# Statistics for detecting climate change

**Serge Guillas**
**School of Mathematics**
**Georgia Institute of Technology**

*Supported by the Statistical and Applied Mathematical Sciences Institute*

# Overview

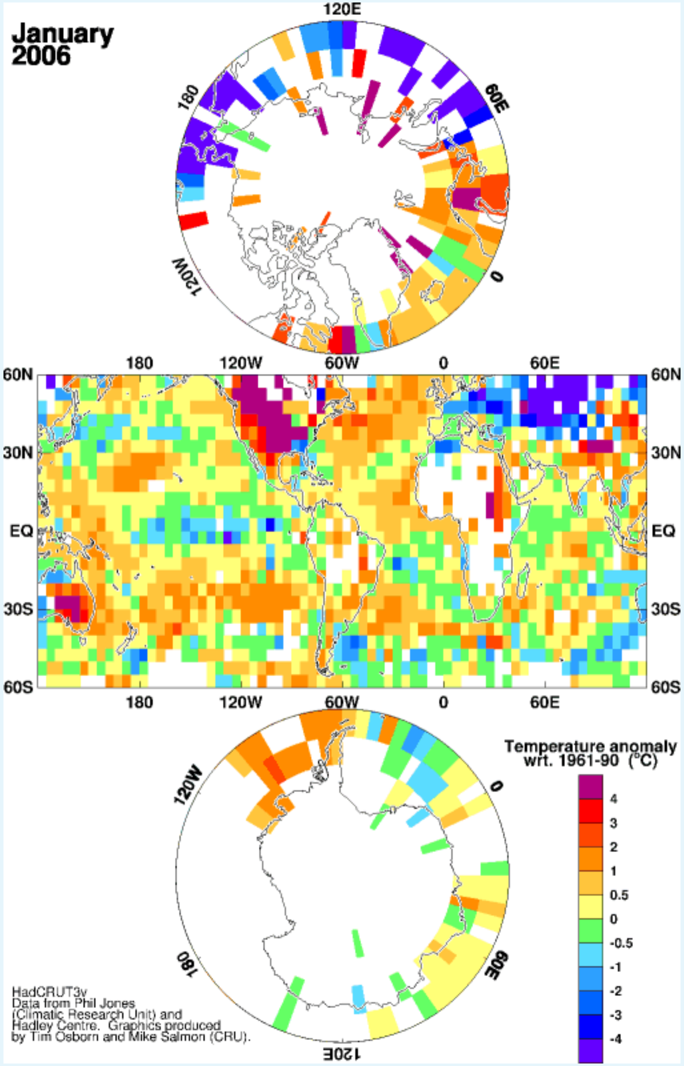
1. Introduction
2. Fingerprints, patterns of climate change
3. Detection and attribution
4. Bayesian approach
5. Future work

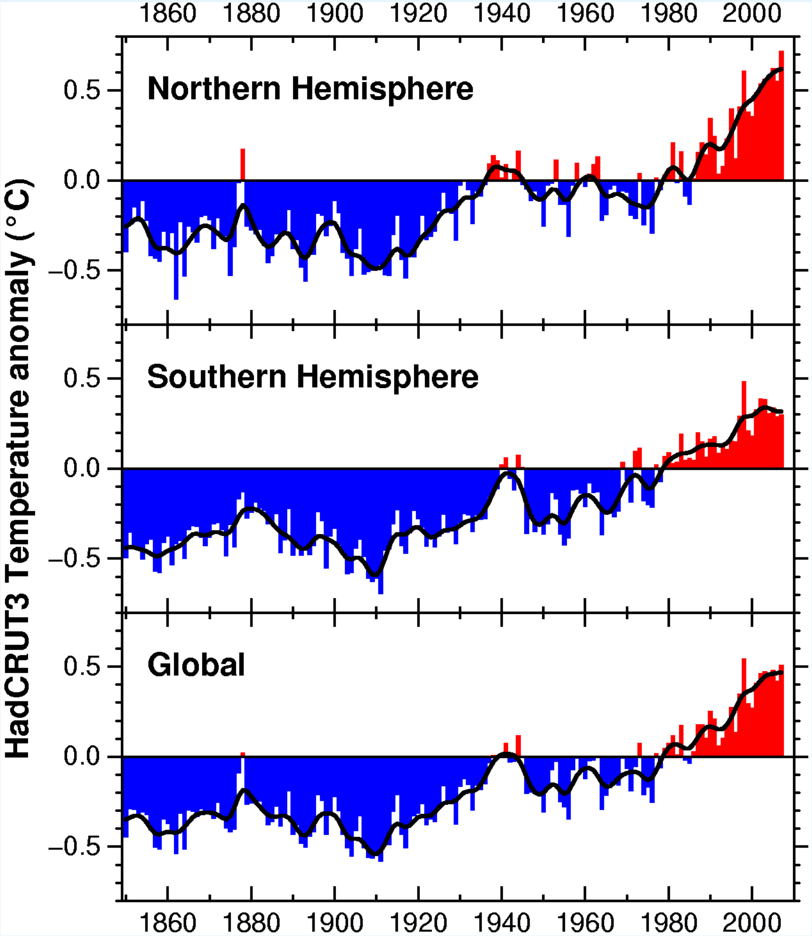
# Introduction

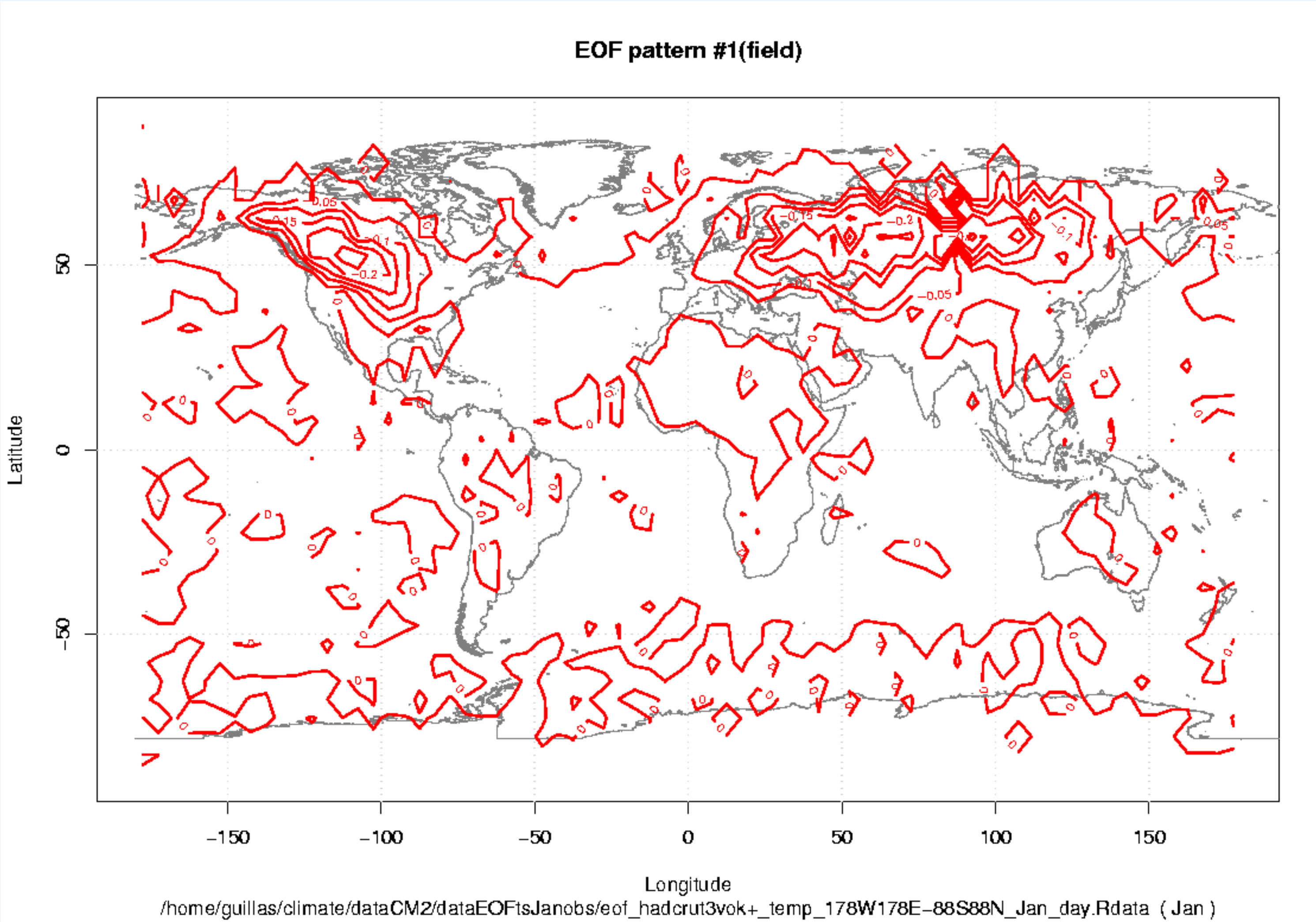
## 2006-2007 SAMSI climate and RM working group:

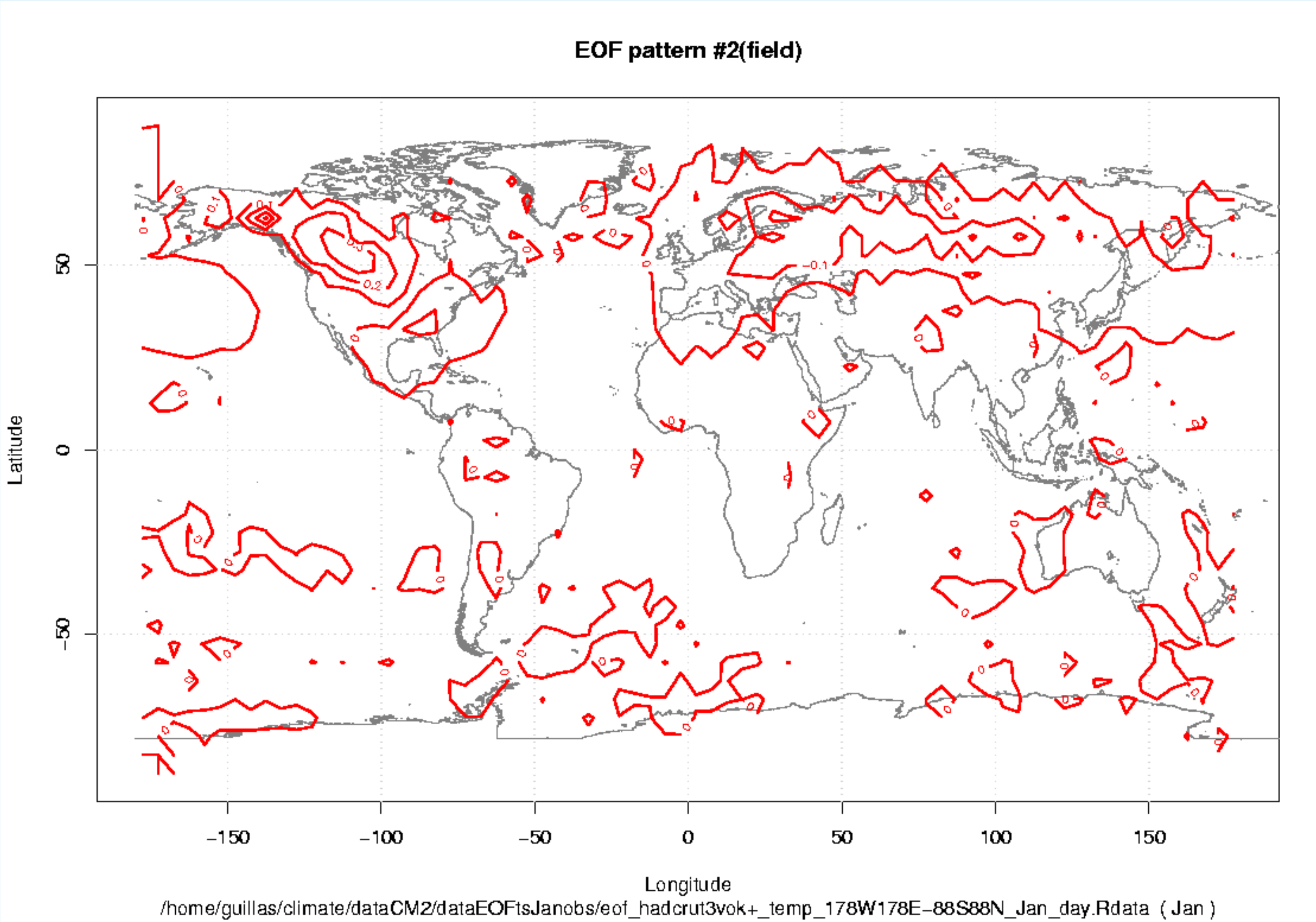
- reading group (EOF, detection & attribution,..)
- Undergraduate workshop
- Data portal: CMIP3 multi-model dataset at PCMDI (ex IPCC AR4 archive)

# Search for evidence of climate change

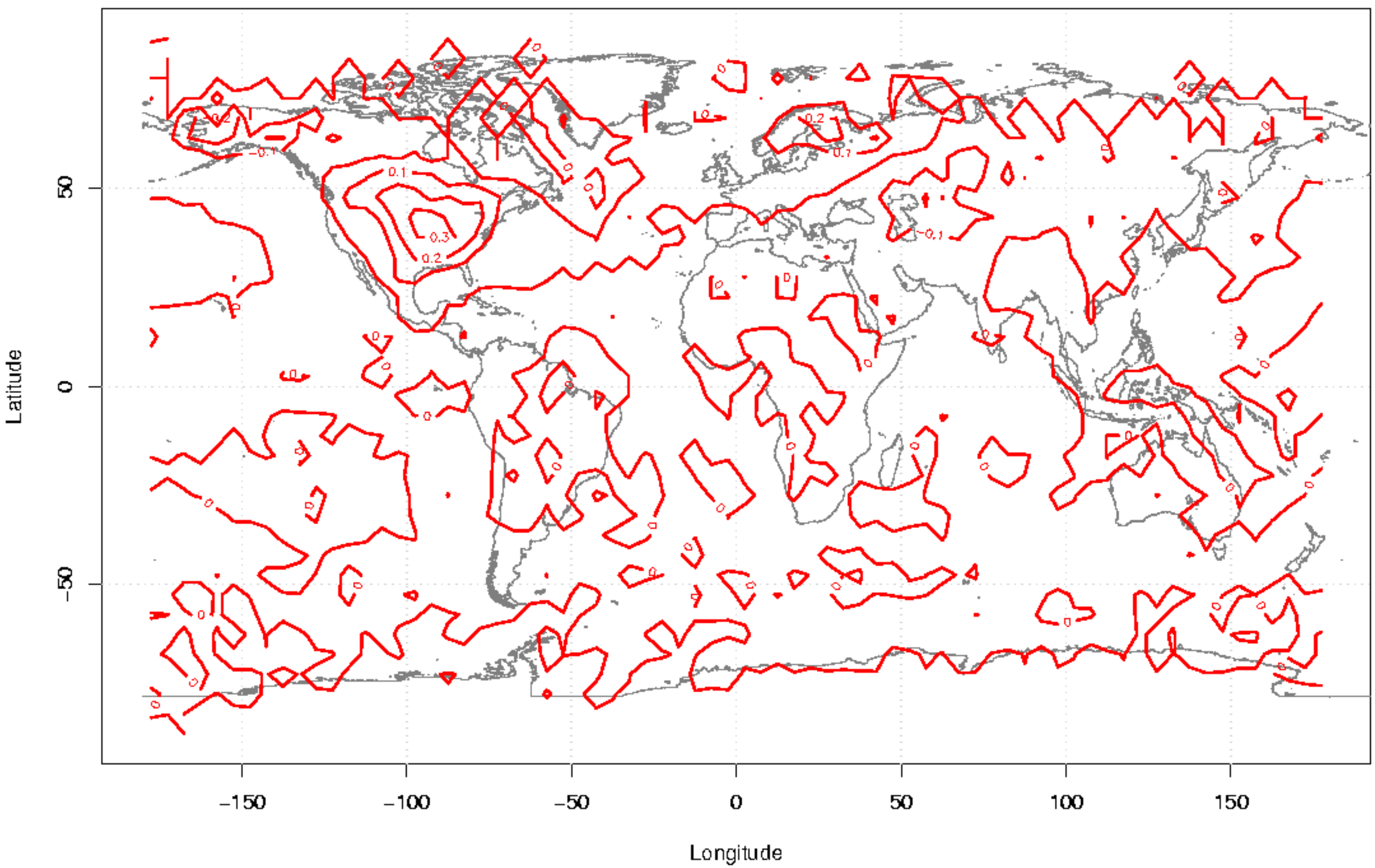



HadCRUT3 Temperature anomaly (°C)
Northern Hemisphere
Southern Hemisphere
Global
0.5
0.0
−0.5
1860
1880
1900
1920
1940
1960
1980
2000


EOF pattern #1(field)
Latitude
Longitude
/home/guillas/climate/dataCM2/dataEOFtsJanobs/eof_hadcrut3vok+_temp_178W178E-88S88N_Jan_day.Rdata ( Jan )
-150
-100
-50
0
50
100
150
50
0
-50
-0.05
-0.15
-0.1
-0.2
-0.15
-0.2
-0.05
-0.1
0.05


EOF pattern #2(field)
Latitude
50
0
-50
-150
-100
-50
0
50
100
150
Longitude
/home/guillas/climate/dataCM2/dataEOFtsJanobs/eof_hadcrut3vok+_temp_178W178E-88S88N_Jan_day.Rdata ( Jan )

**EOF pattern #3(field)**




/home/guillas/climate/dataCM2/dataEOFtsJanobs/eof_hadcrut3vok+_temp_178W178E-88S88N_Jan_day.Rdata ( Jan )

# Fingerprints, patterns of climate change

- Patterns are spatio-temporal responses of the climate to a certain forcing
- e.g. [only spatial] surface temperature change from increasing GHG or from anthropogenic sulfate aerosols
- Chosen a priori (physics or ensemble of runs of a model)

# Detection and attribution

**Hasselman (1997), Hegerl & North (1997), Allen & Tett (1999), Levine &Berliner (1999), Hegerl & Allen (2001),..**

## Detection of anthropogenic signals

**Decompose the observations $y$ into a linear combination of climate change "signals" $g_i$.**

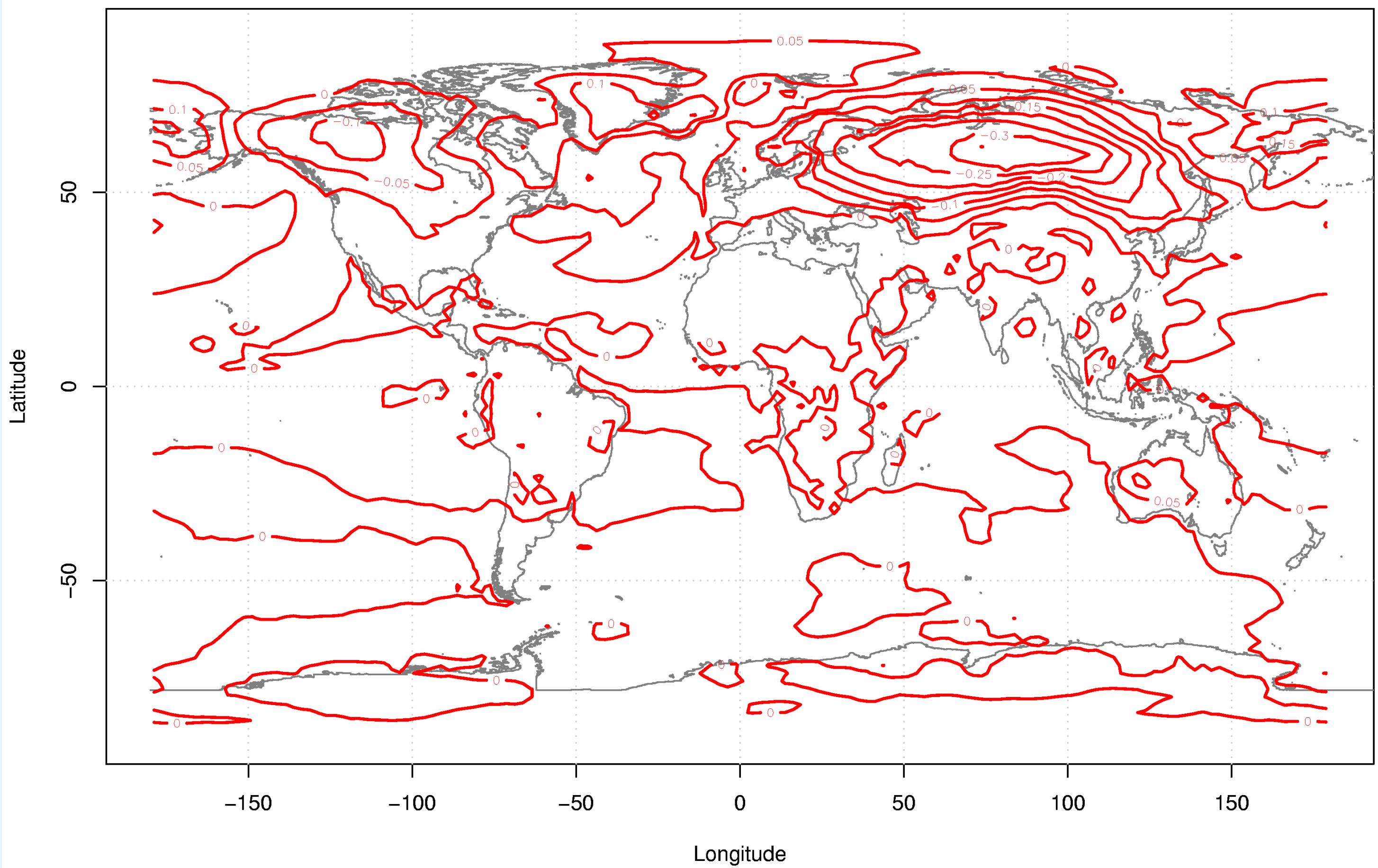
**$g_i$ : patterns (or building blocks for fingerprints) of responses to individual or combined forcings (solar, sulfate aerosol, GHG,..).**

$$y = \sum_{i=1}^{m} b_i g_i + \varepsilon \quad (1)$$

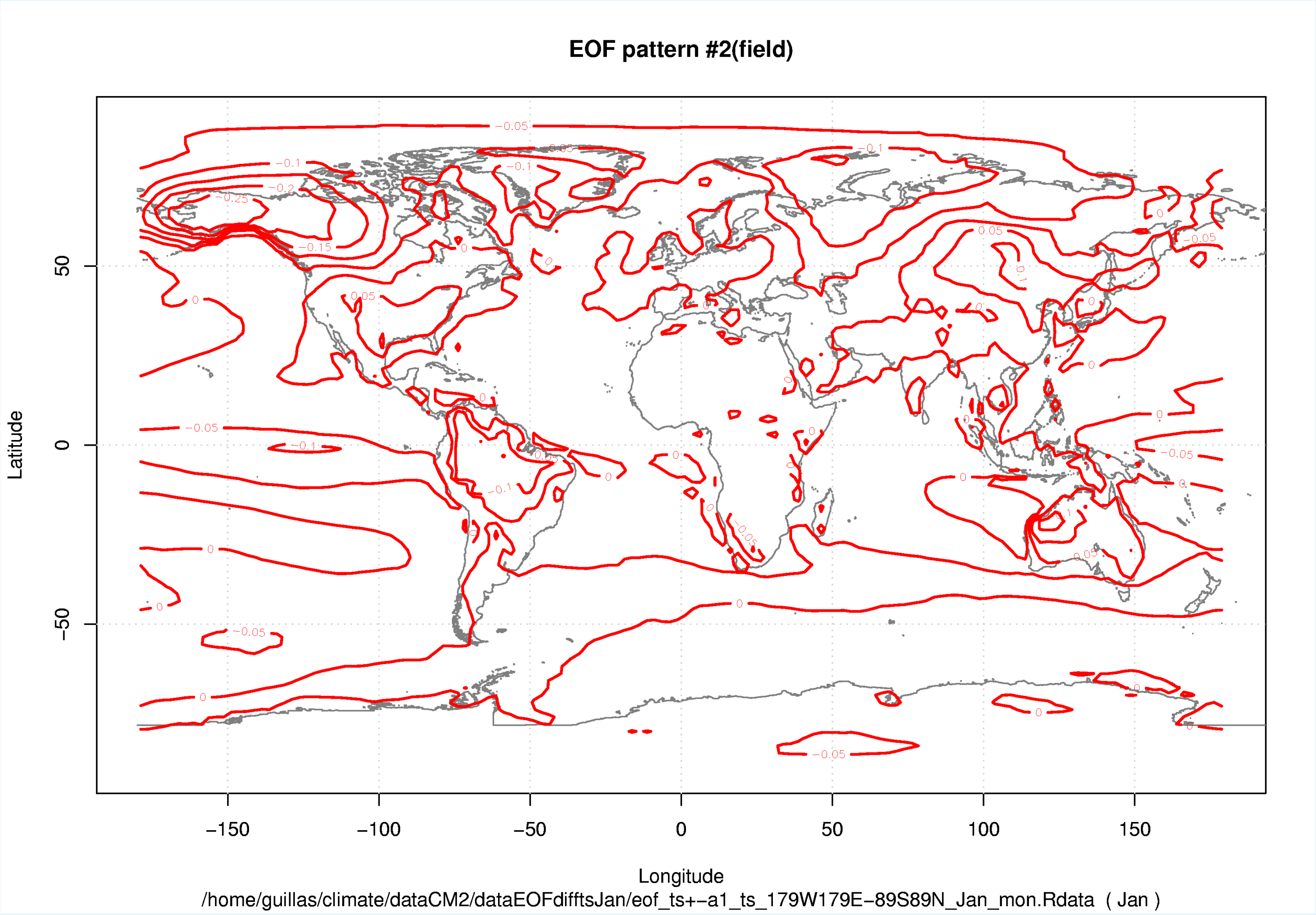
**Note: Linearity good assumption (Gillett et al., 2004)**

**Option: "orthogonalize" some signals prior to the regression (e.g. sulfate aerosol, greenhouse gas).**

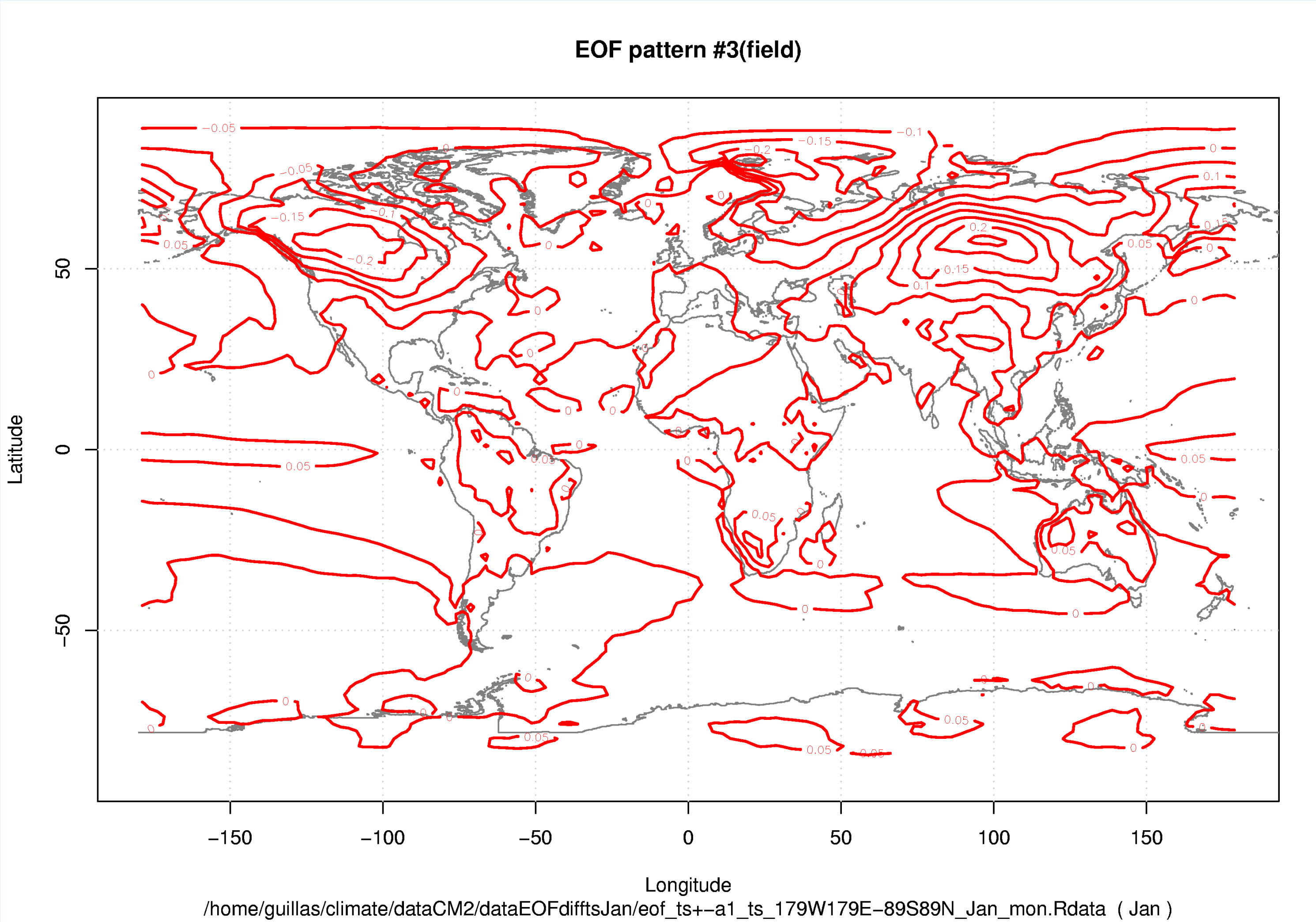
EOF pattern #1(field)

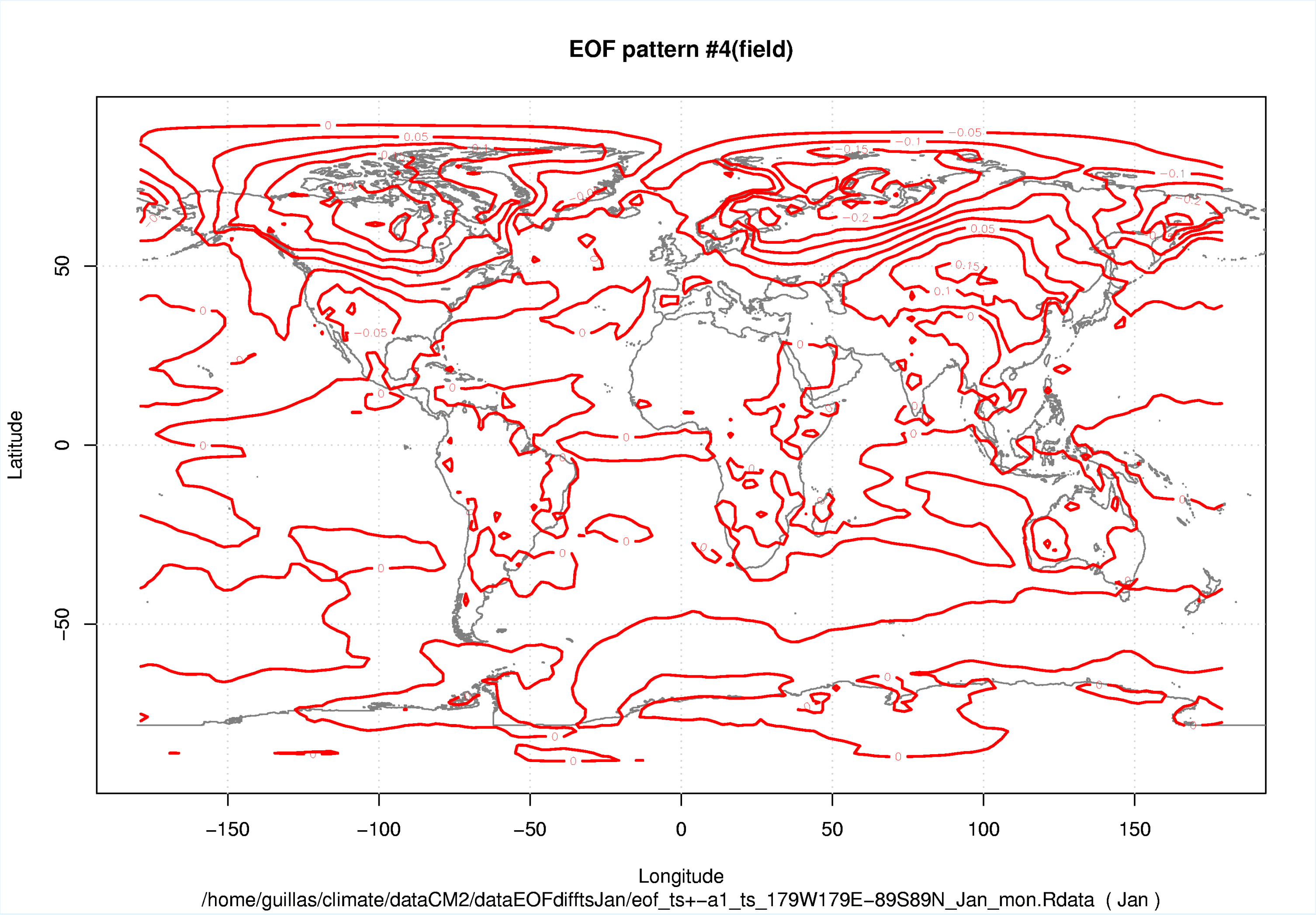



/home/guillas/climate/dataCM2/dataEOFdifftsJan/eof_ts+-a1_ts_179W179E-89S89N_Jan_mon.Rdata ( Jan )



EOF pattern #2(field)
Latitude
Longitude
/home/guillas/climate/dataCM2/dataEOFdifftsJan/eof_ts+-a1_ts_179W179E-89S89N_Jan_mon.Rdata ( Jan )


EOF pattern #3(field)
Latitude
Longitude
/home/guillas/climate/dataCM2/dataEOFdifftsJan/eof_ts+-a1_ts_179W179E-89S89N_Jan_mon.Rdata ( Jan )


EOF pattern #4(field)
Latitude
Longitude
/home/guillas/climate/dataCM2/dataEOFdifftsJan/eof_ts+-a1_ts_179W179E-89S89N_Jan_mon.Rdata ( Jan )

**Similar results with more general approach:**

## Regression

- $Y$: observations $p$-dimensional
- $X$: patterns of climate change from model, known $p \times m$ matrix of ranq $q$.
- each column of $X$ is a vector representing a response-pattern spatially

**Linear regression: (Mardia et al. 1979)**

$$Y = X\beta + u \tag{2}$$

**with $u$ is the climate noise, $Eu = 0$, $C(u) = \sigma^2 I$.**

## Ordinary Least Squares:

**Let** $\widehat{\beta} = (X'X)^{-1} X'Y$.

**Th. [Gauss-Markov Th]:** $\widehat{\beta}$ **is the BLUE of** $\beta$.

## Generalized Least Squares:

**When** $C(u) \neq \sigma^2 I$, **Consider the transformed model:**

$$Z = C(u)^{-1/2} X\beta + v \tag{3}$$

**where** $Z = C(u)^{-1/2}Y$, $v = C(u)^{-1/2}u$.

**"Pre-whitening" with any matrix P such that:**

$$E\left(Puu'P'\right) = PC(u)P' = I. \tag{4}$$

$C(v) = I$, **so we can apply G-M to this model:**

**Let**

$$\tilde{\beta} = \left(X'C(u)^{-1}X\right)^{-1} X'C(u)^{-1/2}z = \left(X'C(u)^{-1}X\right)^{-1} X'C(u)^{-1}y. \tag{5}$$

$\tilde{\beta}$ **is the BLUE of** $\beta$, **and** $C(\tilde{\beta}) = \left(X'C(u)^{-1}X\right)^{-1}$.

**Possible to look at scalar diagnostics $\phi = w'Y$. (e.g. global mean, or focus on one mean for one grid cell)**

**Issue: noise in $X$ inflates the variance of $\tilde{\beta}$ by approximately $1 + 1/M$ ($M =$ ens. size)**

## Fingerprints

**The columns of $C(u)^{-1}X$ are the optimal fingerprints.**

# Climate noise

**Note**: $C(u)$ is unknown, so:

- $\widehat{C}_n(u) = \frac{1}{n} Y_n Y_n'$ can be plugged in.
- $Y_n$ are "pseudo-observations" from a control run with..
  features as close as possible to the observations (locations of missing data,..)

**Problem:** $\widehat{C}_n(u)$ **not invertible,** $(p > n)$**, so:**

1. **use** $k$ **EOFs of control runs (or sometimes of forced runs)**
2. **Define** $P_{(k)}$ **as matrix of** $k$ **highest variance EOFs weighted by** $\sqrt{\lambda_i}$
3. **use the Moore-Penrose pseudo-inverse** $P'_{(k)}P_{(k)}$ **in place of** $\widehat{C}_n^{-1}$

**So** $P'_{(k)}P_{(k)} = I_k$

**Issue: depends on** $k$**!**

## Tests and confidence regions

**Under normality assumption for u,**

$$\left(\tilde{\beta} - \beta\right)\left(X'C(u)^{-1}X\right)^{-1}\left(\tilde{\beta} - \beta\right) \sim \chi^2_m \tag{6}$$

**Using EOFs:**

**With an estimated Covariance matrix (often on another sample),**
**test becomes a $T^2$-test, using $F-$distributions.**

$$\left(\tilde{\beta} - \beta\right)\left(X'C(u)^{-1}X\right)^{-1}\left(\tilde{\beta} - \beta\right) \sim T^2 \tag{7}$$

# Bayesian approach

**Berliner et al. (2000)**

**True vector of temperatures** $T_t$

**Observations** $Y_t$

$$Y_t|T_t, D_t \sim N\left(L_tT_t, D_t\right) \tag{8}$$

**with** $L_t$ **location matrix (only 0 except 1 for the location)**

$$T_t|a, g, \Sigma \sim N\left(a.g, \Sigma^s\right) \tag{9}$$

**with** $g$ **spatial fingerprint,** $\Sigma^s$ **spatial covariance.**

**Assumption of space-time separability.**

**Prior on $a$: (actually collection of)**

$$\pi(a) = pn(0, \sigma^2) + (1-p)n(\mu_A, \tau_A^2) \tag{10}$$

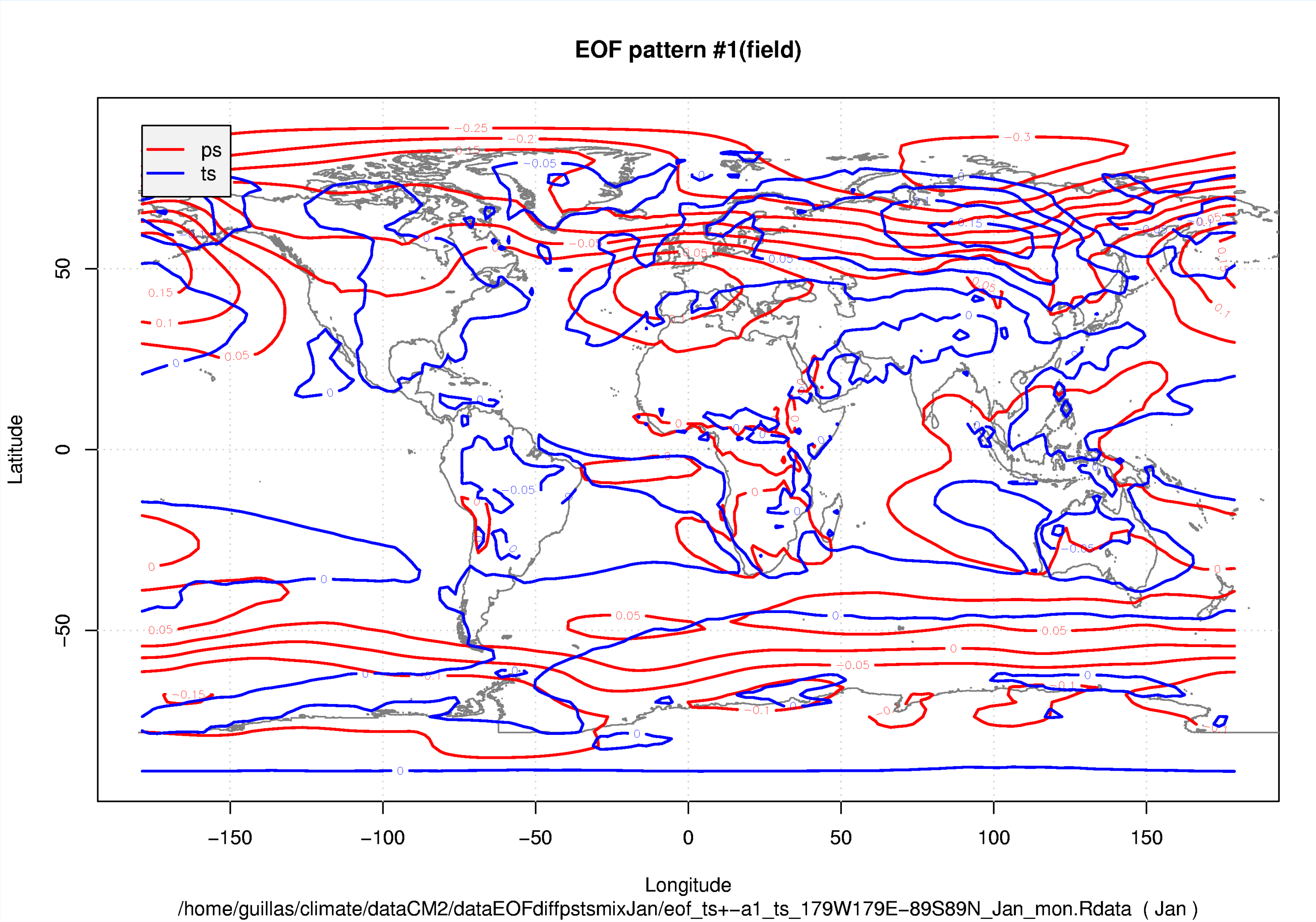
# Future work

## Common EOFs

Benestad (2001) [downscaling studies]

- concatenate the fields (temp, pressure)
- Carry out the EOF decomposition alltogether.
- EOFs more representative of the patterns.

..not yet used for detection & attribution


EOF pattern #1(field)
ps
ts
Latitude
50
0
−50
−150
−100
−50
0
50
100
150
Longitude
/home/guillas/climate/dataCM2/dataEOFdiffpstsmixJan/eof_ts+−a1_ts_179W179E−89S89N_Jan_mon.Rdata ( Jan )

## Spatio-temporal EOFs
## (North and Wu, 2001)

**Issues:**

- size of covariance matrices
- type of correlation
- truncation