# Guidelines for the use of multivariate Granger causality

Esther Florin

December $10^{th}$, 2014

MVAR Causality

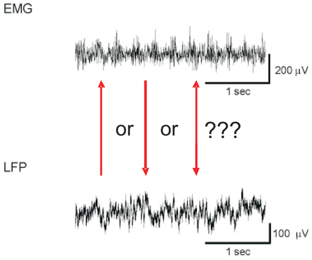
Simulations

Preoprcessing

# Granger Causality

**Idea:**

Temporal preceding signals can be causal for another signal.




$$X_t = \sum_{j=1}^{\infty} a_j X_{t-j} + \varepsilon_t$$

$$Y_t = \sum_{j=1}^{\infty} a_j Y_{t-j} + \eta_t$$

$$X_t = \sum_{j=1}^{\infty} c_j X_{t-j} + \sum_{j=1}^{\infty} d_j Y_{t-j} + \varepsilon_t$$

$$Y_t = \sum_{j=1}^{\infty} e_j Y_{t-j} + \sum_{j=1}^{\infty} f_j X_{t-j} + \eta_t$$

→ F-Test

# Causality Measures

Based on Granger causality several measures to tackle the question of directionality in Neuroscience have been proposed:

- Directed Transfer Function (DTF) (Kaminski et al. 1991)
- Direct Directed Transfer Function (dDTF) (Korzeniewska et al. 2003)
- Partial Directed Coherence (PDC) (Baccala et al. 2001)
- Squared Partial Directed Coherence (sPDC) (Astolfi et al. 2006)
- H - Transfer Matrix

# Multivariate Autoregressive Model

MVAR:

$$\mathbf{X}(t) = \sum_{L=1}^{p} \mathbf{A}(L)\mathbf{X}(t-L) + \mathbf{E}(t)$$

Transfer to the frequency domain:

$$\mathbf{A}(f) = \sum_{L=0}^{p} \mathbf{A}(L) e^{-j2\pi f \Delta t L}$$

Transfer function:

$$\mathbf{H}(f) = \mathbf{A}(f)^{-1}$$

# Directed Transfer Function

$$\gamma_{ij}^2 = \frac{|H_{ij}(f)|^2}{\sum_{m=1}^{k} |H_{im}(f)|^2}$$

- DTF describes the proportion of information transfer from channel j to channel i with respect to all information, which flow to channel i (Kus 2004, Kaminski 2001).
- Takes values between 0 and 1

**Problem:**

No distinction between direct and indirect connections

**Solution:**

- Direct Directed Transfer Function (dDTF)
- Partial Directed Coherence (PDC)

# Partial Directed Coherence

PDC describes the proportion of information transfer from channel j to i with respect to the total information flow to channel j (Baccala and Samechima 2001).

$$\pi_{ij}(f) = \frac{A_{ij}(f)}{\sqrt{\sum_{k=1}^{N} A_{ki}(f) A^{*}_{kj}(f)}}$$

- Takes values between 0 and 1

**Advantage**

Distinction between direct and indirect connections

Astolfi et al. 2006 suggested in addition the squared partial directed coherence(sPDC)

# Evaluation of the causality measures

- Comparison of the different methods
- Determination of a significance level
- Influencing factors such as data length and noise level
- Different data types
- Preprocessing

# Model to test the causality measures

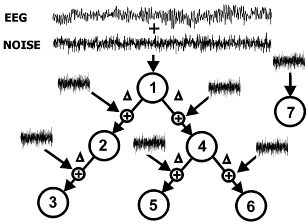



Figure: Kus Model

## Datatypes

EEG, MEG, EMG and LFP

## Parameters to be varied

- Data Length:
  100 to 1000 in steps of 100
  1000 to 13500 in steps of 500
- Model Order:
  1 to 49 in steps of 2
- Noise Level:
  0 to 10 in steps of 0.2
- Coupling Strength:
  0.2 to 2 in steps of 0.2

# Significance

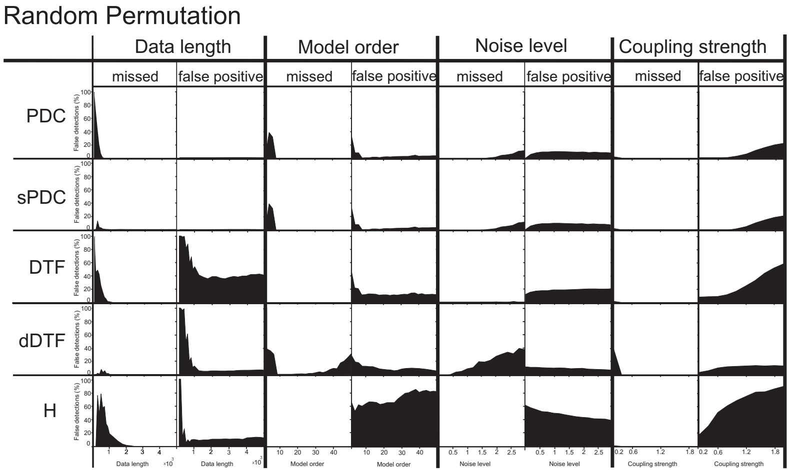
## Shuffling of the data:

- Time course of the data randomized
- Determination of causality measure
- 200 repetitions
- 99 % percentile

## Leave One Out Method (LOOM)

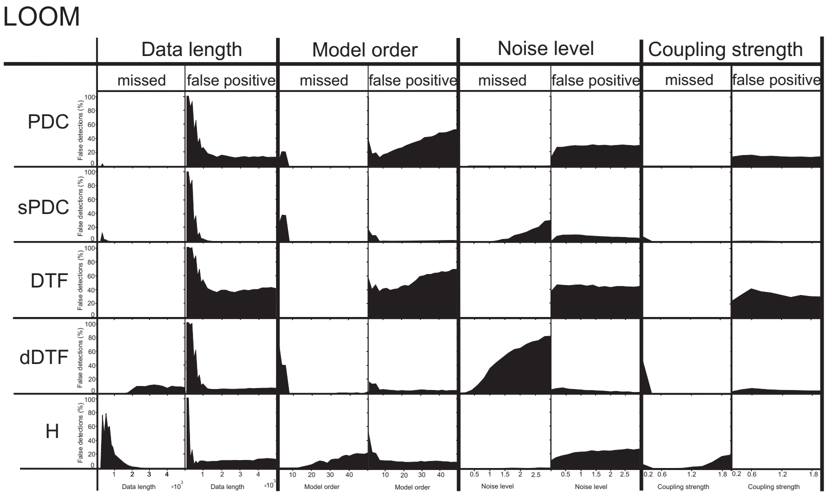
- One data segment left out
- Determination of causality measure
- As many repetitions as data segments
- 99 % percentile

# Results with shuffling




Florin et al., Journal of Neuroscience Methods 2011

# Results with LOOM




Florin et al., Journal of Neuroscience Methods 2011

# Summary

## Dependence on the varied parameters

- Increasing data length leads to better results
- Only LOOM shows dependence on model order
- No dependence on noise level
- Coupling strength: Increasing differences in signal to noise ratio lead to wrong detections

## LOOM vs. Shuffle

Obtaining the significance with shuffling leads to better results.

## Recommended Method

The most reliable results are obtained with the sPDC.

Florin et al., Journal of Neuroscience Methods 2011

# Preprocessing

## AR-Model

$$y_t = \sum_{j=0}^{\infty} a_j x_{t-j} + u_t$$

## Estimate of a in $y_t = \sum_{j=0}^{\infty} a_j x_{t-j} + u_t$:

$$a(L) = \frac{g_{yx}(L)}{g_x(L)} = \frac{covariance}{variance}$$

## Lag Operator

Lag-Operator L shifts a time series by one point in time.

# Basics

## Wold infinite moving average representation

Every covariance-stationary process has an infinite moving average representation:

$$x_t = \sum_{i=0}^{\infty} b_j u_{t-i} = B(L) u_t$$

# Covariance generating function

$$g_x(L) = \sigma^2 B(L^{-1})B(L) \text{ mit: } B(L) = \sum_{j=-\infty}^{\infty} b_j L^j$$

## a with only positive lags

$$a = [\frac{g_{yx}(z)}{B(L^{-1})}]_+ \frac{1}{B(L)}$$

# Filtering

## Filtering and Lag-representation

$$x_t^f = f(L)x_t \text{ mit: } f(L) = \sum_{j=-\infty}^{\infty} f_j L^j$$

## Example: Filtering with a moving average

$y_t = \frac{1}{3}x_{t+1} + \frac{1}{3}x_t + \frac{1}{3}x_{t-1}$
$y_t = x_t(\frac{1}{3}L^{-1} + \frac{1}{3}L^0 + \frac{1}{3}L)$
$y_t = D(L)x_t$

# Filtering

**Covariance-generating function after the application of a filter**

$g_{y^f} = f(z)f(z^{-1})g_y(z)$
$g_{y^f x^f} = f(z)f(z^{-1})g_{yx}(z)$

**AR representation of the filtered function**

$$x_t^f = \sum_{j=0}^{\infty} a_j^f x_{t-j}^f + \tilde{u}_t$$

Same regression results only if:
$a_j = a_j^f$

# $a_j = a_j^f$?

## Insertion of the terms for the covariance matrix:

$$a_j^f(L) = \left[\frac{g_{y^f x^f}}{g_{x^f}}\right]$$

$$\Rightarrow a_j^f(L) = [\frac{f(L)f(L^{-1})g_{yx}(L)}{c(L^{-1})}]_+ \frac{1}{c(L)} \neq [\frac{g_{yx}(L)}{B(L^{-1})}]_+ \frac{1}{B(L)} = a_j$$

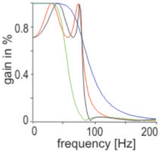
$$g_{y^f x^f} = f(L)f(L^{-1})g_{yx}(L)$$

$$g_{x^f} = c(L) * c(L^{-1}) = f(L)f(L^{-1})B(L)B(L^{-1})$$

In general the following does not hold: c(L)=f(L)B(L)

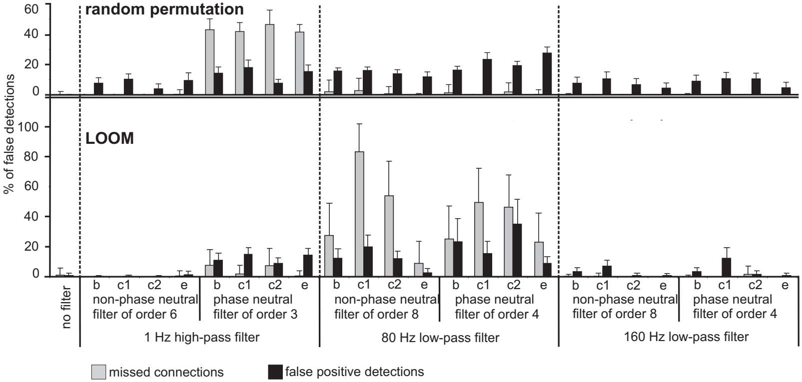
# Simulation Filtering

- Butterworth, Chebychev type I + II, and elliptic filter



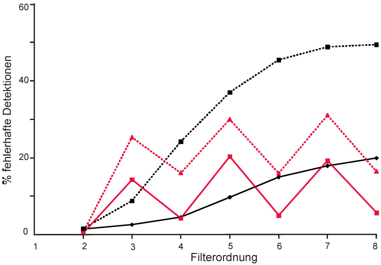

- phase-neutral and non phase-neutral filters
  - 1 Hz high pass
  - 80 Hz low pass
  - 160 Hz low pass
- Filter order
- 50 Hz line noise
- decimation and interpolation

# Results: sPDC

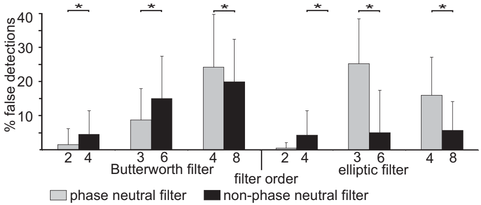



Florin et al., Neuroimage 2010

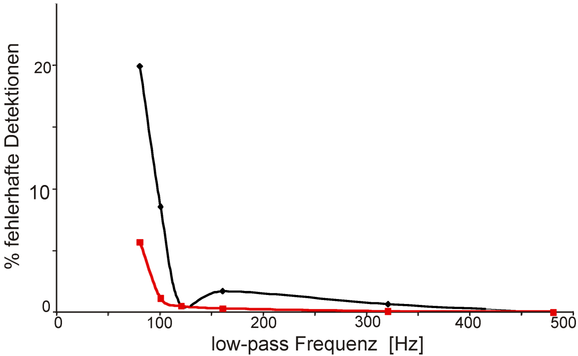
# Influence of the filter order and phase-neutral filtering




Florin et al., Neuroimage 2010



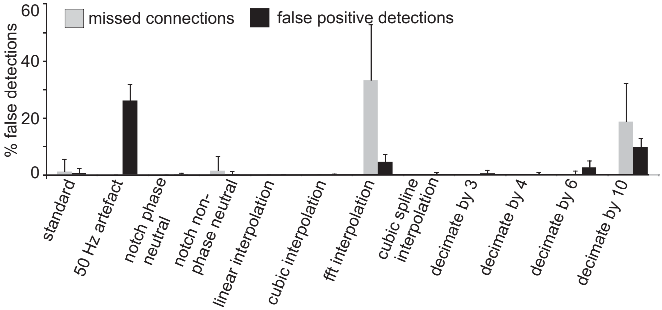
# Different Low-Pass Filters




Florin et al., Neuroimage 2010

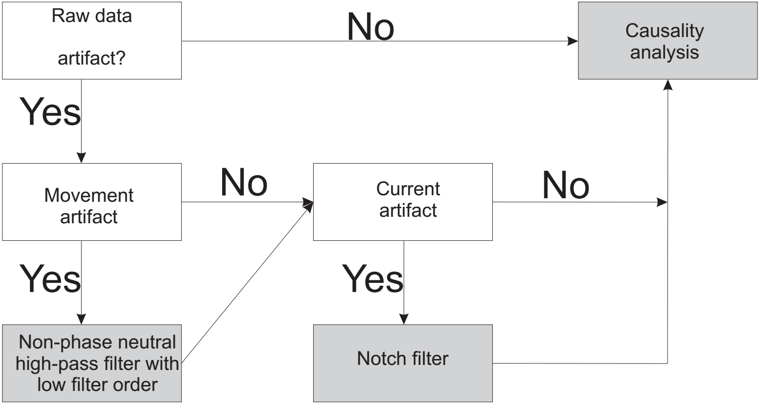
# Decimation and interpolation

Temporal aggregation may result in the loss of information contained in the variance-covariance matrix (Breitung and Swanson 2002).




Florin et al., Neuroimage 2010

# Recommendation pre-processing




Florin et al., Neuroimage 2010

# Summary

## Preprocessing

- Most filtering techniques will lead to false detections.
- Interpolation and downsampling will also lead to false detections.

## Determination of the significance level

Both Leave one out method and random permutation yield similar results.

## Recommended multivariate method

Squared partial directed coherence in combination with leave one out method

Florin et al., NeuroImage 2010; Florin et al., J Neuroscience Methods, 2011

# Thank you for your attention!