A Bayesian semi-parametric model for functional near-infrared spectroscopy data

Timothy D. Johnson

Department of Biostatistics
University of Michigan

Joint work with
Marco Benedetti, Ioulia Kovelman, Xiaosu (Frank) Hu

Founded in 2012

- Founding Officers:
  - Chair: Dan Rowe (Marquette University)
  - Chair-elect: Hongtu Zhu (University of North Carolina)
  - Secretary: Brian Caffo (Johns Hopkins University)
  - Treasurer: DuBois Bowman (Emory University)
  - Program Chair: Martin Lindquist (Columbia University)
  - Program Chair-elect: Hernando Ombao (Brown University)
  - Publications Liaison: Ranjan Maitra (Iowa State University)
  - COS Representative: Kary Myers (Los Alamos National Laboratory)
  - Tom Nichols

Charge of the Section

The Statistics in Imaging Section promotes statistics and statisticians' work in all areas of the imaging sciences. We foster research, education, and influence of statistics and statisticians on imaging science and its associated areas of application.

Our Goals

- Increase the influence of statistics and statisticians on imaging science and its associated areas of application
- Produce a focal organization and meeting-place for statisticians working in imaging science and its associated areas of application
- Introduce statisticians and students to modern problems in statistical imaging science and its associated areas of application
- Organize statisticians for addressing key issues in imaging science

Our Primary Functions

- Sponsorship, including joint sponsorship with other organizations, sessions at ASA meetings, sessions at imaging conferences, meetings, seminars and short-courses
- Planning, in cooperation with the Program Committee of the ASA, sessions in national and regional meetings of the ASA, IMS, IBS, ENAR and WNAR
- Promotion of academic, translational and non-journal research in statistics in imaging sciences
- Service as a resource for public and private agencies seeking assistance in these fields of interest

Past Section Chairs

- Dan Rowe, 2012
- Hongtu Zhu, 2013
- Tim Johnson, 2014
- Ciprian Crainiceanu, 2015
- John Kornak, 2016
- Martin Lindquist, 2017
- Nicole Lazar, 2018
- Hernando Ombao, 2019
- Anuj Srivastava, 2020
- Tingting Zhang, 2021
- Ying Guo (chair-elect) 2022

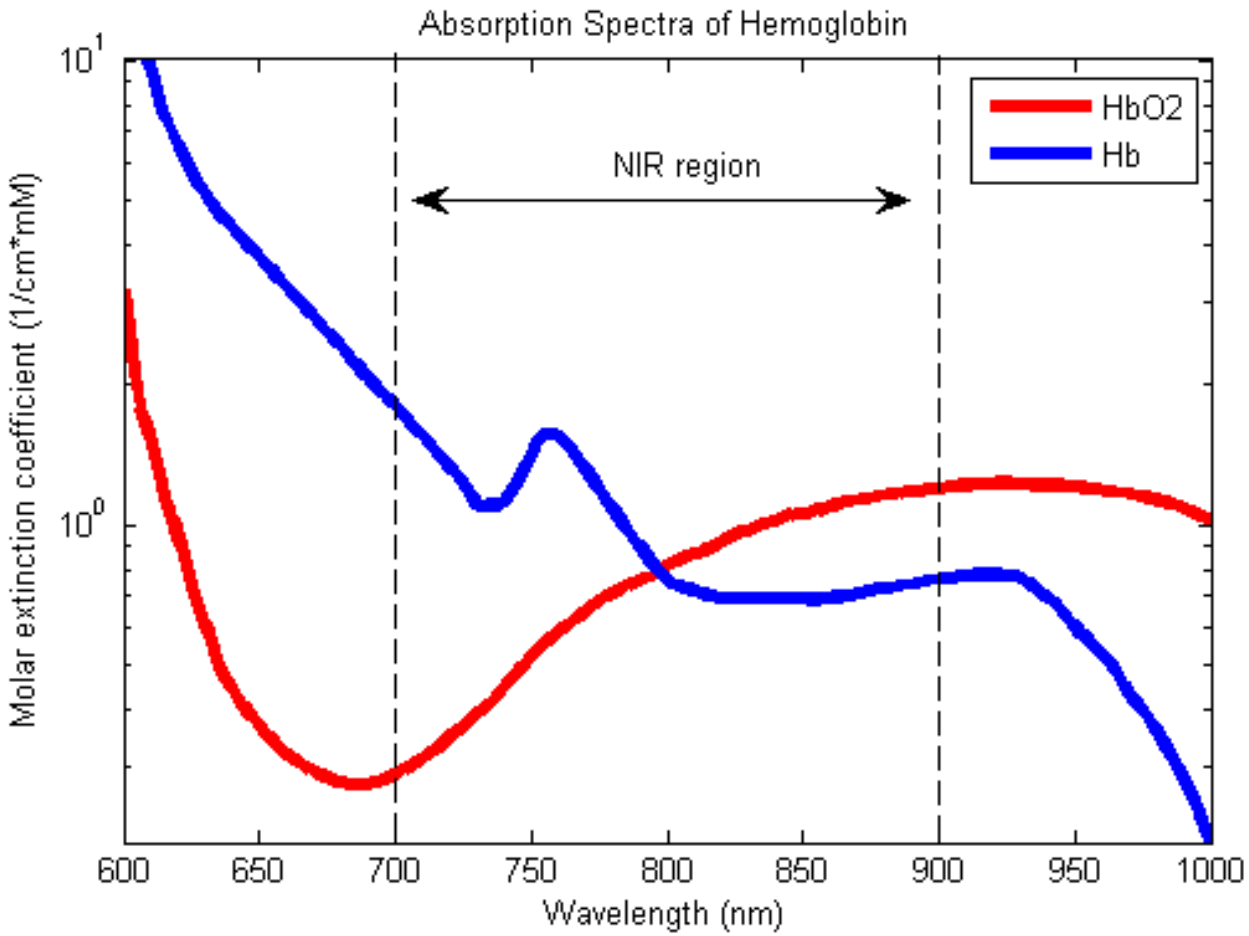
History of this meeting

- 2015: University of Michigan
- 2016: University of Colorado
- 2017: University of Pittsburg
- 2018: University of Pennsylvania
- 2019: University of California, Irvine
- 2020: (Covid-19 hiatus)
- 2021: Emory University

Our primary source of funding. Proceeds used to sponsor student paper competitions, JSM section meeting, and other functions as deemed appropriate by current officers.

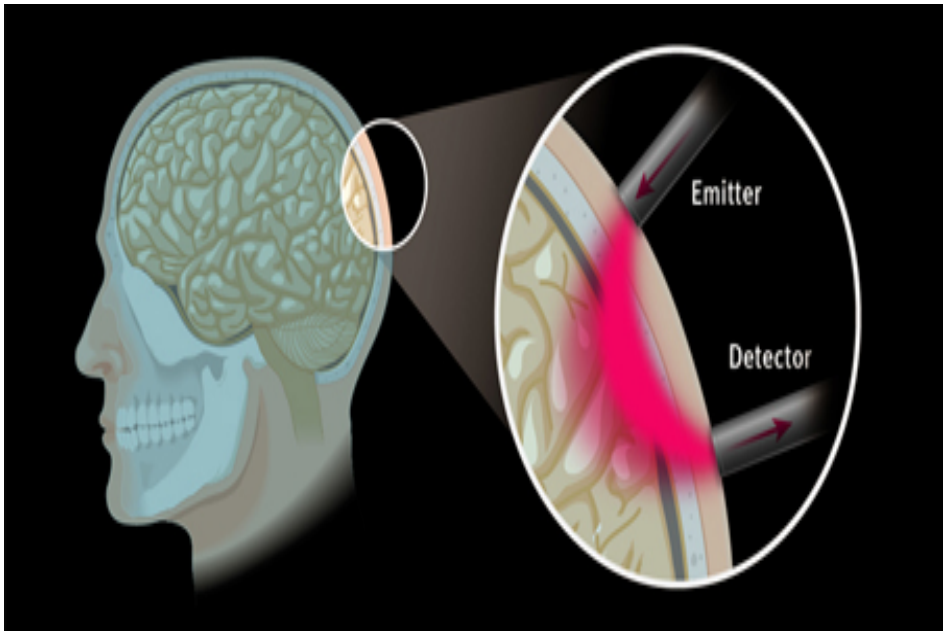
What is fNIRS?

Oxy/Deoxy. blood NIR light absorption



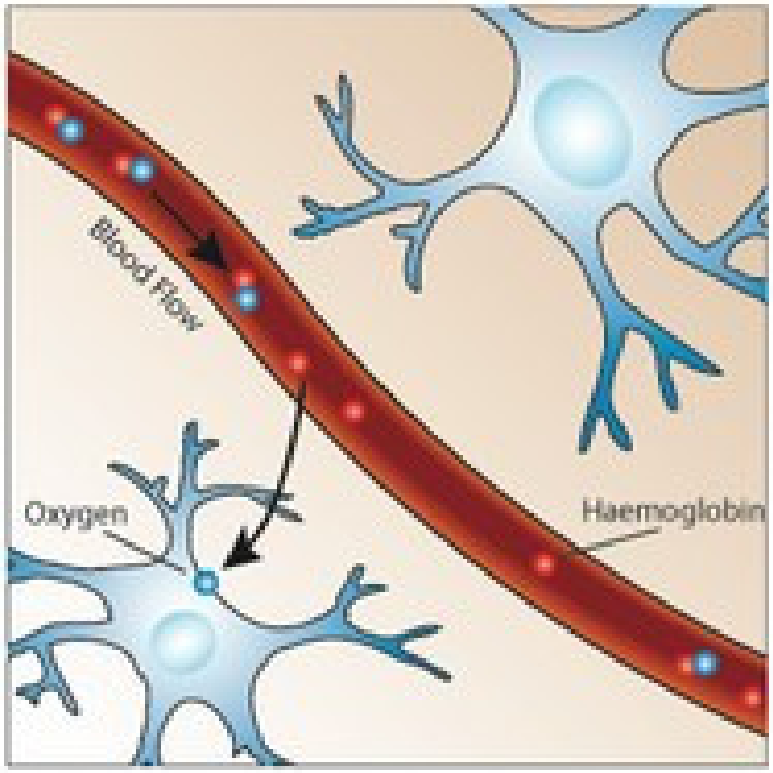
What is fNIRS?

Functional Near Infrared Spectroscopy
uses light to measure concentration of Oxy/Deoxy hemoglobin

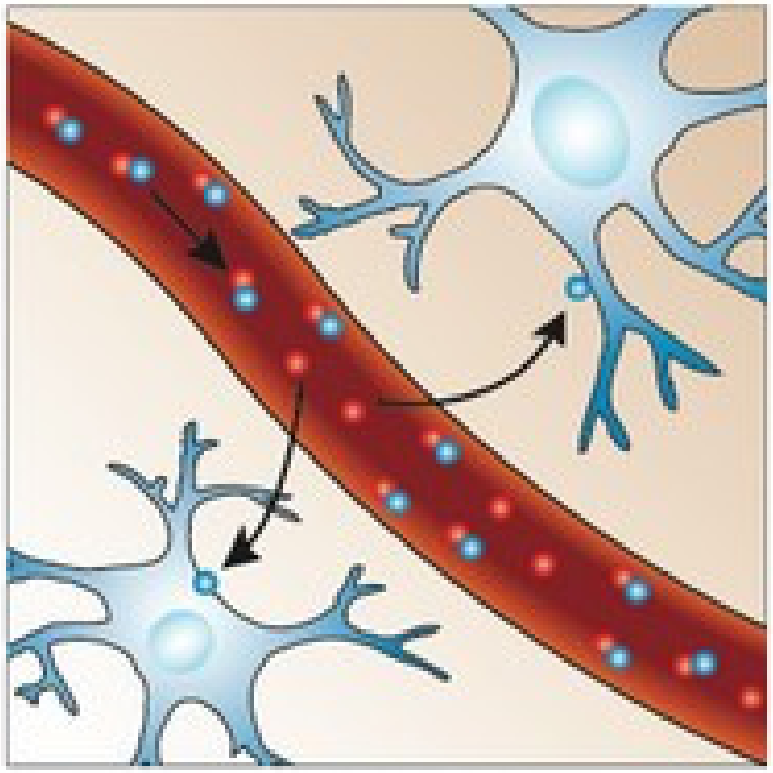


Physiological Response to Stimuli (BOLD response)

BOLD-Blood Oxygen Level Dependent



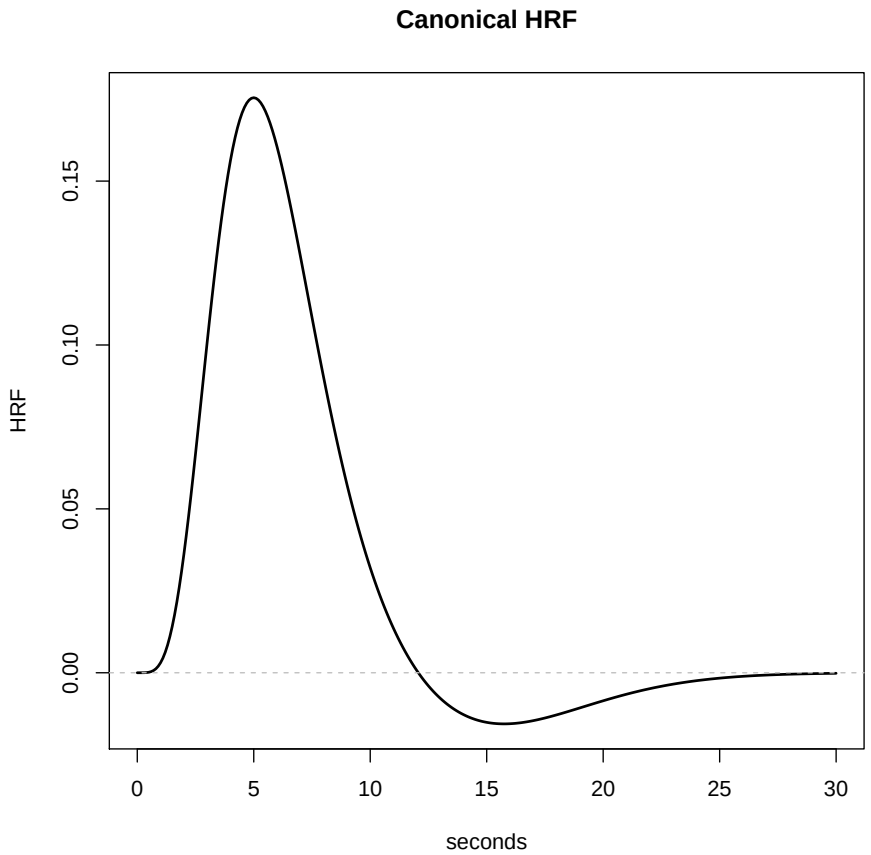
Resting



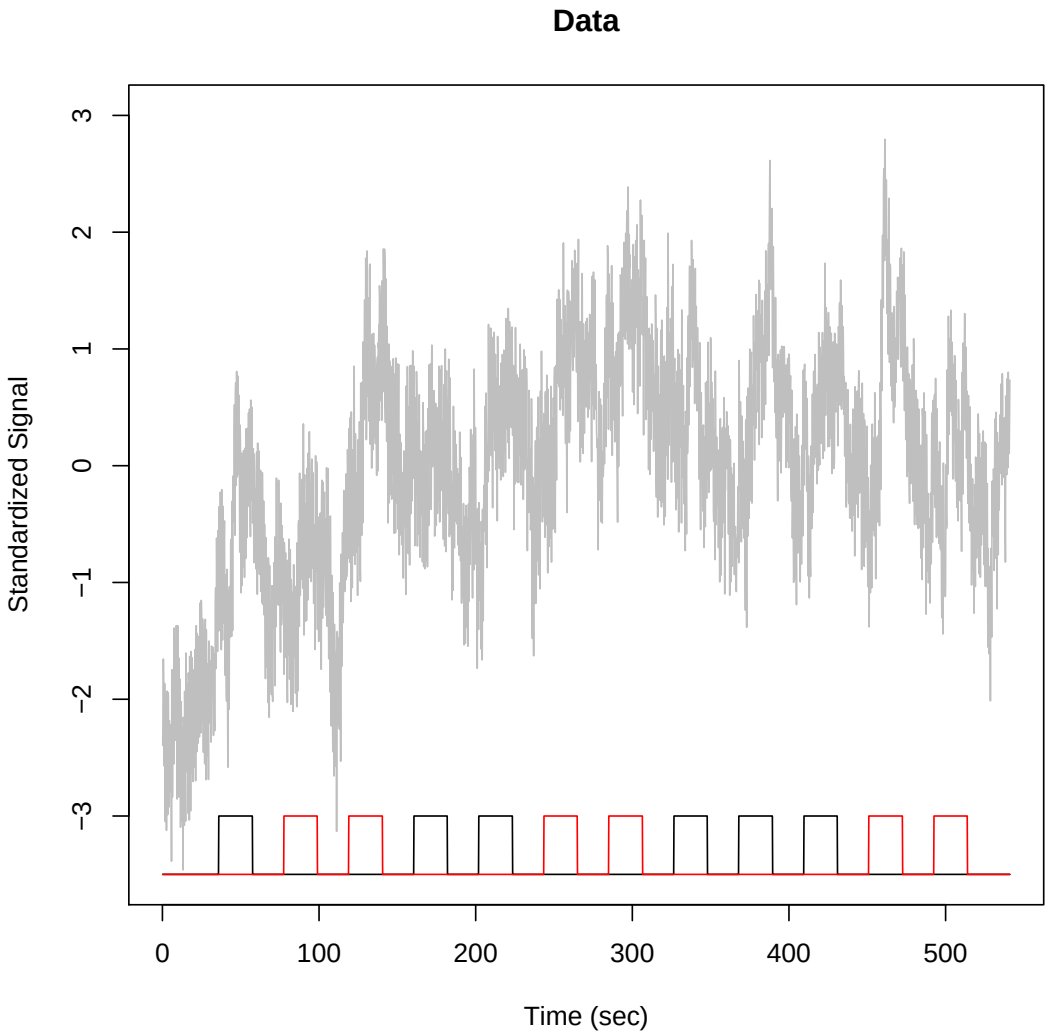
Activated

Physiological Response to Stimuli (BOLD response)

- BOLD response not instantaneous
- BOLD response modeled by the HRF
- Canonical HRF used - often with time and dispersion derivatives



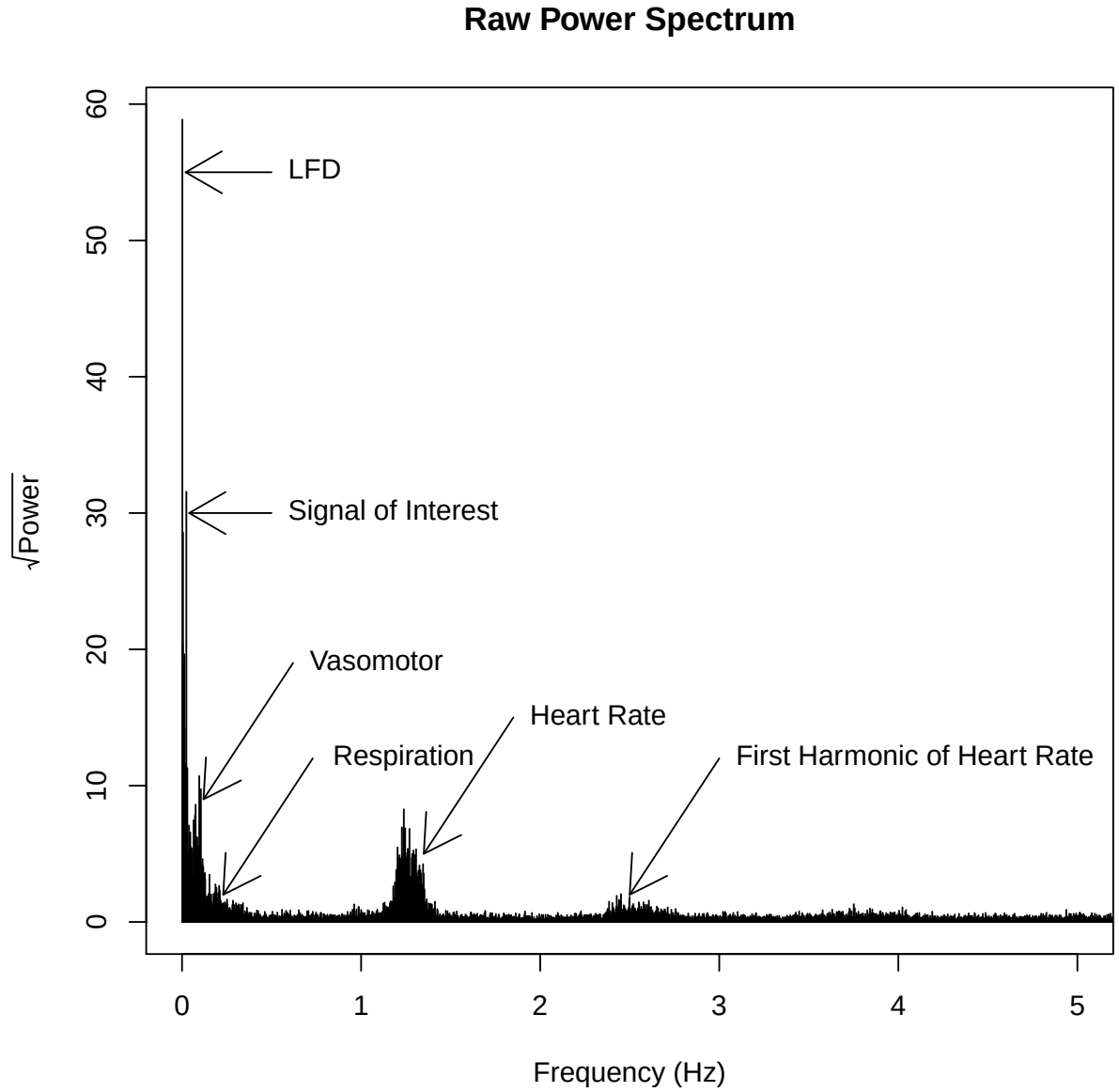
Example of fNIRS data



Challenges with fNIRS statistical modeling

- Physiological signals can be strong and picked up by fNIRS and are quasi-periodic
  - vasomotor—frequency very close to block design frequency
  - heart beat
  - respiration
- Fast temporal sampling resulting in high auto-correlation
- Motion artifacts
- Heteroscedasticity

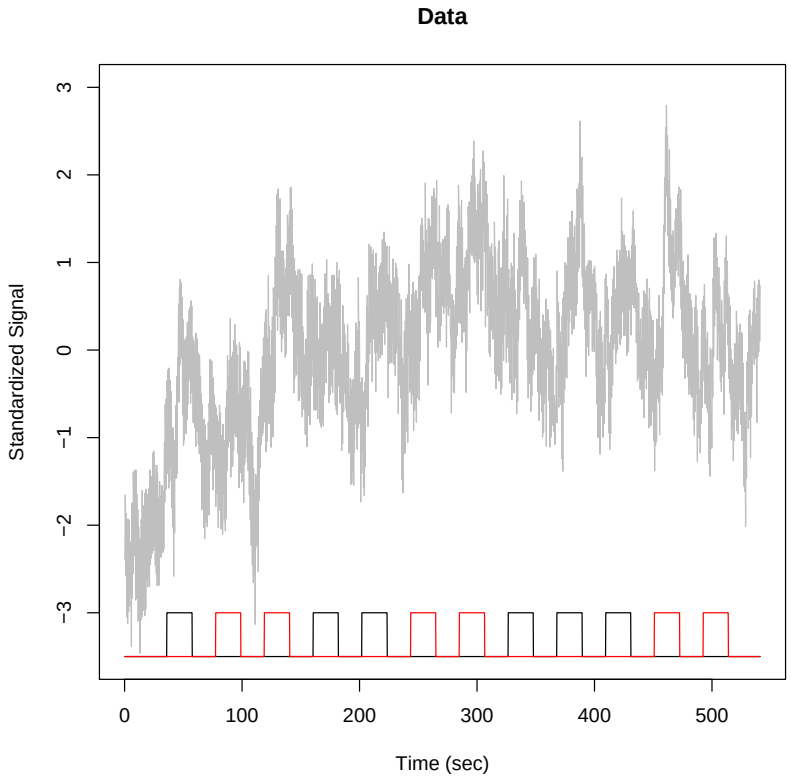
Power Spectrum Components



Bayesian semi-parametric model

$$Y_t$$

- The time series data: $Y = (Y_1, Y_2, \ldots, Y_T)'$

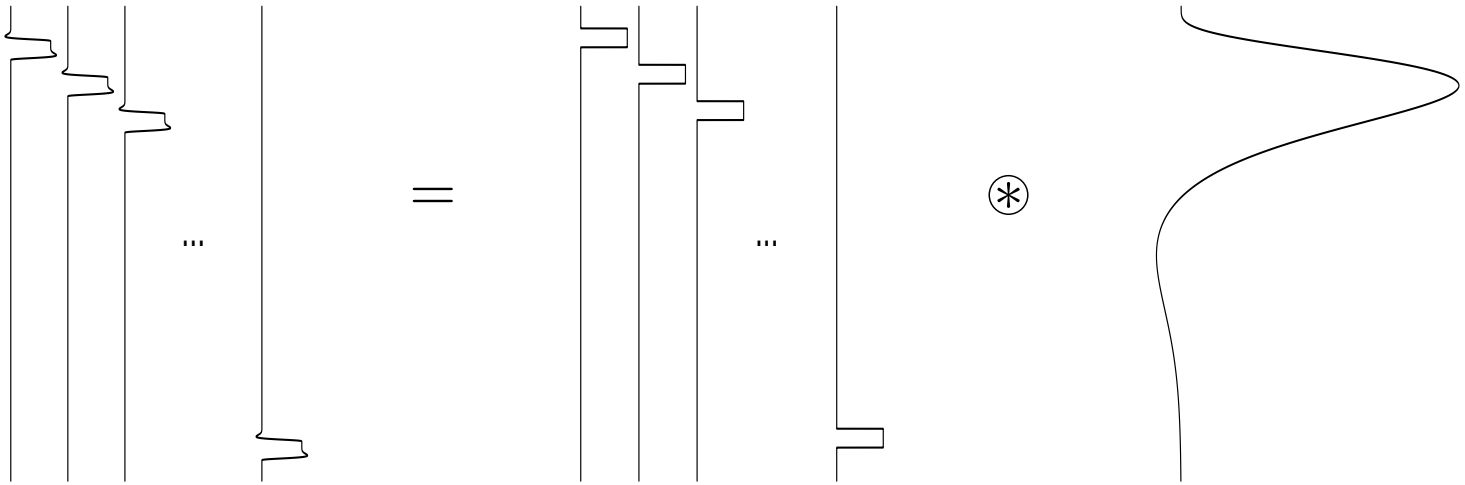


Bayesian semi-parametric model

$$Y_t = (\mathbf{X}\beta)_t$$

- **X** is the design matrix convolved with the HRFs

X = D $\circledast$ HRF

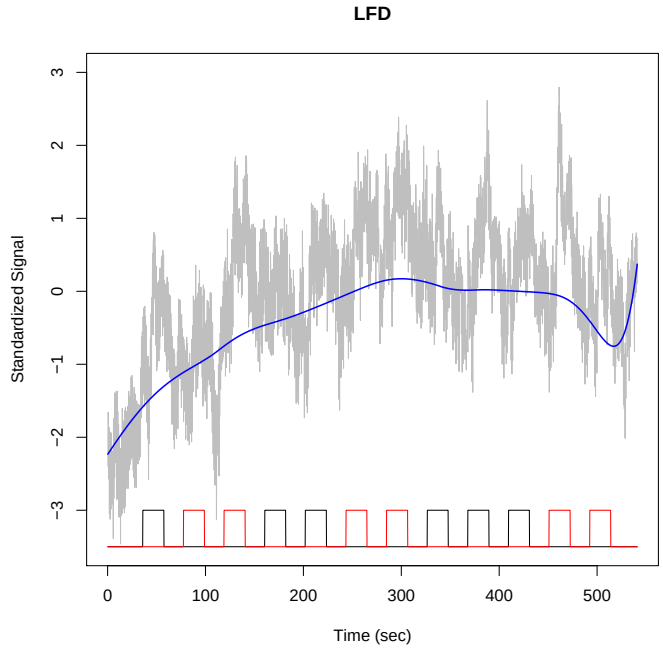


- HRFs modeled only w/canonical HRF

Bayesian semi-parametric model

$$Y_t = (\mathbf{X}\beta)_t + (\mathbf{V}\eta)_t$$

- $V\eta$ models this drift nonparametrically with B-spline basis
- Number and location of bases random
  - Use general birth-death MCMC algorithm (Stephens, 1998, Annals of Statistics)



Bayesian semi-parametric model

$$Y_t = (\mathbf{X}\beta)_t + (\mathbf{V}\eta)_t + \mathbf{R}_t^{\mathrm{T}}\phi_t$$

- $\mathbf{R}_t$ is the P-vector of lags at time t

$$\mathbf{R}_t = (R_{t-1}, \ldots, R_{t-P})^{\mathrm{T}}$$

where

$$R_t = Y_t - (\mathbf{X}\beta)_t + (\mathbf{V}\eta)_t$$

- ϕ_t is the P-vector of TVAR coefficients at time t

$$\phi_t = (\phi_{t1}, \ldots, \phi_{tP})^{\mathrm{T}}$$

Bayesian semi-parametric model

$$Y_t = (\mathbf{X}\beta)_t + (\mathbf{V}\eta)_t + \mathbf{R}_t^{\mathrm{T}}\phi_t + \varepsilon_t$$

- ε_t is the model error at time t:

$$\varepsilon_t \sim \mathrm{N}(0, \sigma_t^2)$$

Bayesian semi-parametric model

$$Y_t = (\mathbf{X}\boldsymbol{\beta})_t + (\mathbf{V}\boldsymbol{\eta})_t + \mathbf{R}_t^{\mathrm{T}}\boldsymbol{\phi}_t + \varepsilon_t$$

- Suitable priors placed on all parameters
  - flat prior for $\boldsymbol{\beta}$
  - $\boldsymbol{\eta} \sim \mathrm{MVN}(0, k\mathrm{I})$
    - k a constant
  - number of basis (knots): $\mathrm{Poisson}(5, 1)$
  - knot locations: uniform on time range of data
  - $\varepsilon_t \sim \mathrm{N}(0, \sigma_t^2)$
  - $\boldsymbol{\phi}_t$ and σ_t^2 discussed next slide

Priors continued

- $Q_t = \mathbf{R}_t^{\mathrm{T}} \phi_t + \varepsilon_t, \;\; t = P+1, \ldots, T$ cast into DLM framework
  - ϕ_t updated with standard forward-filtering, backward-smoothing algorithm for DLMs (Gibbs)

$$\phi_t = \phi_{t-1} + \mathbf{w}_t, \quad \mathbf{w}_t \sim \mathrm{MVN}(0, W_t)$$

  - Discount factor, δ_1 used for evolution variance W_t
  - Beta-gamma evolution model for σ_t^2 (Gibbs update):

$$\sigma_t^2 = \delta_2 \sigma_{t-1}^2 / \gamma_t$$

where

$$(\gamma_t \mid \mathcal{D}_t) \sim \mathrm{Beta}\left(\delta_2 n_{t-1}/2, (1-\delta_2) n_{t-1}/2\right)$$

and $\delta_2 \in (0, 1)$ acts as a discount factor.

Posterior Sampling

- β and η jointly updated via Gibbs
- knot locations and number of knots—BDMCMC
- locations further updated via MCMC

Harrison and West:

- Suggest values of the discount factors and TVAR order be selected a-priori by maximizing the log likelihood.
- Cannot do this in our model as the mean structure (i. e. $\mathbf{X}\beta + \mathbf{V}\eta$) is not know, a-priori

Posterior Sampling

- β and η jointly updated via Gibbs
- knot locations and number of knots—BDMCMC
- locations further updated via MCMC

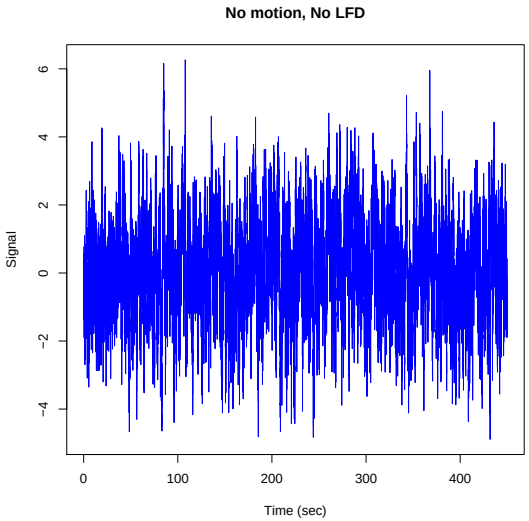
Harrison and West:

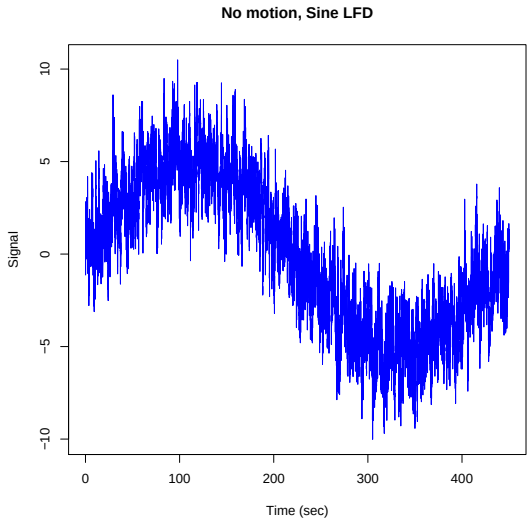
- Suggest values of the discount factors and TVAR order be selected a-priori by maximizing the log likelihood.
- Cannot do this in our model as the mean structure (i. e. $\mathbf{X}\beta + \mathbf{V}\eta$) is not know, a-priori

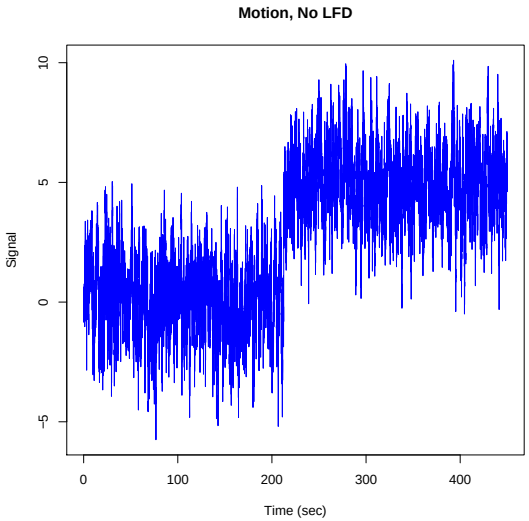
We take a different, fully Bayesian approach:

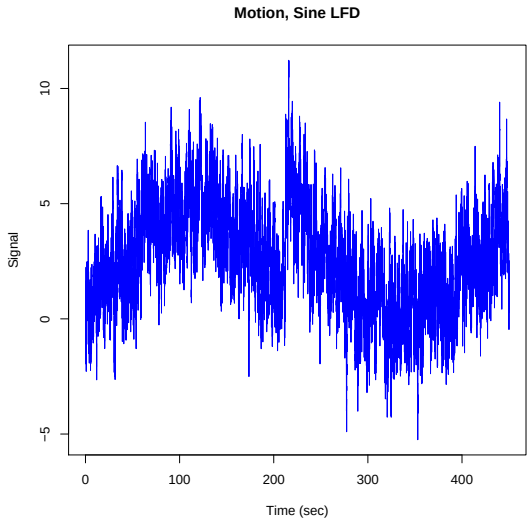
- Both discount factors given uniform priors on $(0.7, 1.0)$.
- The TVAR order P given a uniform prior on the integers $\{1, 2, \ldots, 80\}$
- All three updated via Metropolis-Hastings algorithm

Simulation Study









Simulation Study

- Simulated 2000 data sets for each of the four settings
- AR(3) model with coefficients 0.5, 0.2, 0.15
- Marginal variance = 1
- Canonical HRF, event related design with 10 events
- $\beta = 0, 0.5$
- Studied our proposed Bayesian model and the AR-IRLS
  - AR-IRLS proposed by Barker, et al. , *Biomedical Optics*, 2013
  - The model used by my collaborators and considered "state-of-the-art"
  - Best model I could find in the literature

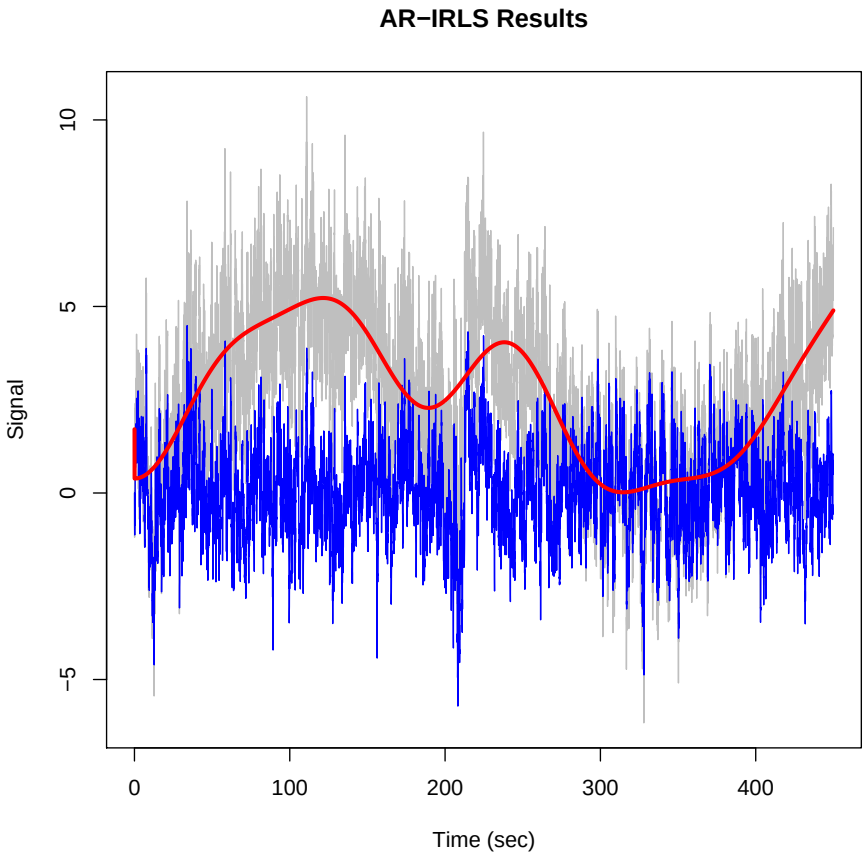
Simulation Results: $\beta = 0$

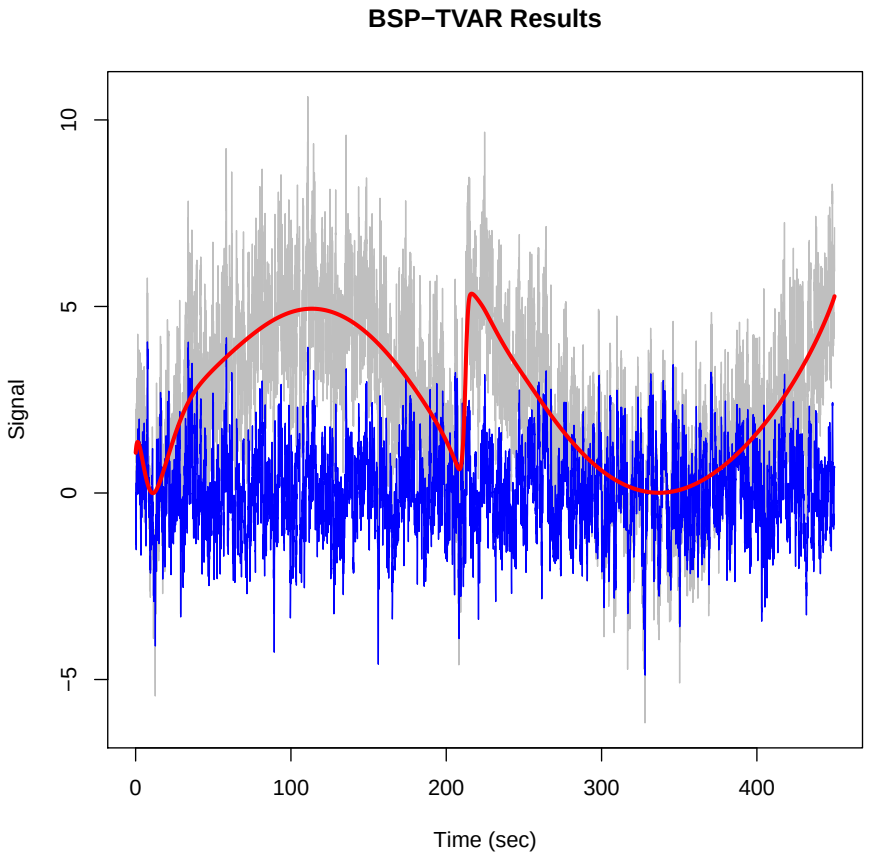
Model	Motion	LFD	Cover.	Bias	$\sqrt{MSE}$	Length
Bayesian	No	No	0.946	-0.0020	0.158	0.628
AR-IRLS			0.933	-0.0002	0.164	0.607
Bayesian	No	Yes	0.953	-0.0043	0.162	0.634
AR-IRLS			0.929	0.0234	0.164	0.609
Bayesian	Yes	No	0.952	0.0290	0.161	0.650
AR-IRLS			0.795	0.4738	0.505	1.240
Bayesian	Yes	Yes	0.934	0.0583	0.171	0.655
AR-IRLS			0.772	0.4883	0.521	1.253

Simulation Results: $\beta = 0.5$

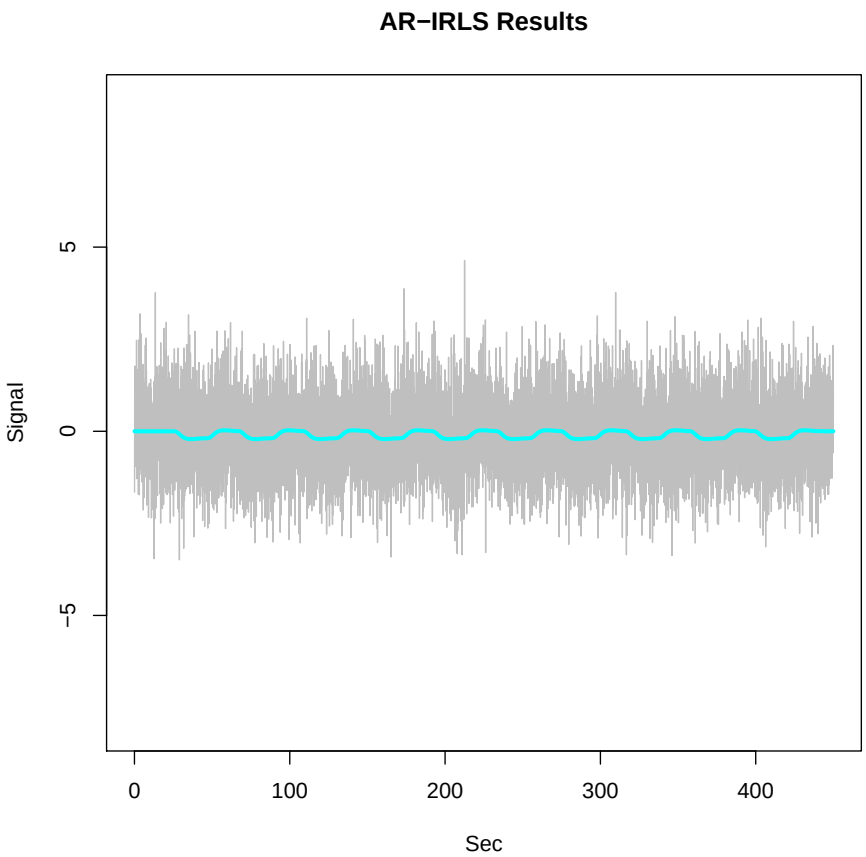
Model	Motion	LFD	Cover.	Bias	Power	$\sqrt{MSE}$	Length
Bayesian	No	No	0.956	0.0022	0.892	0.154	0.629
AR-IRLS			0.935	0.0003	0.886	0.164	0.605
Bayesian	No	Yes	0.941	-0.0047	0.861	0.165	0.634
AR-IRLS			0.929	0.0234	0.913	0.164	0.609
Bayesian	Yes	No	0.954	0.0299	0.903	0.161	0.649
AR-IRLS			0.796	0.4719	0.979	0.503	1.239
Bayesian	Yes	Yes	0.927	0.0547	0.916	0.178	0.654
AR-IRLS			0.779	0.4892	0.989	0.521	1.252

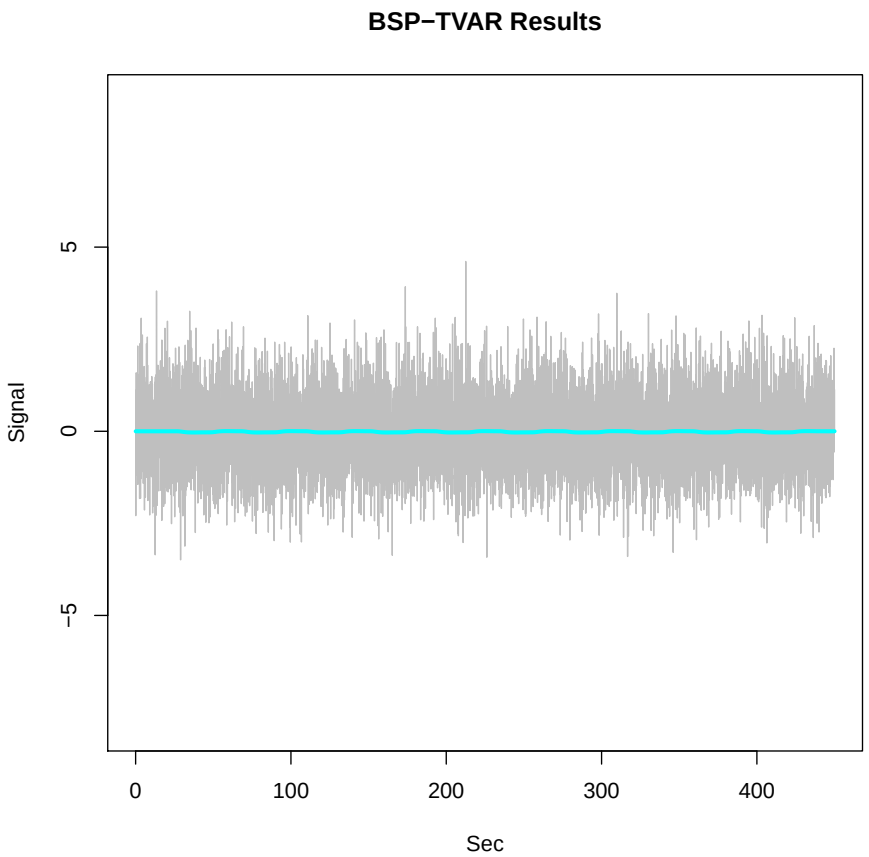
Simulation Results: Motion, Sine LFD, $\beta = 0$





Simulation Results: Motion, Sine LFD, $\beta = 0$

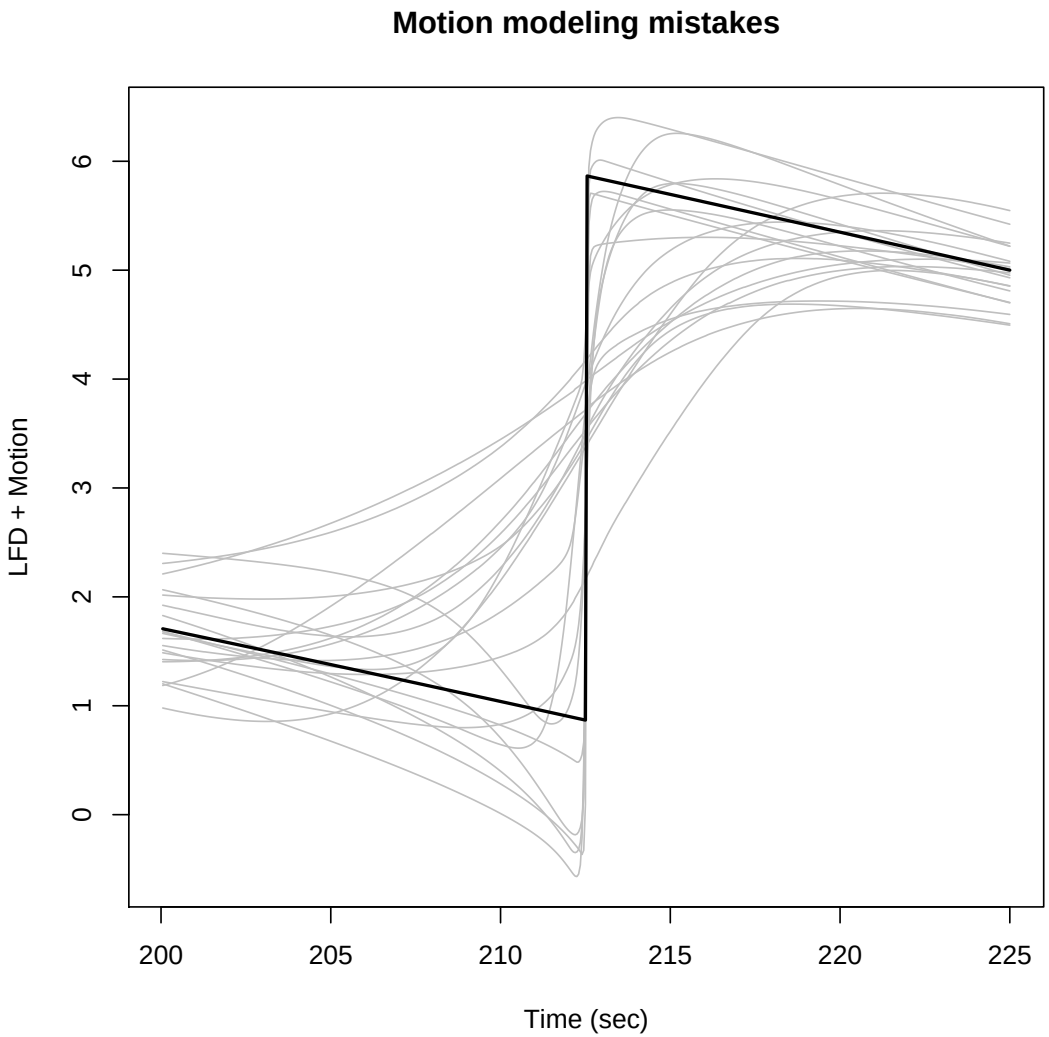




$\hat{\beta} = -0.21, \quad 95\%\ \text{CI} = (-0.53,\ 0.11)$

$\widehat{\text{E}(\beta \mid y)} = -0.03, \quad 95\%\ \text{Cred.Int.} = (-0.35,\ 0.29)$

Motion not always captured correctly



Priors for B-splines

1. Poisson prior on number of splines
2. Uniform prior on locations
   - Can result in poor fit to motion

- Homogeneous Poisson process has properties (1) and (2)

- Perhaps an inhomogeneous Poisson process prior

New prior for B-splines, simple case: 1 shift motion artifact

- Let m_1 denote the location of the motion
- Let $m_1 \in (m_1 - \delta, m_1 + \delta]$
- Define a piecewise homogeneous Poisson process on the three intervals

$$(0, m_1 - \delta], \quad (m_1 - \delta, m_1 + \delta], \quad (m_1 + \delta, T)$$

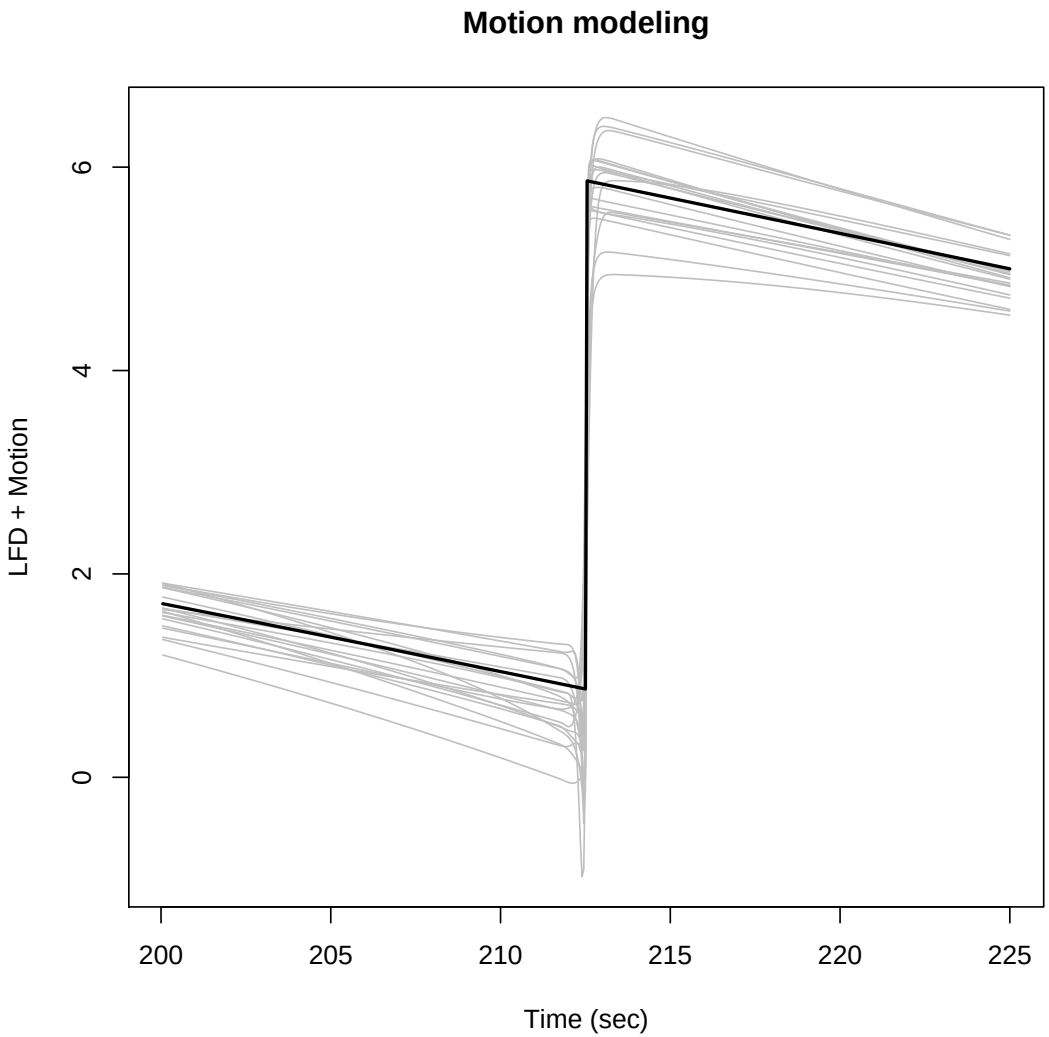
with intensity functions

$$\begin{aligned} \lambda_0 &= \frac{r}{(m_1 - \delta) + (T - m_1 - \delta)} \\ \lambda_1 &= \frac{r_m}{2\delta} \\ \lambda_2 &= \frac{r}{(m_1 - \delta) + (T - m_1 - \delta)} \end{aligned}$$

New prior for B-splines, simple case: 1 shift motion artifact

- The above intensity functions define a piecewise homog. Poisson process with rate r_m on $(m_1 - \delta, m_1 + \delta]$ and total rate r on $(0, m_1 - \delta] \cup (m_1 + \delta, T)$
- The rate r is shared propotionally to the length of the two intervals
- Take r moderately size, say 5 or 10.
- $r_m = 4$
  - For cubic B-splines, 4 bases that coalesce results in a discontinuity at the point of coalescence.
- Requires user input to define motion locations
- δ taken to be 1 second

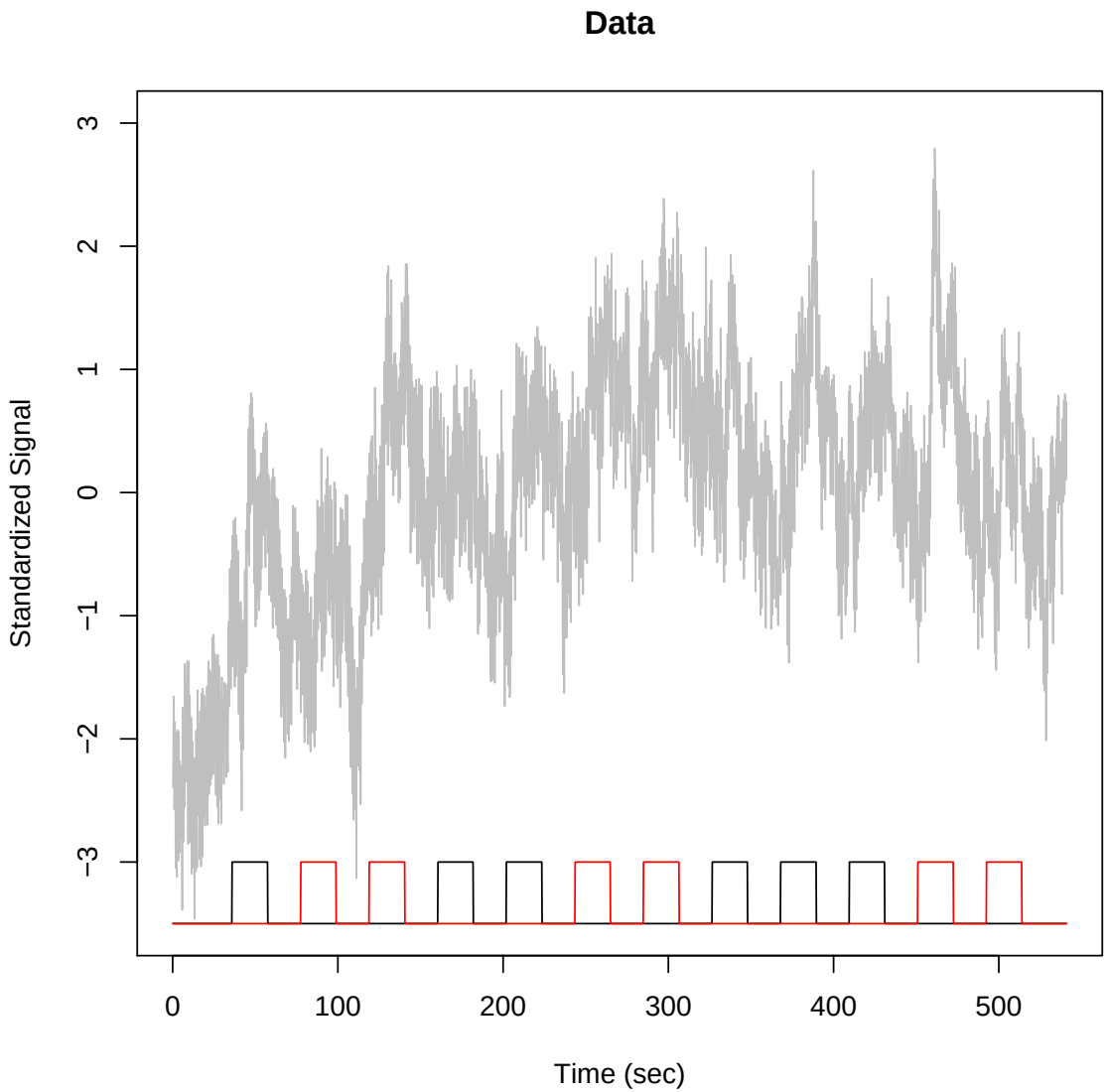
Better motion modeling via prior



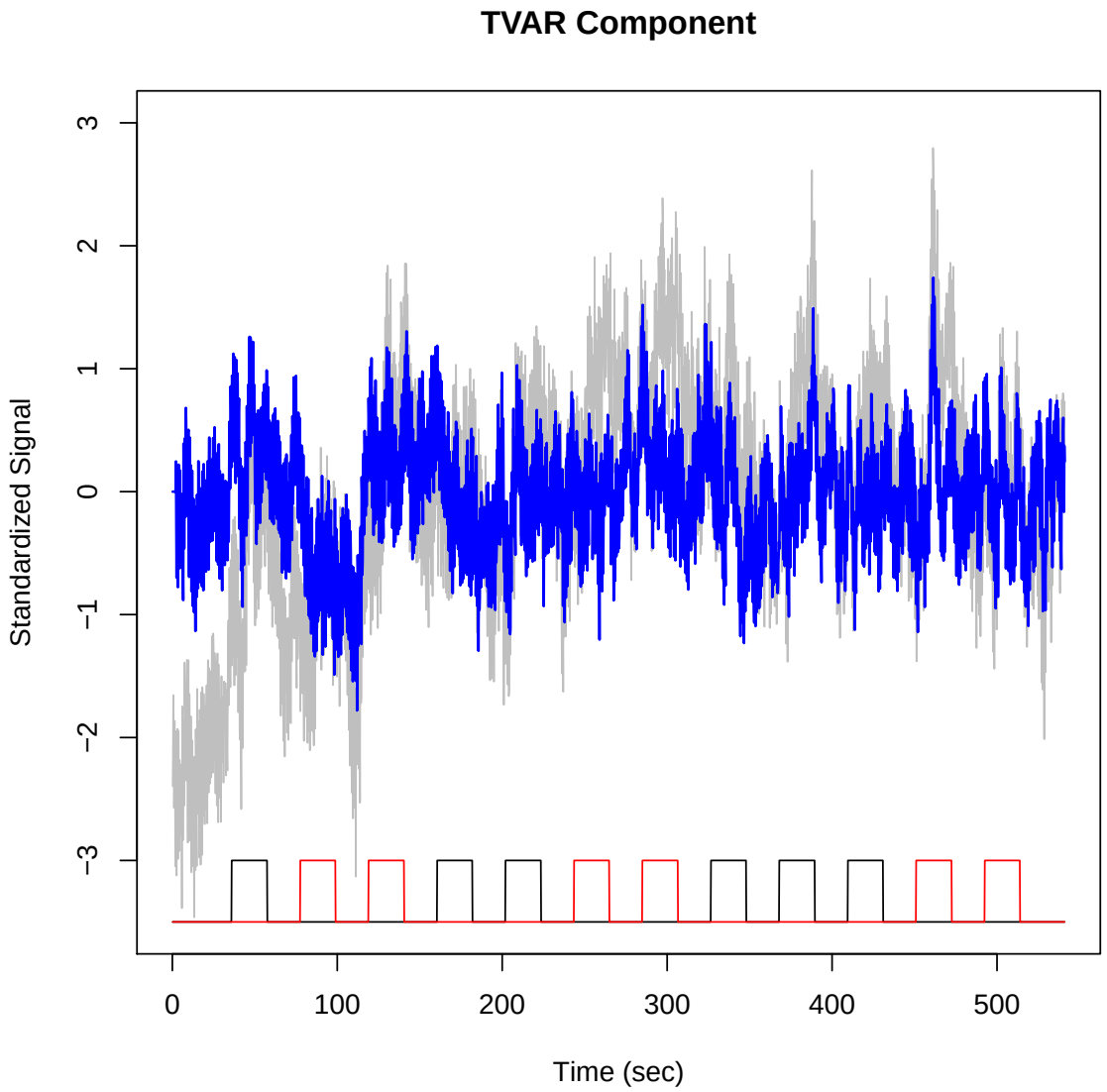
Simulation comparison, $\beta = 0.5$

Model	Motion	LFD	Cover.	Bias	Power	$\sqrt{MSE}$	Length
Homog. PP	Yes	Yes	0.927	0.0547	0.916	0.178	0.654
Inhomog. PP			0.948	0.0504	0.932	0.163	0.637

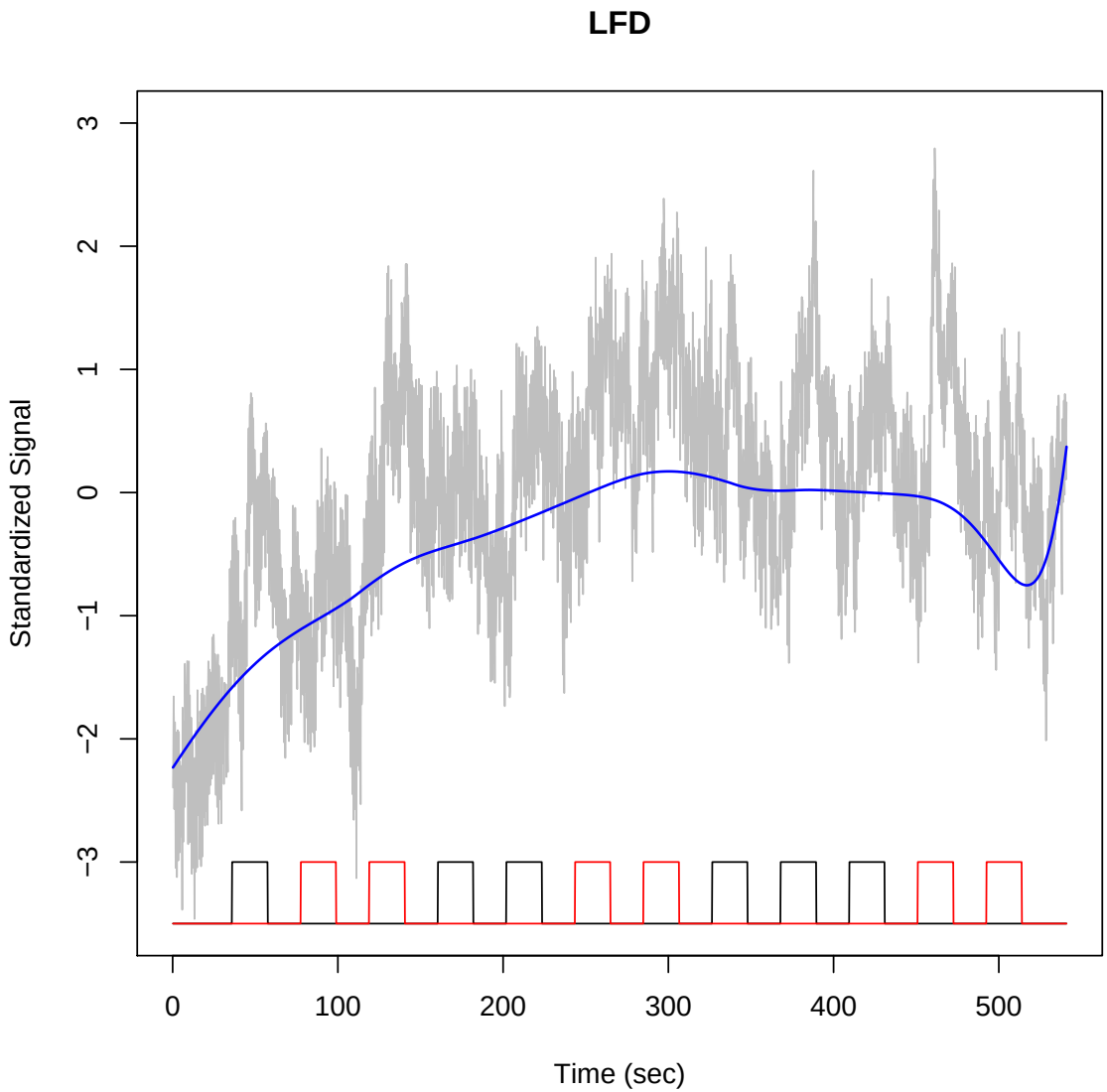
Data Analysis: Example 1



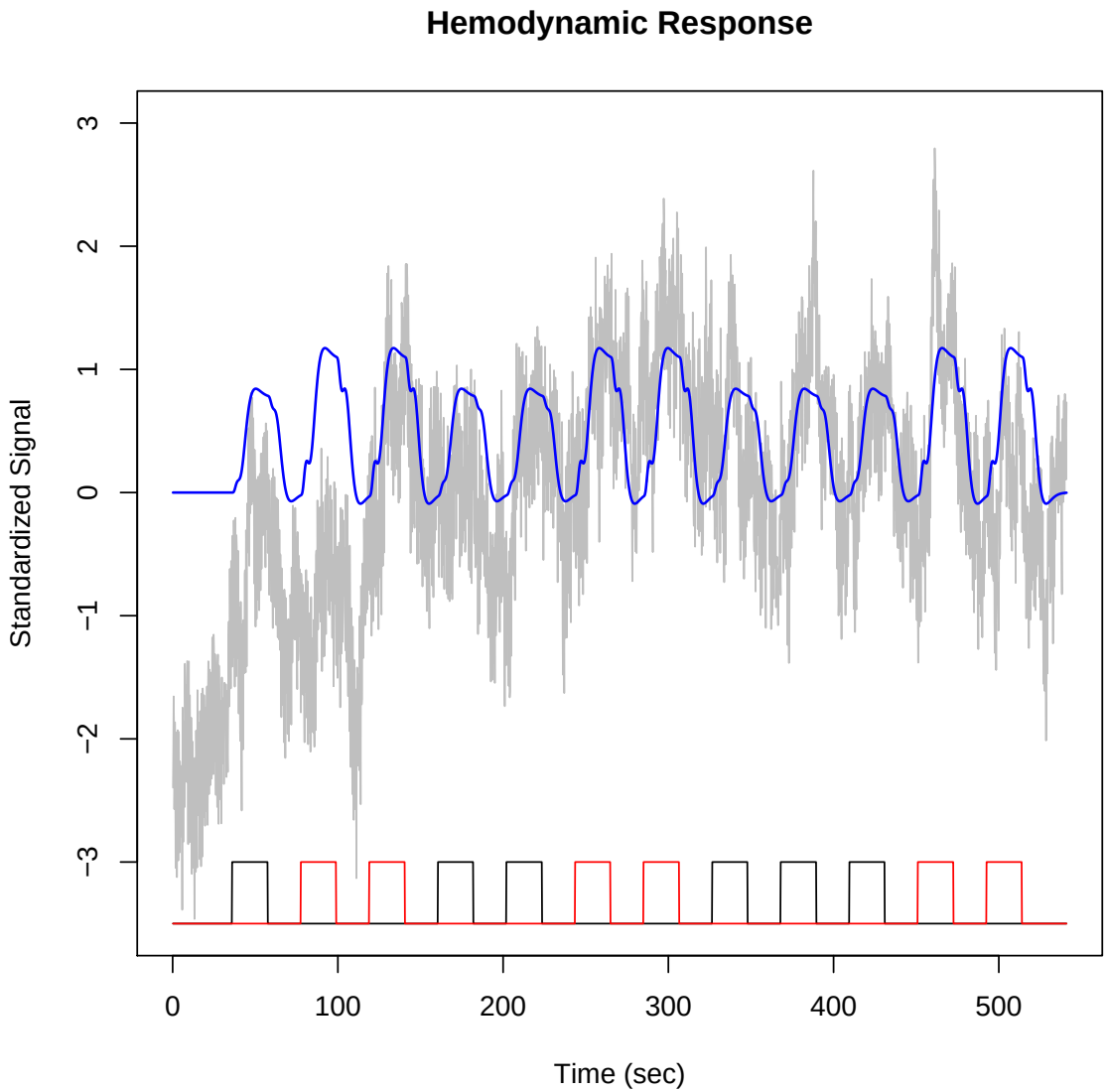
Data Analysis: Example 1



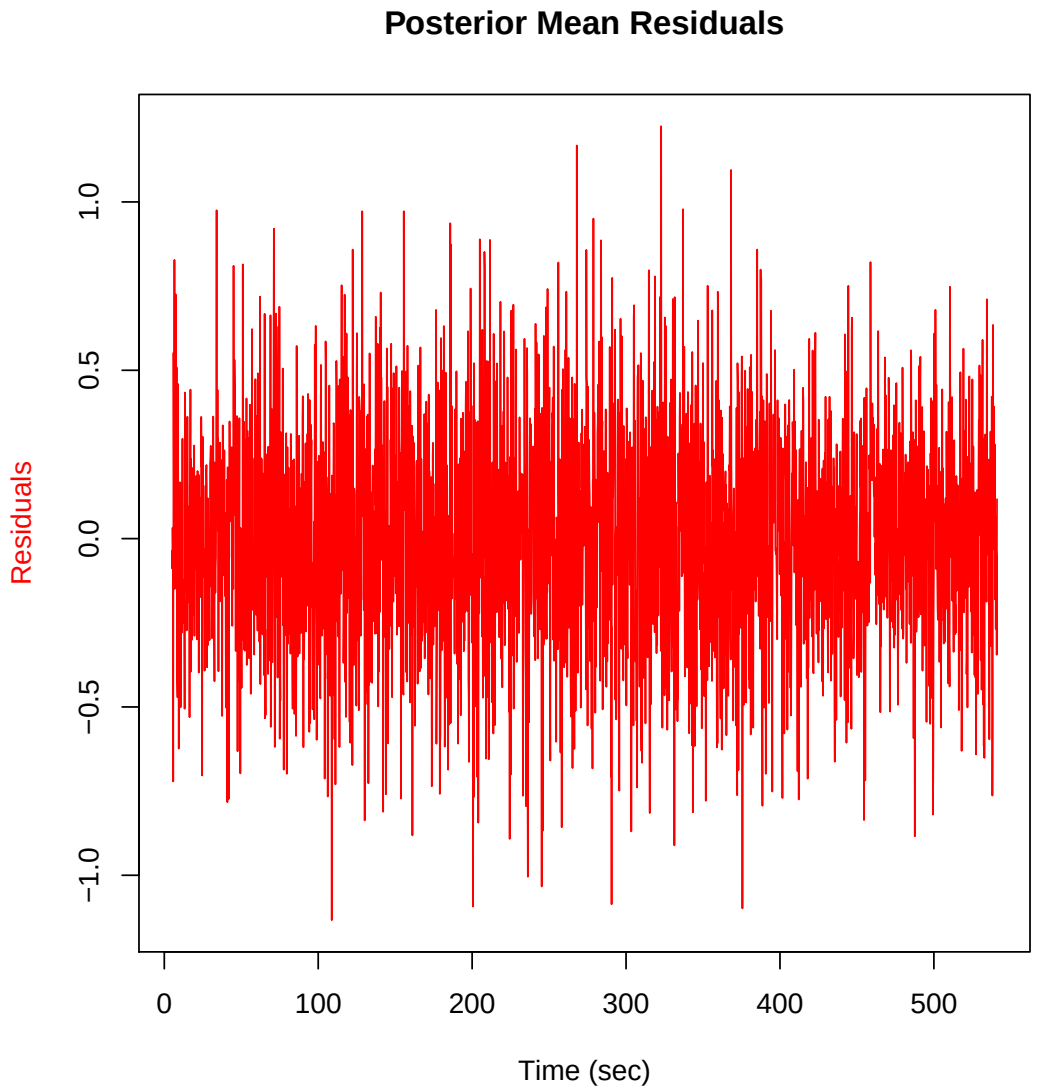
Data Analysis: Example 1



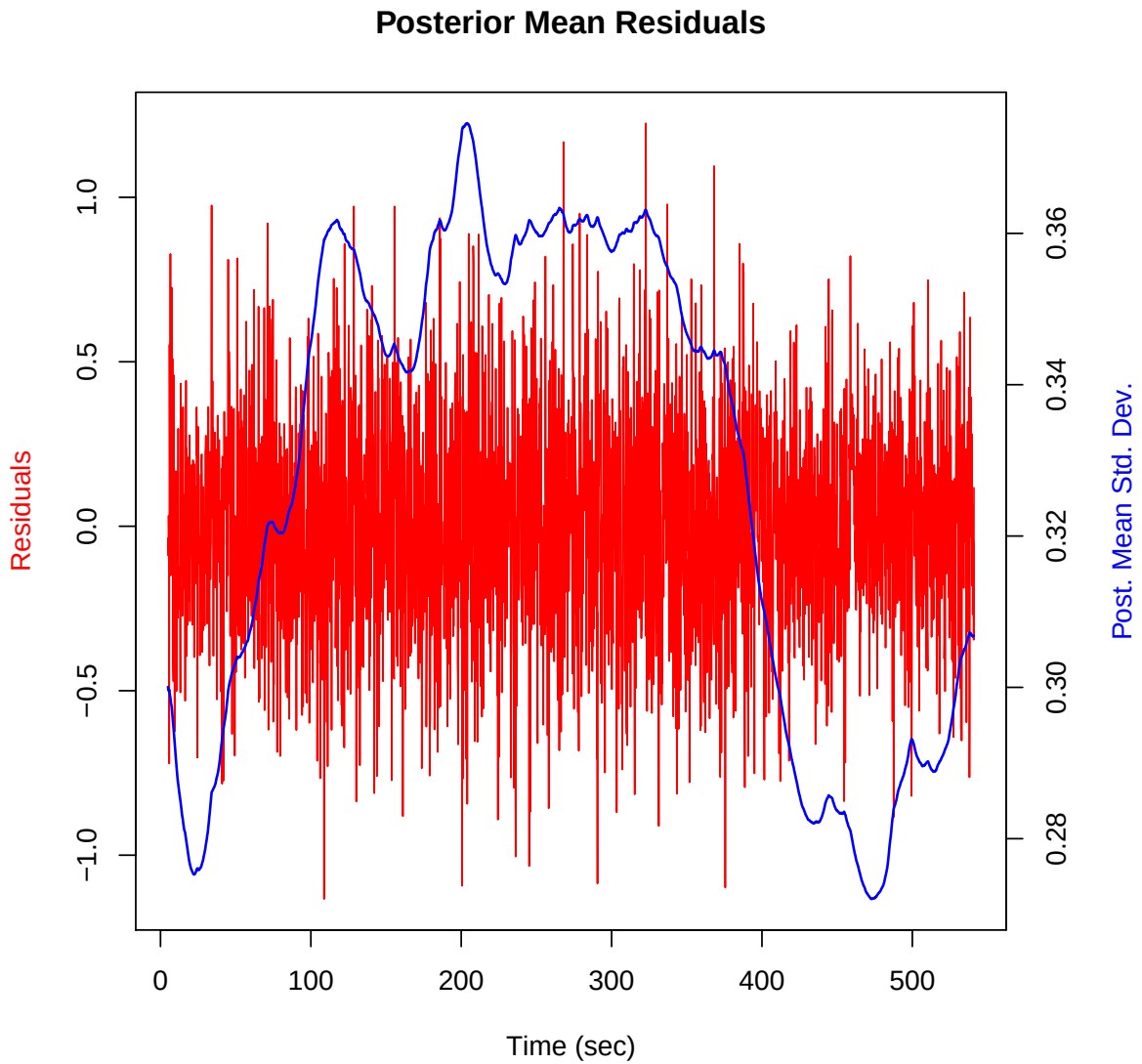
Data Analysis: Example 1



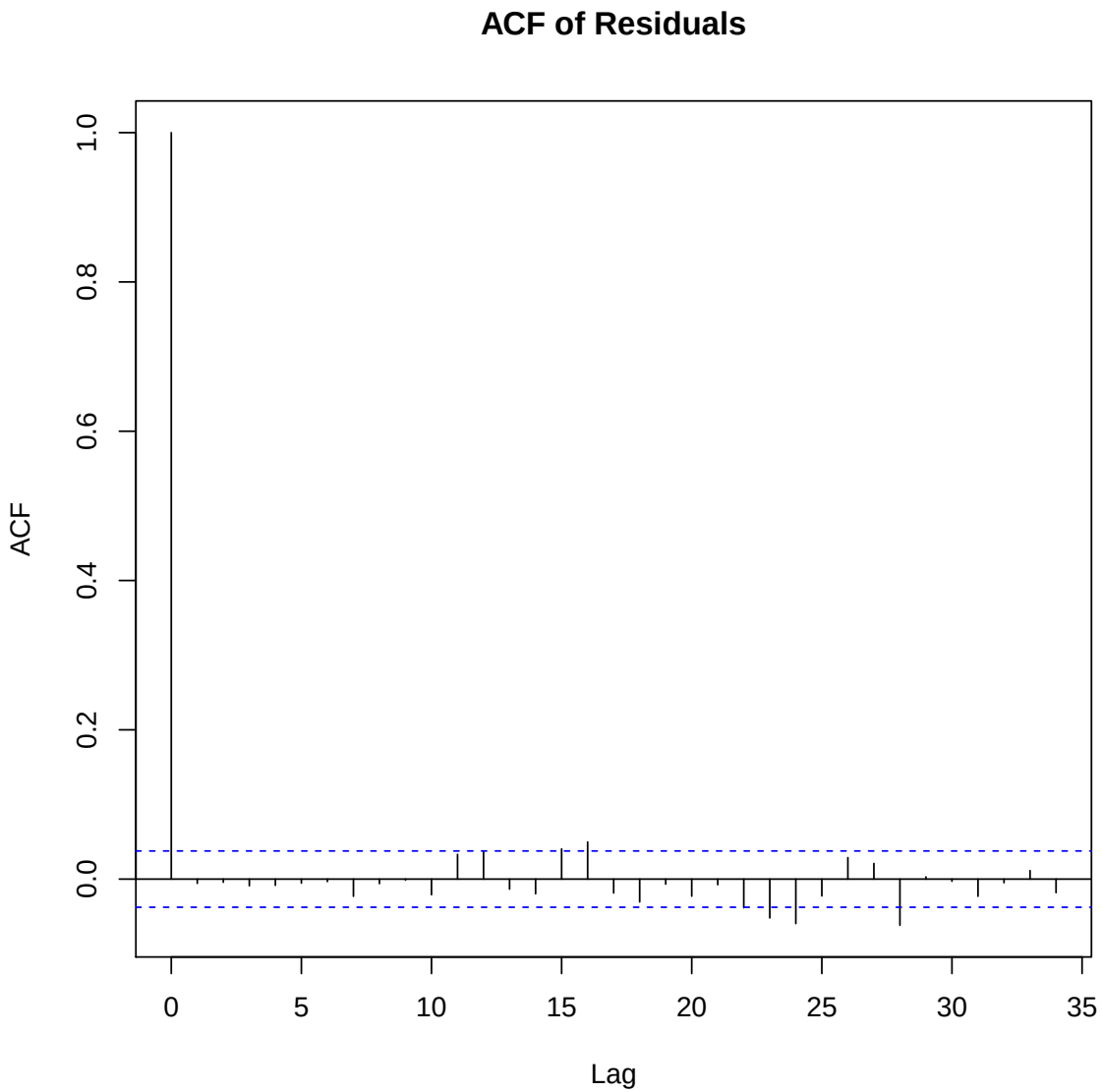
Data Analysis: Example 1



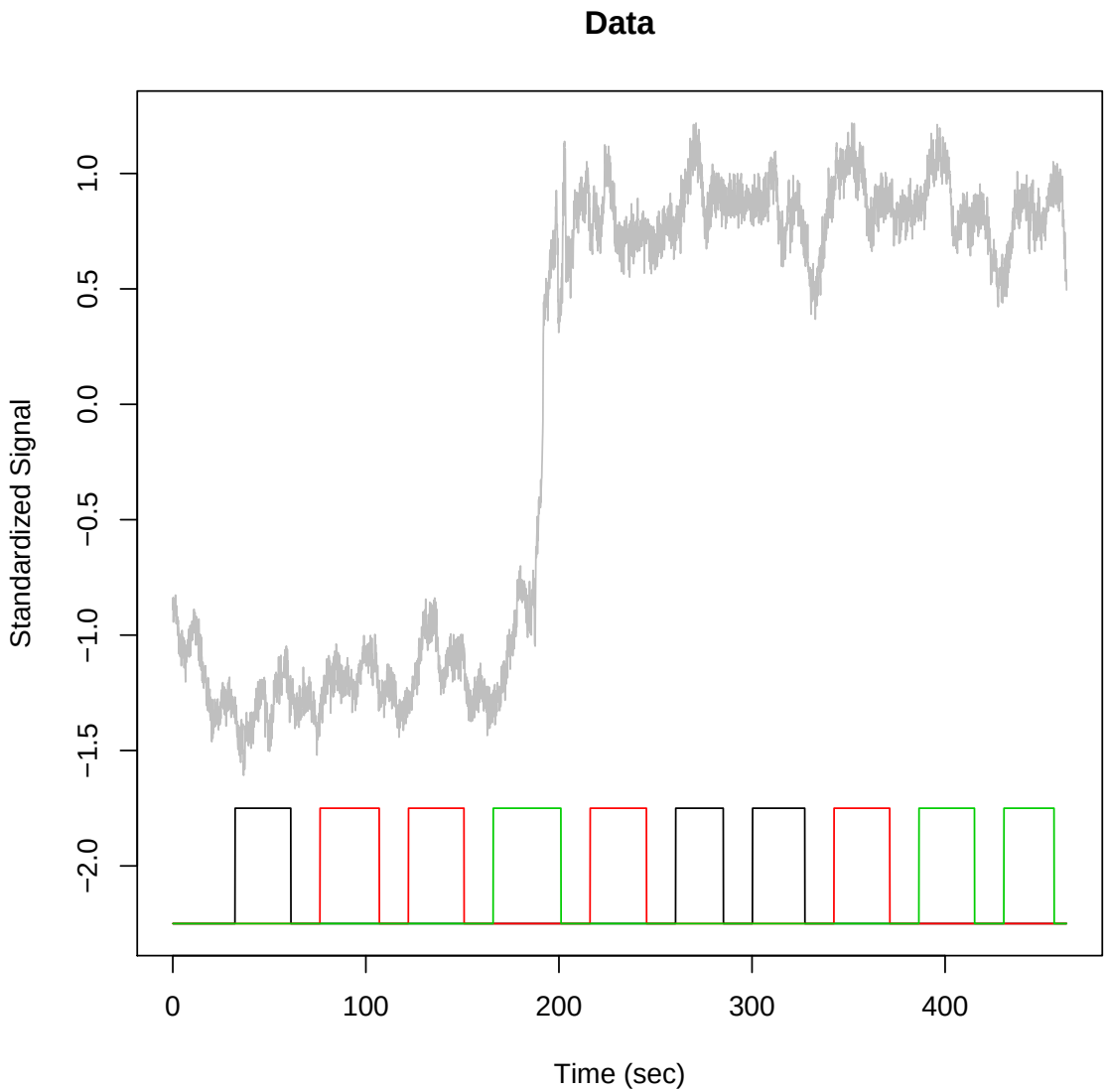
Data Analysis: Example 1



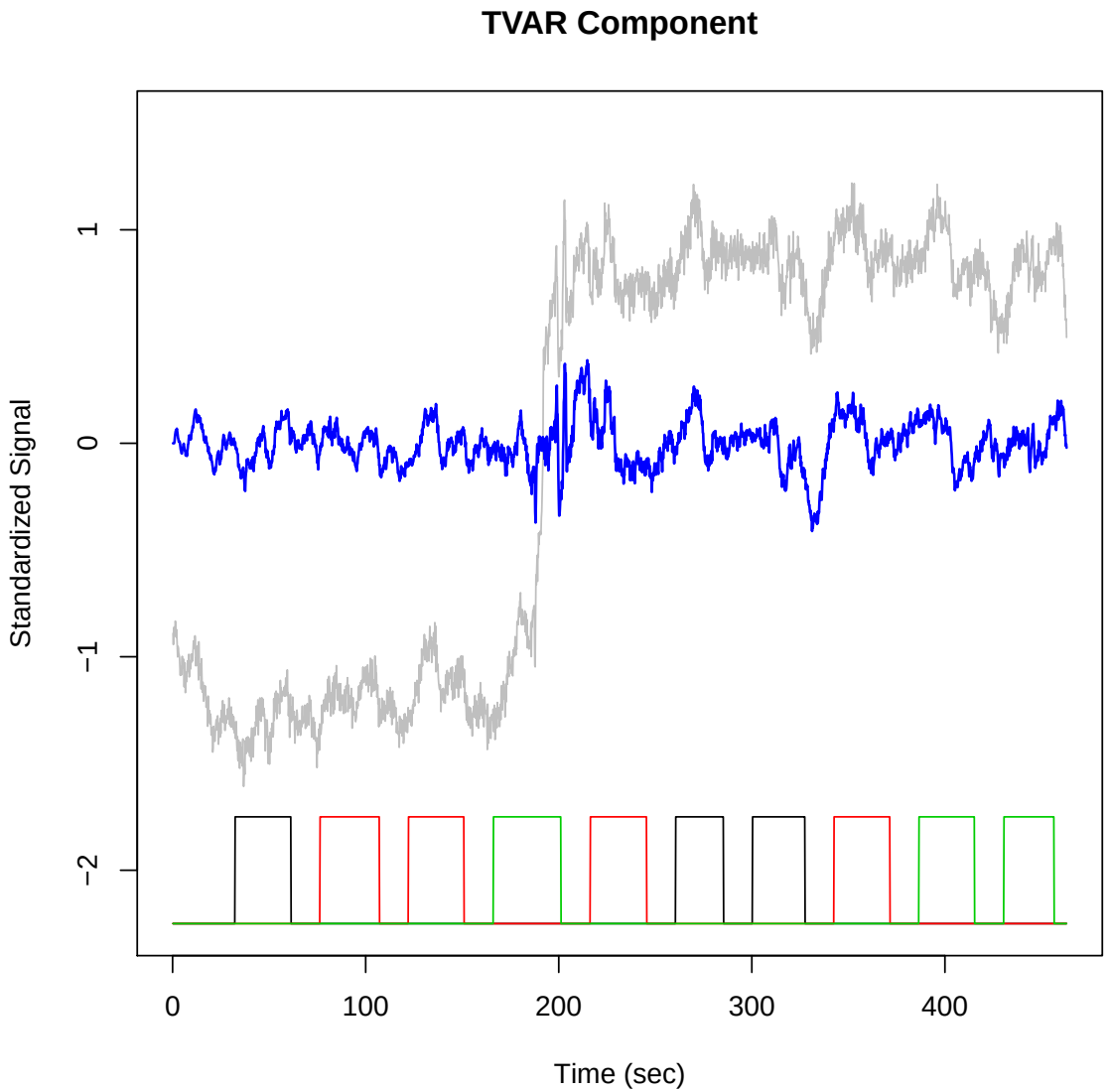
Data Analysis: Example 1



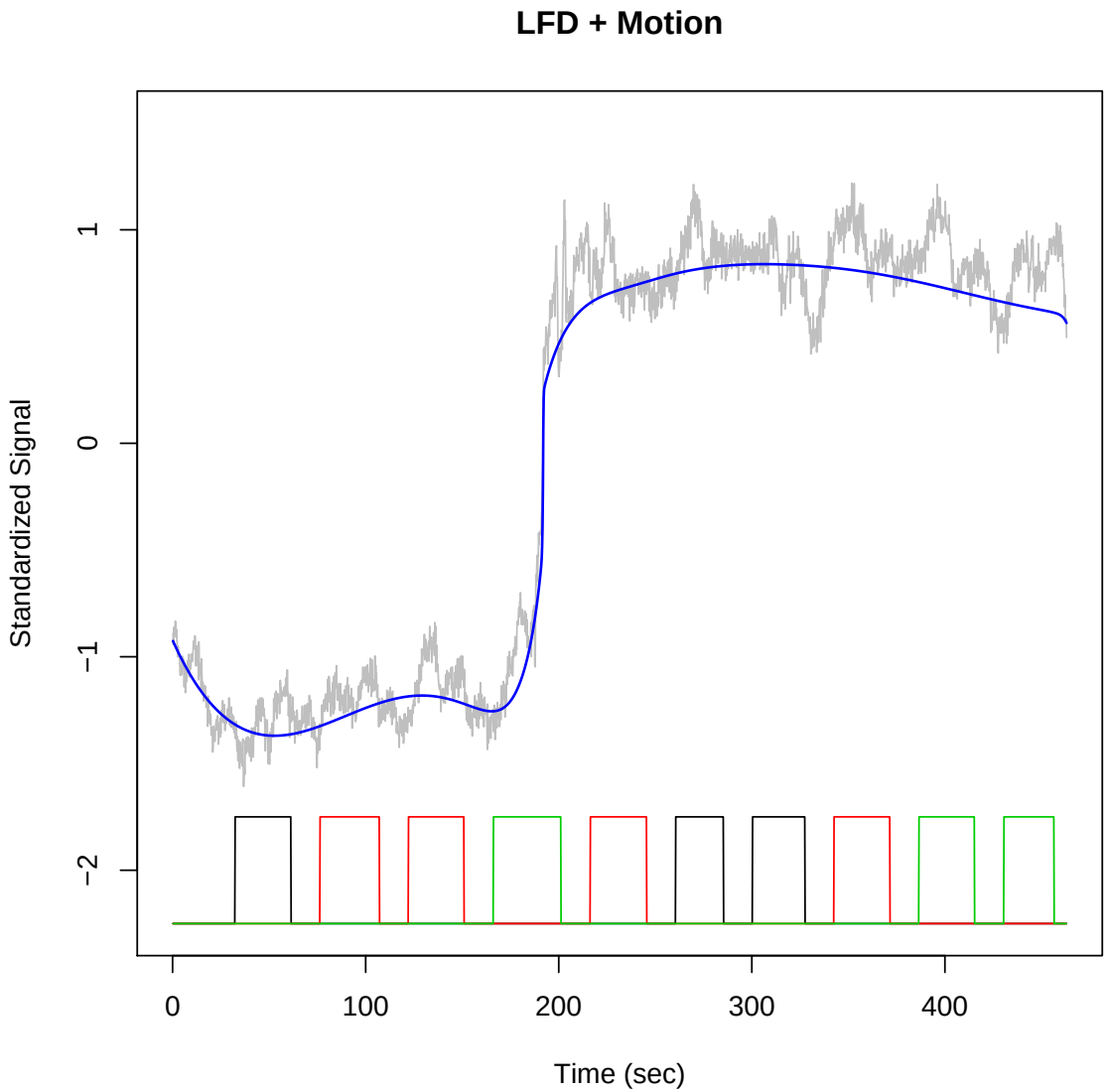
Data Analysis: Example 2



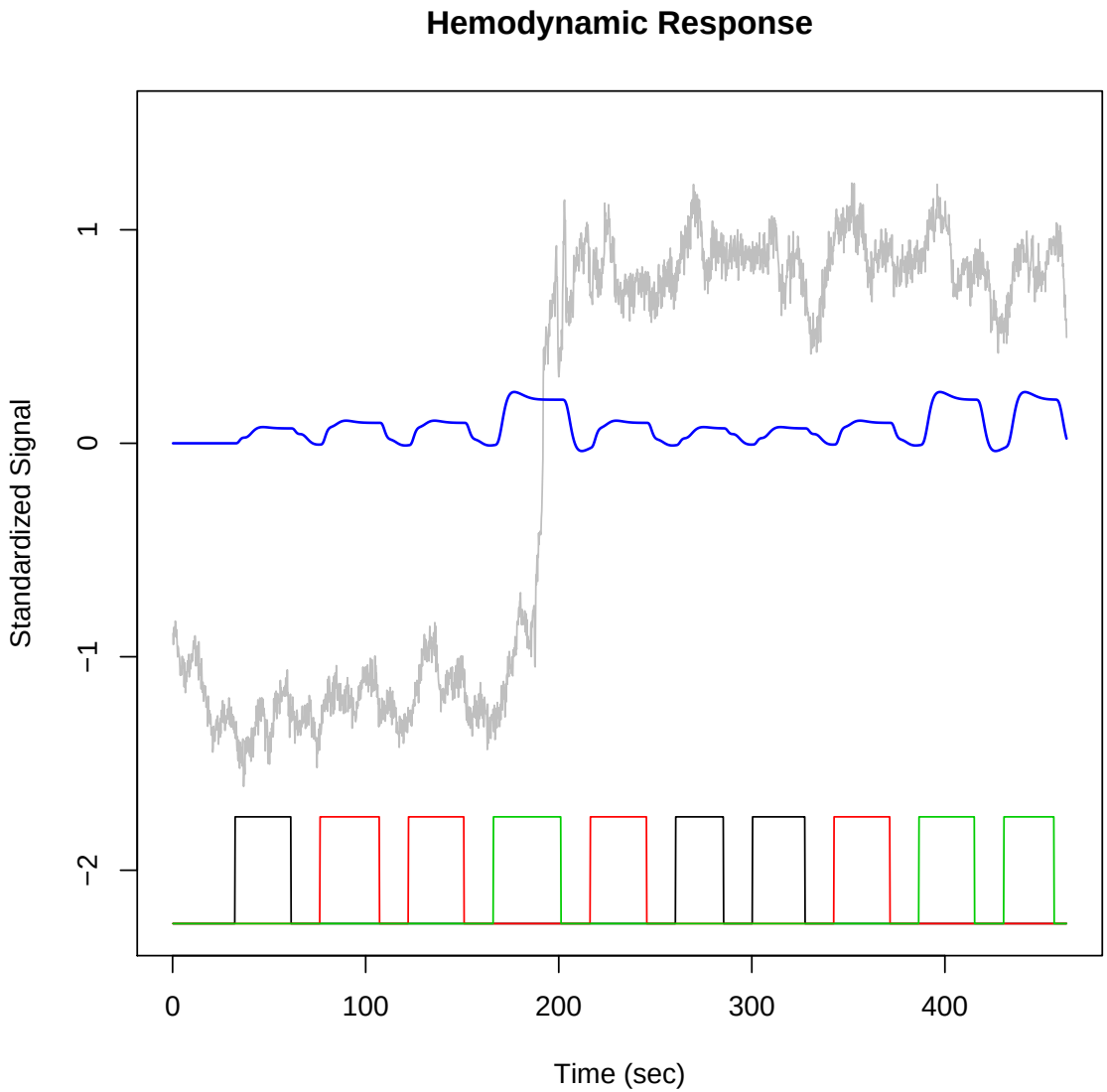
Data Analysis: Example 2



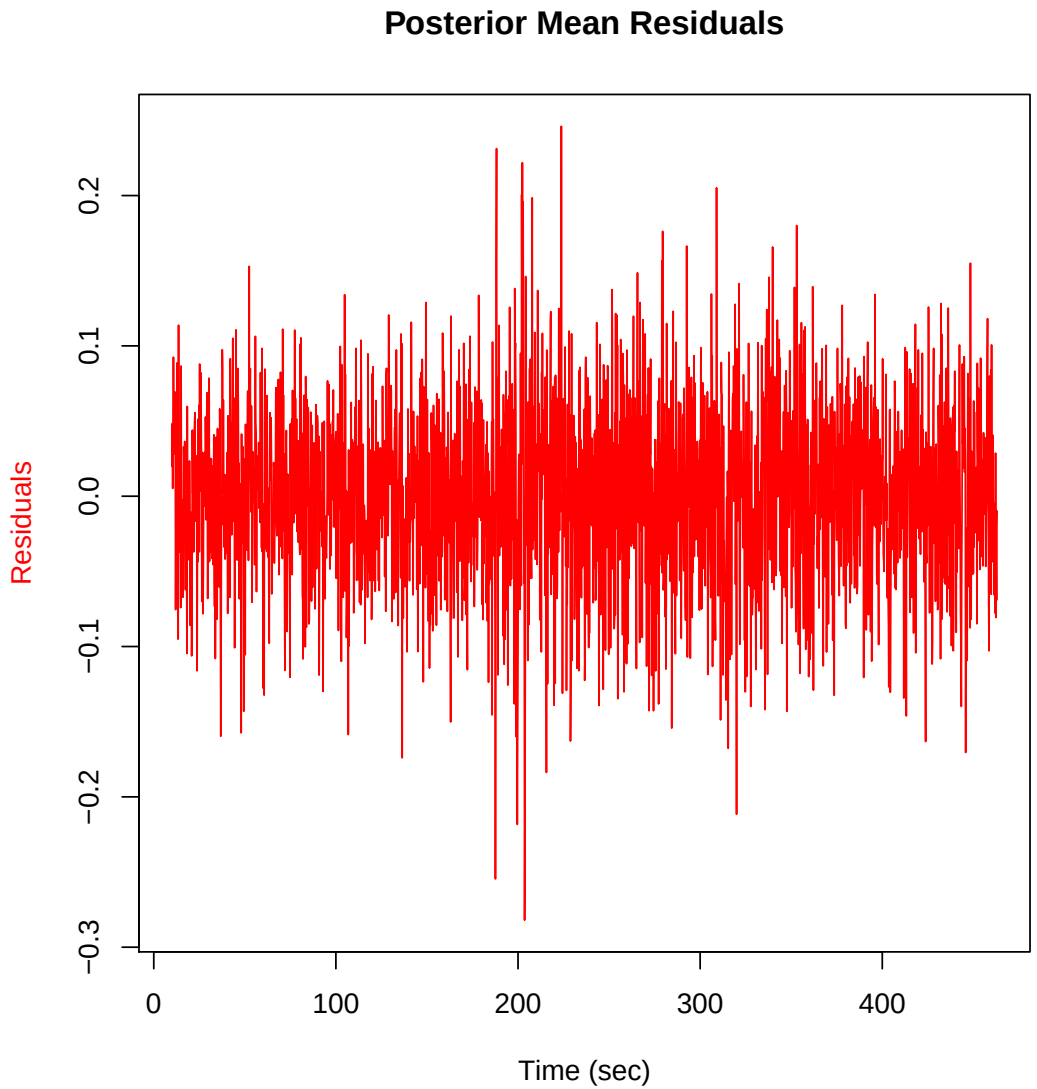
Data Analysis: Example 2



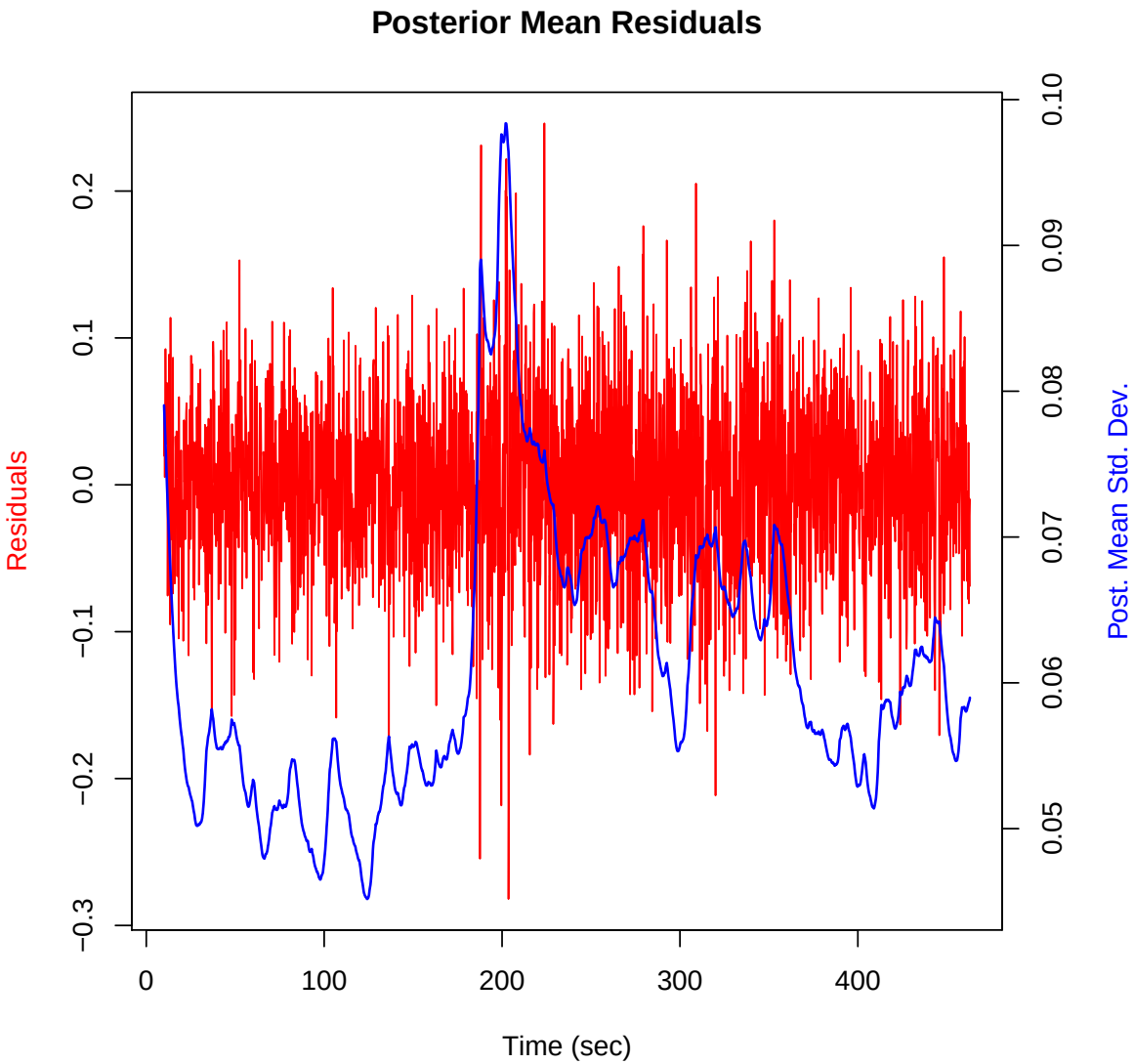
Data Analysis: Example 2



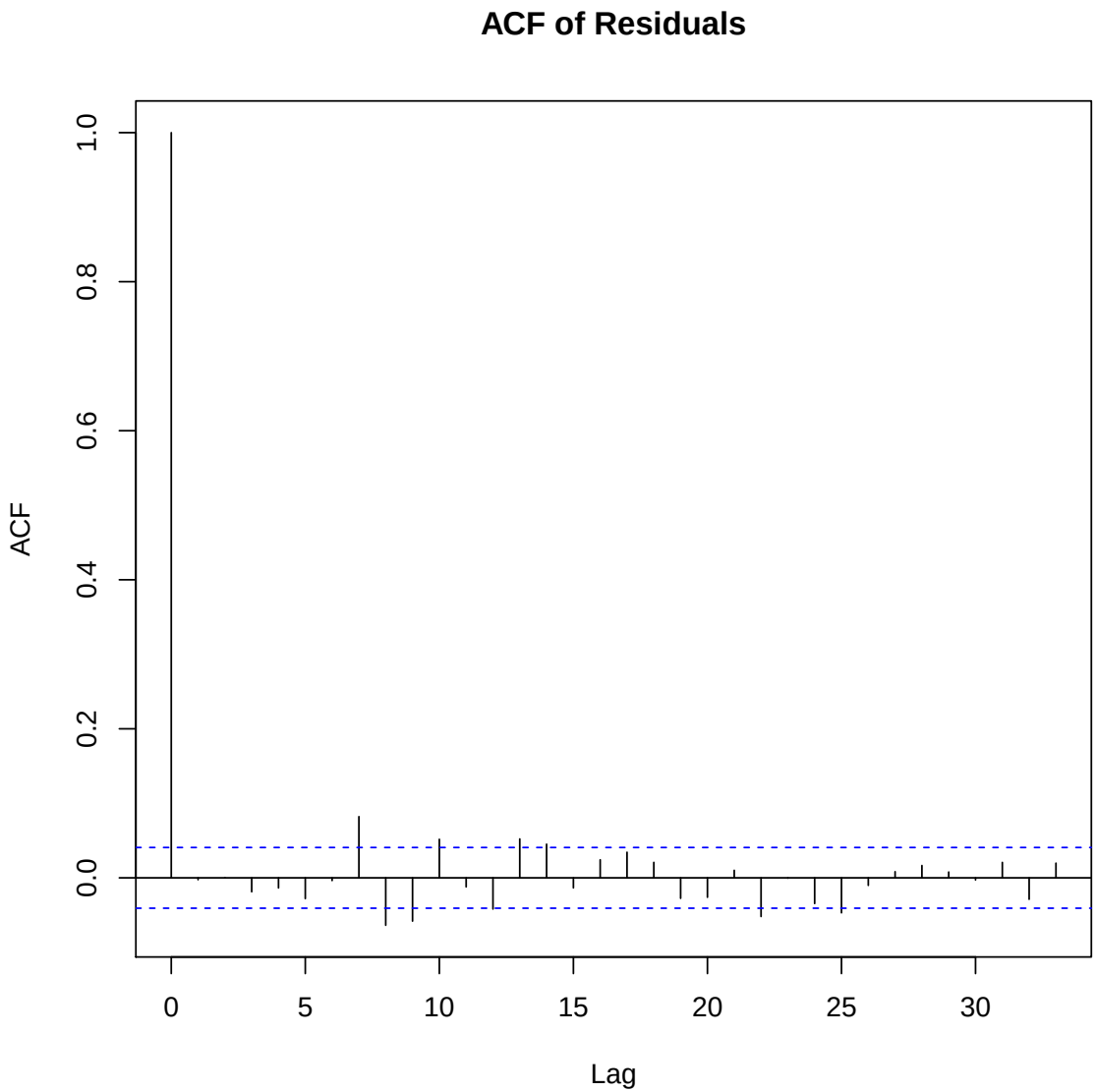
Data Analysis: Example 2



Data Analysis: Example 2



Data Analysis: Example 2



Conclusions/Future Work

- Showed BSP-TVAR model outperforms AR-IRLS
  - Especially in the presence of motion artifacts
- Future work
  - Modeling of the HRF
  - Bivariate TS model
    - deoxy. hem. may be important
    - may help overall model fit
  - Spatial modeling of the array of detectors
  - Group model
    - Inclusion of clinical covariates

References

1. Barker, J. W. Aarabi, A., Huppert, T. J. (2013), Autoregressive model base algorithm for correcting motion and serially correlated errors in fNIRS, *Biomedical Optics Express*, **4**(8), DOI:10.1364/BOE.4.001366
2. Prado, R. and West, M. (2010) Time Series Modeling, Computation and Inferenence. Chapman Hall/CRC, Boca Raton.
3. Stephens, M. (2000), Bayesian analysis of mixture models with an unknown number of components—an alternative to reversible jump methods, *Annals of Statistics*, **28**(1), 40–74.
4. West, M. and Harrison, J. (1997), Bayesian Forecasting and Dynamic Models, 2nd edition. Springer, New York.