom.", sy_title = "SST anom.",
  tlim = c(as.Date("1940-01-01"), as.Date("2000-01-01")),
    t_title = paste0("CCA", modeNum, "; ", "rho = ",
    round(diag(CCA$ccor)[modeNum], 2)))
```
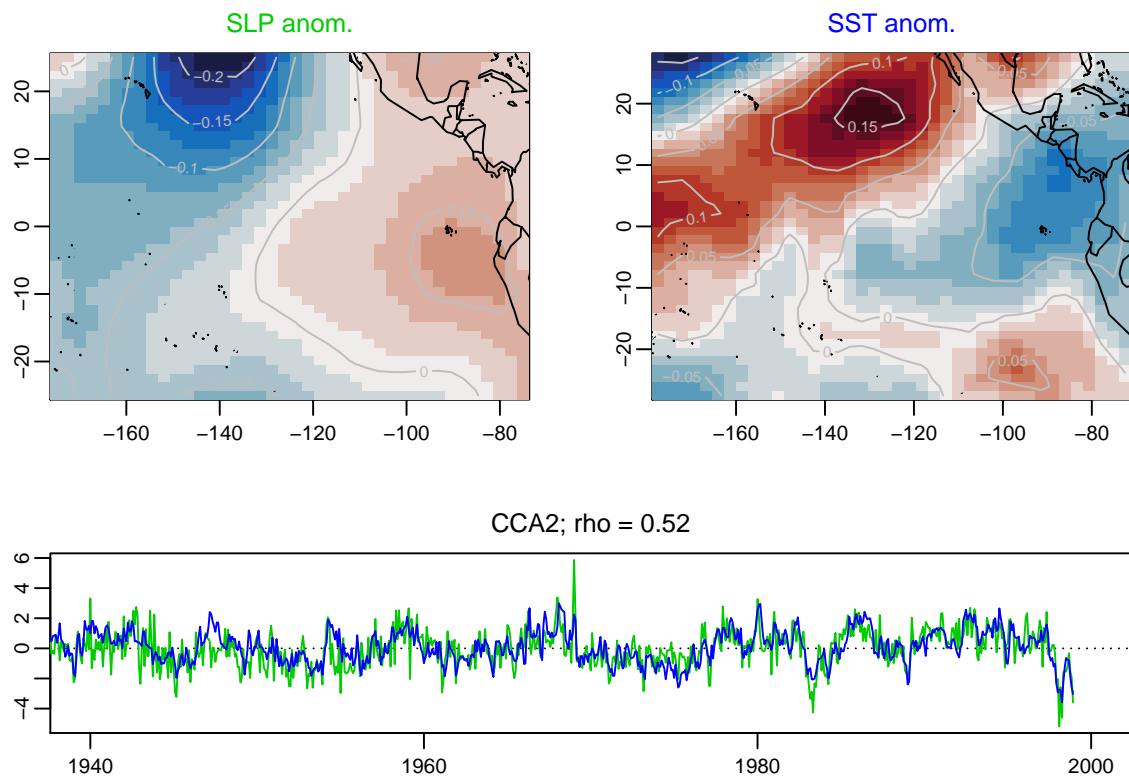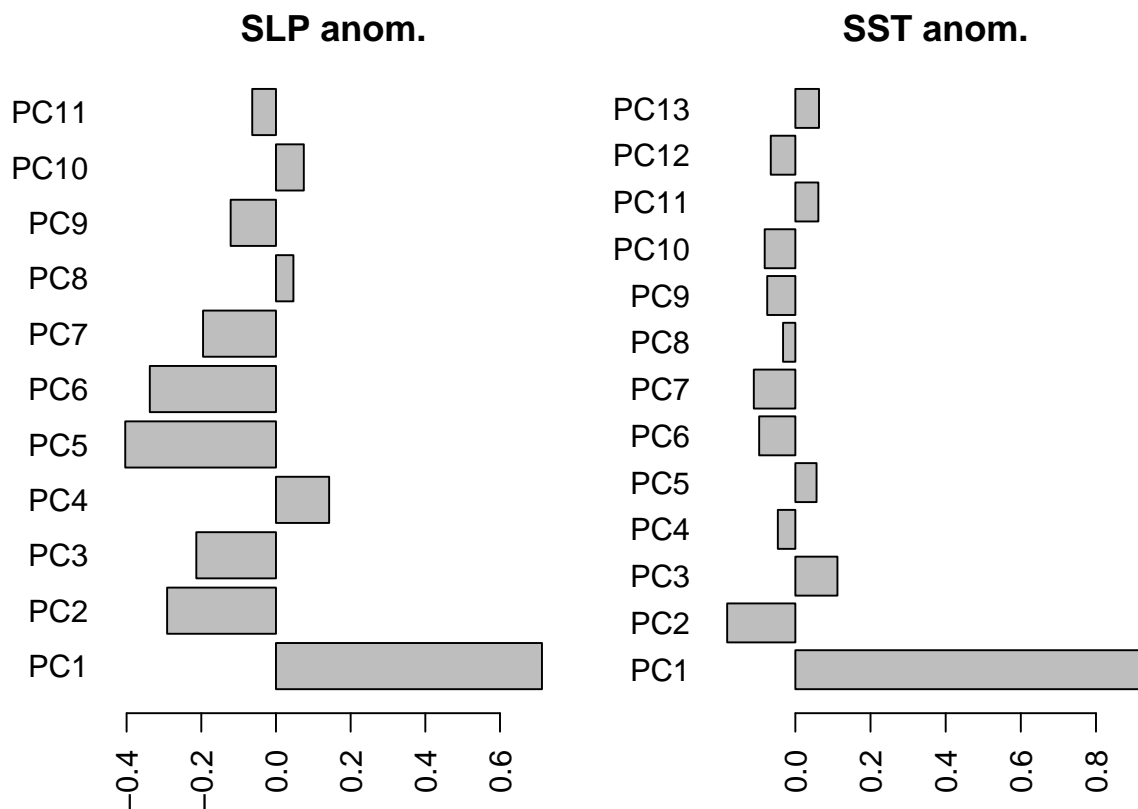
5.4 PC coefficients in CCA1

```
par(mfcol = c(1,2), mar = c(3,4,2,1))
barplot(height = CCA$A[,1], names.arg = paste0("PC", seq(CCA$A[,1])),
  horiz = T, las = 2, main = "SLP anom.")
barplot(height = CCA$B[,1], names.arg = paste0("PC", seq(CCA$B[,1])),
  horiz = T, las = 2, main = "SST anom.")
```

6 Redundancy Analysis (RDA)

The use of RDA in coupled field analysis is less common than CCA, but may be an attractive option when the focus is on maximizing the explained variance of the response dataset, or when there is clearer evidence of a causal relationship. A mathematical basis can be found in the following sources: Rao (1964), Tyler (1982), Hans von Storch and Zwiers (1999). F. Kauker and Meier (2003) and Frank Kauker et al. (2008) provide a similar application (with accompanying appendices on RDA), to that shown here. The approach is similar to the CCA above, whereby EOF is first used to filter out a smaller subset of PCs, which are subsequently used as input to the analysis. The `vegan` package is used, which has become on of the standard R packages for multivariate community analysis.

```r
RDA <- vegan::rda(Yx ~ Xx, scale = F)


S <- summary(RDA, axes = 5)
explVar <- RDA$CCA$eig / sum(S$constr.chi + S$unconst.chi)

## extract modes - variable (PC) and time component scores
# orthonormal environmental scores (columns of X; i.e. PC loadings)
RDA$A <- S$biplot
# orthonormal species scores (columns of Y; i.e. PC loadings)
RDA$B <- S$species
# orthonormal site scores (rows of X; i.e. time loadings)
RDA$V <- S$constraints
# orthonormal site scores (rows of Y; i.e. time loadings)
RDA$W <- S$sites


# spatial loadings (project EOFs onto RDA variable loadings)
x_spatial <- PX$rotation[, x_pc_incl] %*% RDA$A
y_spatial <- PY$rotation[, y_pc_incl] %*% RDA$B
```

6.1 RDA mode significance

The `vegan` package includes a sophisticated set of permutation tests that can be used to test various aspects of model significance. One can also test the significance of the RDA modes (i.e. axes). The following routine permutes the data for a prescribed number of times and evaluates whether the explained variance is above a random noise threshold.

```r
# ANOVA-like test
clus <- parallel::makePSOCKcluster(min(3, parallel::detectCores()-1))
set.seed(1111)
rdaAx <- anova(RDA, by = "axis", permutations = 199, parallel = clus, cutoff = 0.05)
parallel::stopCluster(clus)
rdaAx
```

```
Permutation test for rda under reduced model
Forward tests for axes
Permutation: free
Number of permutations: 199

Model: rda(formula = Yx ~ Xx, scale = F)
         Df Variance       F Pr(>F)
RDA1      1   22.612 629.3490  0.005 **
RDA2      1    1.769  49.2338  0.005 **
RDA3      1    0.455  12.6498  0.005 **
RDA4      1    0.391  10.8689  0.005 **
RDA5      1    0.207   5.7750  0.105
```

```
RDA6        1   0.122   3.3857
RDA7        1   0.098   2.7229
RDA8        1   0.055   1.5401
RDA9        1   0.036   0.9970
RDA10       1   0.017   0.4707
RDA11       1   0.005   0.1445
Residual 1524  54.757
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```
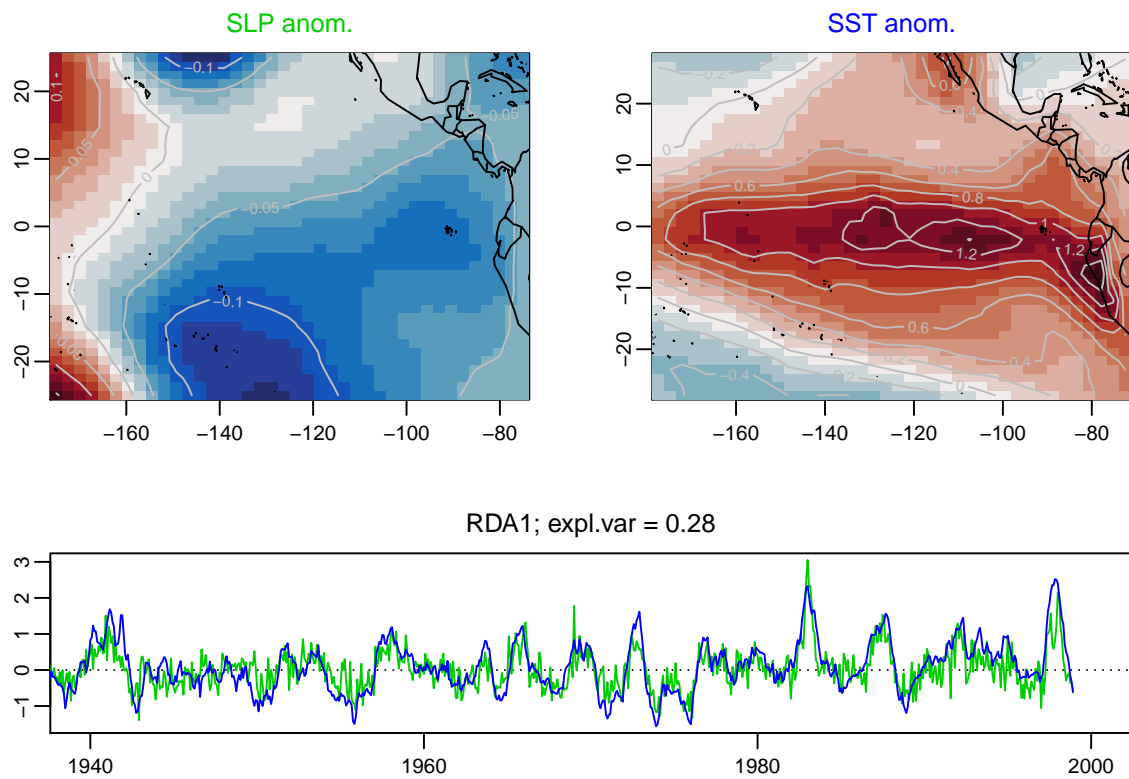
6.2 RDA mode 1

```
modeNum <- 1
sx <- interp(x = slp$grid$lon[x_grd_incl], y = slp$grid$lat[x_grd_incl],
  z = x_spatial[,modeNum])
sy <- interp(x = sst$grid$lon[y_grd_incl], y = sst$grid$lat[y_grd_incl],
  z = y_spatial[,modeNum])
tx <- data.frame(x = slp$date[x_date_incl], y = RDA$V[,modeNum])
ty <- data.frame(x = sst$date[y_date_incl], y = RDA$W[,modeNum])

coupledModes(sx, sy, tx, ty, sx_title = "SLP anom.", sy_title = "SST anom.",
  tlim = c(as.Date("1940-01-01"), as.Date("2000-01-01")),
  t_title = paste0("RDA", modeNum, "; ",
    "expl.var = ", round(explVar[modeNum], 2)))
```

6.3 RDA mode 2

```r
modeNum <- 2
sx <- interp(x = slp$grid$lon[x_grd_incl], y = slp$grid$lat[x_grd_incl],
  z = x_spatial[,modeNum])
sy <- interp(x = sst$grid$lon[y_grd_incl], y = sst$grid$lat[y_grd_incl],
  z = y_spatial[,modeNum])
tx <- data.frame(x = slp$date[x_date_incl], y = RDA$V[,modeNum])
ty <- data.frame(x = sst$date[y_date_incl], y = RDA$W[,modeNum])

coupledModes(sx, sy, tx, ty, sx_title = "SLP anom.", sy_title = "SST anom.",
  tlim = c(as.Date("1940-01-01"), as.Date("2000-01-01")),
  scale_tx = T, scale_ty = T,
  t_title = paste0("RDA", modeNum, "; ",
    "expl.var = ", round(explVar[modeNum], 2)))
```
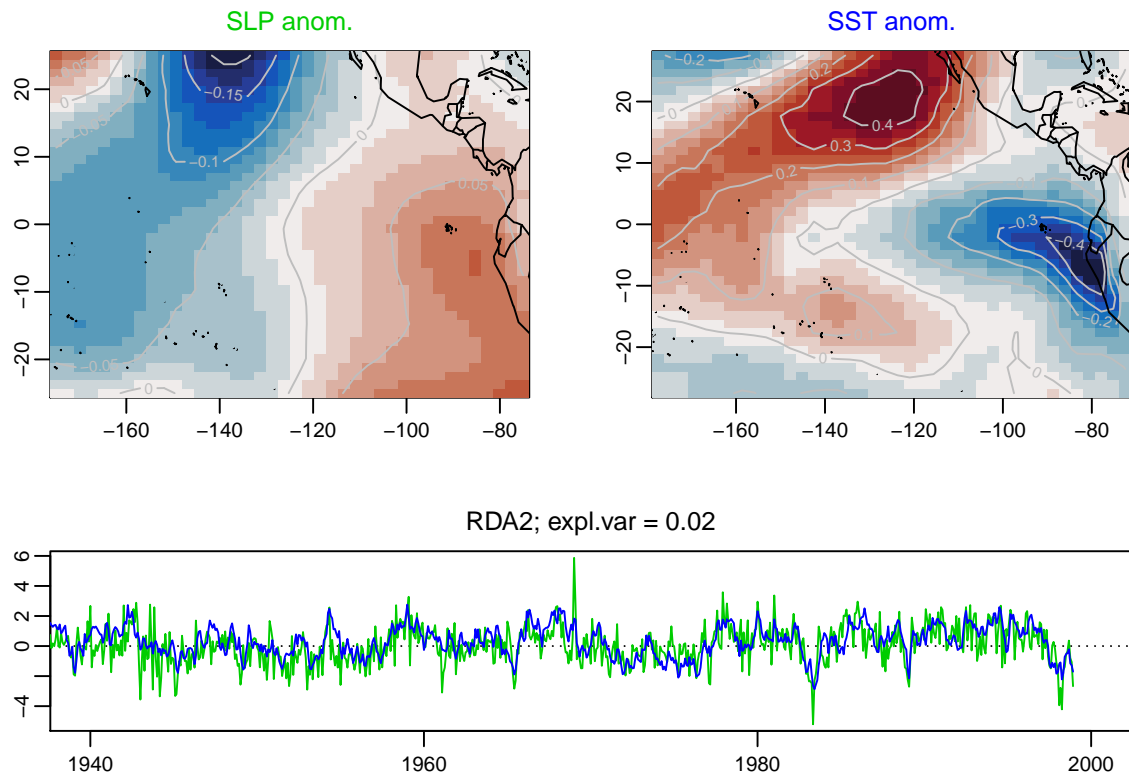
7 Further topics

- Lags - All three methods could technically evaluate whether a mode's importance increases (i.e. in terms of explained variance or correlation coefficient) when lags are used between the fields.
- Prediction - CCA and RDA have both been used to predict the state of one field when the other is known, by using the model fit to historical data. My experience with CCA is that the predictive power of the model is not very sensitive to the EOF truncation level, as unimportant PCs are downweighted by the model and contribute little to the prediction (Taylor et al. 2013). Truncation level may be more relevant to RDA, and may specifically influence the overall explained variance of the model, even if the leading modes remain largely unchanged.
- Gappy data - If missing data is interspersed in the data field, interpolation methods may be required as an initial step before proceeding to the model fitting; e.g. DINEOF (Data Interpolating Empirical Orthogonal Functions) (`sinkr::dineof()`).
- The methods shown here deal with standing oscillations, not propagating patterns. For those cases, there are related methods; e.g. Principal Oscillation Patterns (POPs) and Principal Interaction Patterns (POPs) (see H. von Storch and Navarra 2013)

8 Software Versions

- R version 3.6.3 (2020-02-29)
- sinkr: 0.6
- vegan: 2.5.6
- parallel: 3.6.3
- calibrate: 1.7.7
- akima: 0.6.2.1
- maps: 3.3.0
- pals: 1.6
- CCP: 1.1
- knitr: 1.29
- **Compiled**: 2021-Mar-14

9 Author

Marc Taylor. Thünen Institute of Sea Fisheries, Marine Living Resources Unit, Herwigstraße 31, 27572 Bremerhaven, Germany. https://www.thuenen.de/en/sf/

References

Barnett, T. P., and R. Preisendorfer. 1987. "Origins and Levels of Monthly and Seasonal Forecast Skill for United States Surface Air Temperatures Determined by Canonical Correlation Analysis." *Monthly Weather Review* 115 (9): 1825–50. https://doi.org/10.1175/1520-0493(1987)115%3C1825:OALOMA%3E2.0.CO;2.

Bjornsson, Halldor, and Silvia A Venegas. 1997. "A Manual for EOF and SVD Analyses of Climate Data." *McGill University, CCGCR Report*, no. 97-1: 52.

Bretherton, Christopher S., Catherine Smith, and John M. Wallace. 1992. "An Intercomparison of Methods for Finding Coupled Patterns in Climate Data." *Journal of Climate* 5 (6): 541–60. https://doi.org/10.1175/1520-0442(1992)005%3C0541:AIOMFF%3E2.0.CO;2.

Kauker, F., and H. E. M. Meier. 2003. "Modeling Decadal Variability of the Baltic Sea: 1. Reconstructing Atmospheric Surface Data for the Period 1902–1998." *Journal of Geophysical Research: Oceans* 108 (C8). https://doi.org/https://doi.org/10.1029/2003JC001797.

Kauker, Frank, Cornelia Köberle, Rüdiger Gerdes, and Michael Karcher. 2008. "Modeling the 20th Century Arctic Ocean/Sea Ice System: Reconstruction of Surface Forcing." *J. Geophys. Res.* 113 (C9): C09027. https://doi.org/10.1029/2006JC004023.

Rao, C Radhakrishna. 1964. "The Use and Interpretation of Principal Component Analysis in Applied Research." *Sankhyā: The Indian Journal of Statistics, Series A*, 329–58.

Storch, H. von, and A. Navarra. 2013. *Analysis of Climate Variability: Applications of Statistical Techniques Proceedings of an Autumn School Organized by the Commission of the European Community on Elba from October 30 to November 6, 1993.* Springer Berlin Heidelberg. https://books.google.de/books?id=aAb9CAAAQBAJ.

Storch, Hans von, and Francis W. Zwiers. 1999. *Statistical Analysis in Climate Research.* Cambridge: Cambridge University Press. http://ebooks.cambridge.org/ref/id/CBO9780511612336.

Taylor, Marc H., Martin Losch, Manfred Wenzel, and Jens Schröter. 2013. "On the Sensitivity of Field Reconstruction and Prediction Using Empirical Orthogonal Functions Derived from Gappy Data." *Journal of Climate* 26 (22): 9194–9205. https://doi.org/10.1175/JCLI-D-13-00089.1.

Tyler, David E. 1982. "On the Optimality of the Simultaneous Redundancy Transformations." *Psychometrika* 47 (1): 77–86. https://doi.org/10.1007/BF02293852.