# Functional Data Analysis for Accelerometer Data

**Jeffrey S. Morris**

University of Texas M.D. Anderson Cancer Center, Houston, TX

**Steve Gortmaker**

Harvard School of Public Health

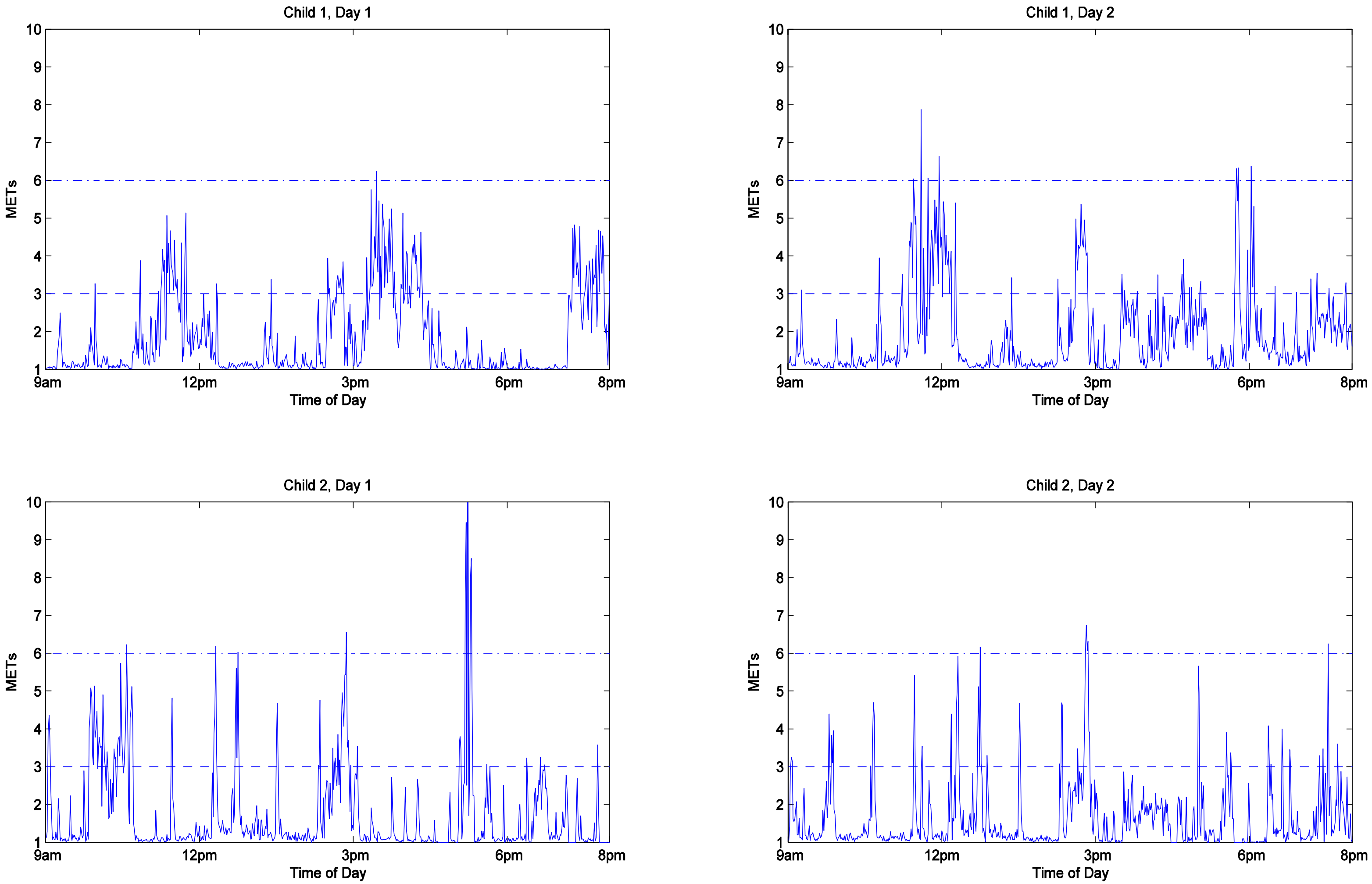
Boston, MA

Other collaborators: **Louise Ryan, Brent Coull, and Cassandra Arroyo**

# Accelerometers

- **Accelerometers:** small motion sensors that digitally record minute-by-minute activity levels
  - Increasingly used in large surveillance studies and intervention trials to objectively monitor activity
  - Children, older adults, Alzheimer patients
  - Revolutionizing the field of physical activity research
- Yield complex functional data
  - Activity level profiles

# Some Accelerometer Profiles



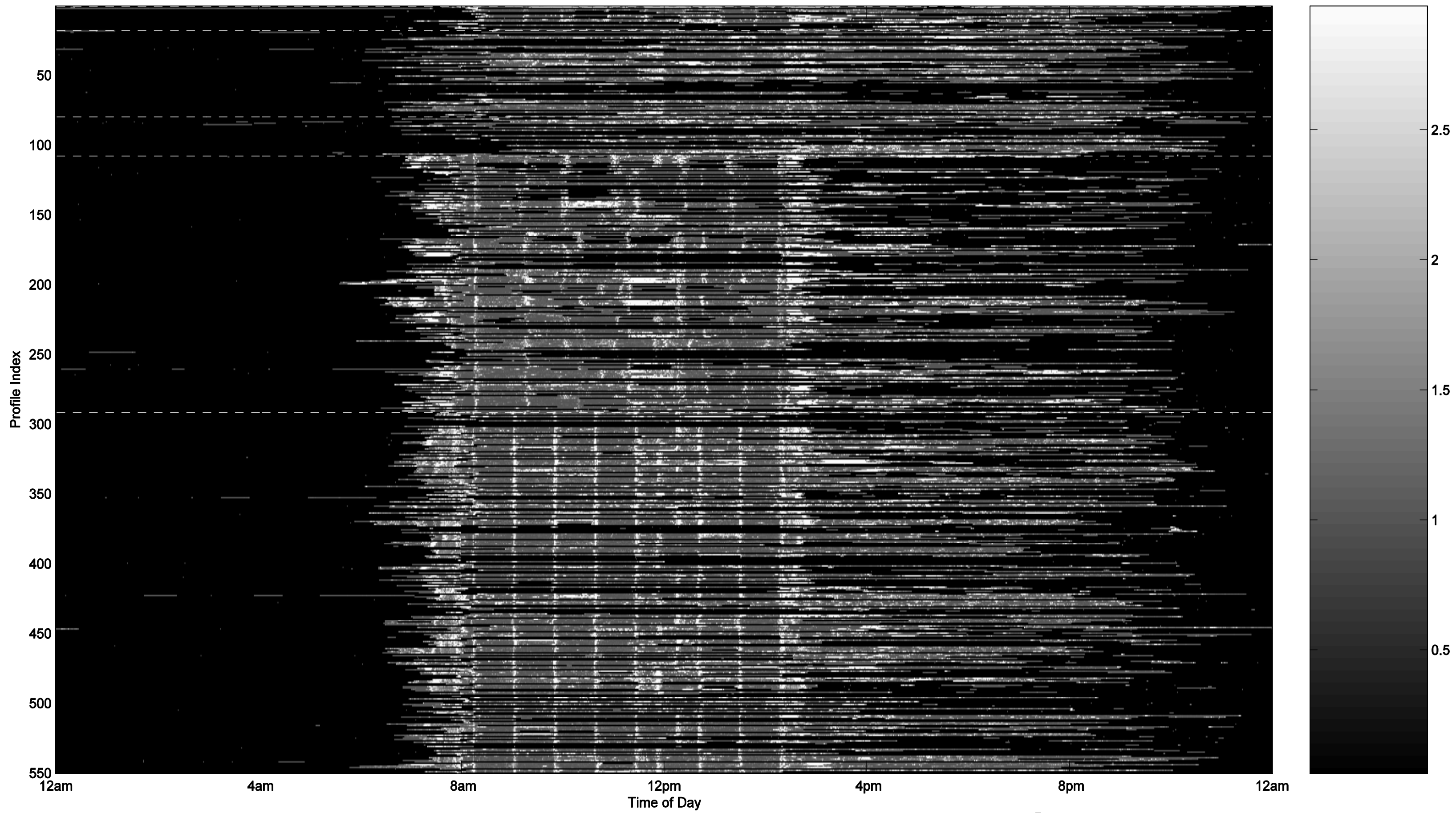
# Analyzing Accelerometer Data

- **Major challenge:** How to analyze these data?
- **Standard strategy:** Extract simple **summary measures** from the profiles, then analyze these
  1. Average daily activity level
  2. Proportion of time with activity levels corresponding to sedentary (<3 MET), moderately vigorous (3-6 MET), and vigorous (>6 MET) activities
- **Benefit:** simplifies data into recognized form.
- **Limitations** of these approaches:
  - Do not make full use of information in functional data
  - Cannot effectively deal with incomplete profiles

# Case Study: Planet Health

- **Planet Health: Boston-area school based intervention**
  - **Subset of children in study wore accelerometer for 4-8 days**
  - **550 profiles from 112 children from 5 schools, 1440 measurements/profile** **(Heatmap)**
- **Lots of *missing data***
  - **Measurement coded as *missing* for time periods for which child was inactive for at least 30 consecutive minutes**
  - **We focus on profiles >50% complete from 9am-8pm**
    - **292 profiles from 106 children**

# Heatmap




Accelerometer Data

# Accelerometer Data

- Other measurements:
  - Child-level covariates: school, race, gender, age, weight, height, BMI, triceps skinfolds, avg hrs of TV/day
  - Day-level covariates: day of week, calendar date
- Questions of Interest:
  1. How do activity levels vary throughout day, across schools, across different days of the week, over time from early to late Spring, and across various child-level covariates?
  2. What proportion of variability in activity levels is from day-to-day and child-to-child? How many days per child?

# Linear Mixed Models

## Linear Mixed Model (Laird and Ware, 1982):

$$\underbrace{Y}_{N\times 1} = \underbrace{X}_{N\times p}\overbrace{\beta}^{p\times 1} + \underbrace{Z}_{N\times m}\overbrace{u}^{m\times 1} + \underbrace{e}_{N\times 1}$$

$$u \sim N(0, \overbrace{P}^{m\times m})$$
$$e \sim N(0, \underbrace{R}_{N\times N})$$

- **Fixed effects** part, $X\beta$, accommodate a broad class of mean structures, including main effects, interactions, and linear coefficients.
- **Random effects** part, $Zu$, provide a convenient mechanism for modeling correlation among the $N$ observations.

# Statistical Challenges

1. Data are **irregular functions**
2. Jointly model ***functional effects*** of multiple covariates
3. Account for **correlation between profiles** from same child
4. Incorporate information from **incomplete profiles**

# Functional Data Analysis

- **Functional Data:**
  - **Ideal units of observation: curves**
  - **Observed data: curves sampled on fine grid**
- **Increasingly encountered in scientific research.**
- **FDA Approach (Ramsay and Silverman 1997):**
  - **Treat functions as single entities, rather than a collection of observations**

# Functional Mixed Model

$Y(t)$ = set of $N$ observed curves, stacked as rows.

$$U_i(t) \sim GP(0, Q)$$

$$E_i(t) \sim GP(0, S)$$

$$\underbrace{Y(t)}_{\substack{\text{N}\\ \text{functions}}} = \overbrace{X}^{N \times p} \underbrace{B(t)}_{\substack{\text{p}\\ \text{functions}}} + \overbrace{Z}^{N \times m} \underbrace{U(t)}_{\substack{\text{m}\\ \text{functions}}} + \underbrace{E(t)}_{\substack{\text{N}\\ \text{functions}}}$$

- ***Q*** and ***S*** are covariance surfaces describing the how the random effect curves/residual error processes vary across replicates.
- Variations of this model were used by Guo (2002) and Morris and Carroll (2004)

# Functional Mixed Model
## (Discrete version)

**Y= *N*-by-*T* matrix containing the observed curves all sampled on a common equally-spaced grid of length *T*, *t*.**

$$\underbrace{Y}_{N\times \mathrm{T}} = \overbrace{X}^{N\times p}\underbrace{B}_{p\times \mathrm{T}} + \overbrace{Z}^{N\times m}\underbrace{U}_{m\times \mathrm{T}} + \underbrace{E}_{N\times \mathrm{T}}$$

$$U_i \sim MVN(0,Q)$$
$$E_i \sim MVN(0,S)$$

- **$B_{ij}$ is the effect of covariate *i* at location $t_j$**
- **$Q$ and $S$ are covariance matrices (*T* x *T*)**
- **Note: Some structure must be assumed on form of Q and S (discussed later)**

# Example: Model

Let $\boldsymbol{Y}$ be 292×660 matrix containing *292* accelerometer profiles for each minute from 9am-8pm.

$$\boxed{Y = XB + ZU + E}$$

- X = 292 × 14 matrix of covariates
  - School effects (5), gender, triceps skinfold, BMI, day-of-week (4), daylight savings time, avg TV hrs/day
- B = 14 × 660 matrix of fixed effects functions
  - $B_{ij}$ is effect of covariate $i$ at time $t_j$
- Z = 292 × 106 matrix indicating child for each profile

  U = 106 × 660 matrix of random effect functions (1/child)
- E = 292 × 660 matrix of residual errors

# Functional Mixed Models

- **Key feature of FMM: Does not require specification of parametric form for curves**
- **Guo (2002) fit FMM using splines**
- **Morris, et al. (2003) and Morris and Carroll (2004) use approach involving wavelet bases**
  - **Accommodates "spiky" functional data**
  - **Allows more general types of covariances**
- **Wavelet Regression: nonparametric regression technique that better preserves local features present in the curves.**

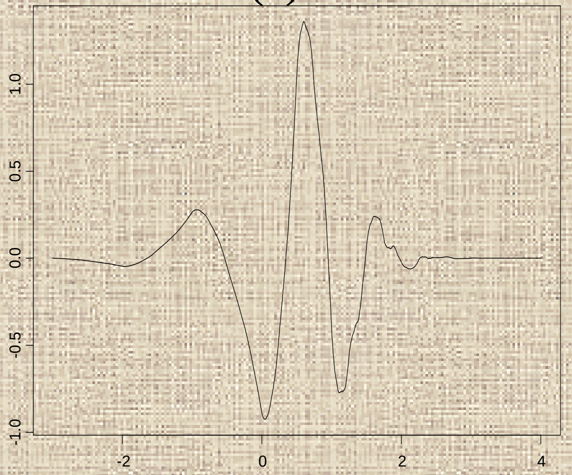
# Introduction to Wavelets

- **Wavelets: families of orthonormal basis functions**

$$g(t) = \sum_{j,k \in \Im} d_{jk} \psi_{jk}(t)$$

$$\psi_{jk}(t) = 2^{-j/2} \psi(2^{-j/2} t - k)$$

$$d_{jk} = \int g(t) \psi_{jk}(t) dt$$

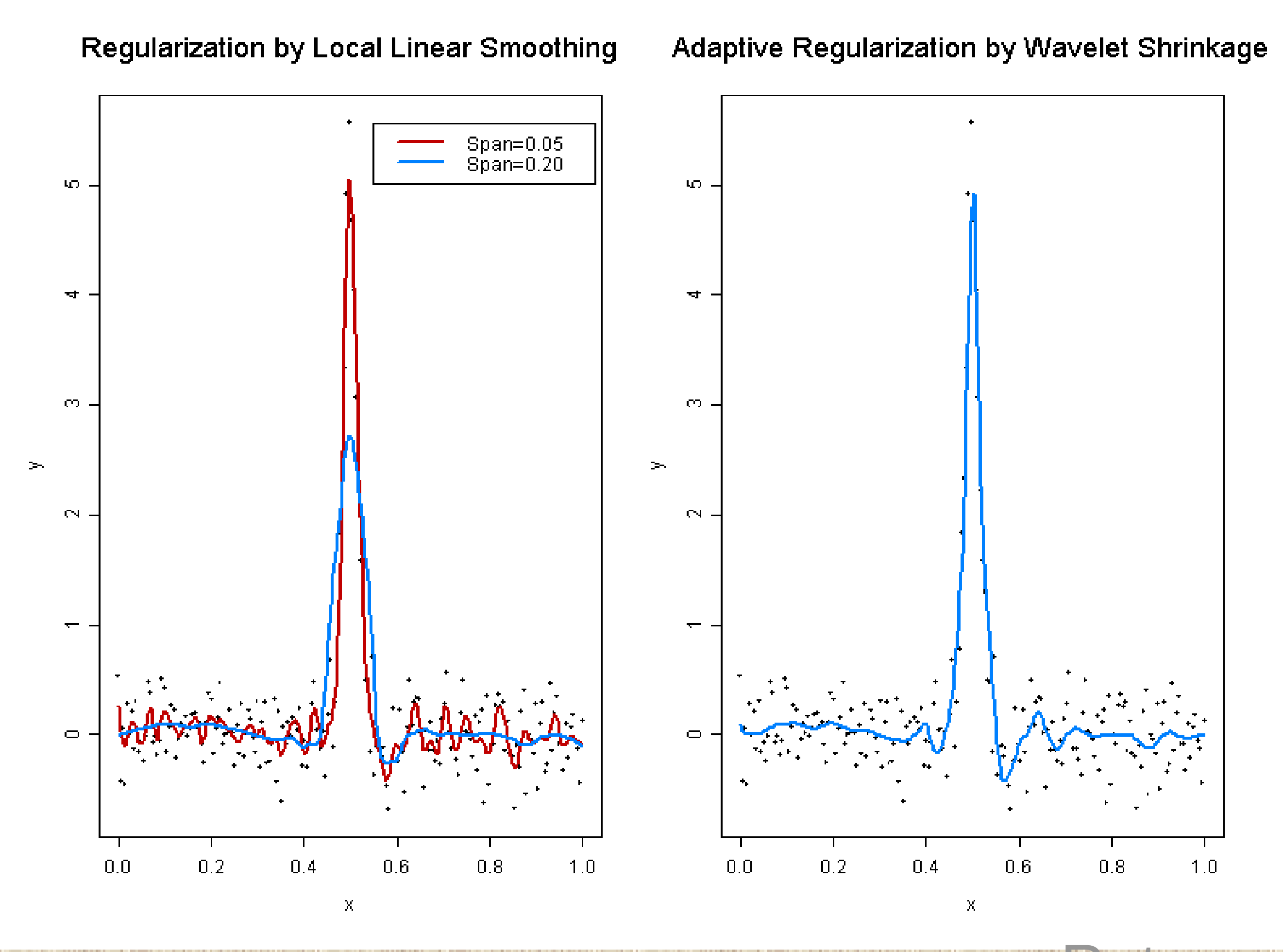



- **Discrete Wavelet Transform (DWT): fast algorithm {O(*T*)} for obtaining *T* empirical wavelet coefficients for curves sampled on equally-spaced grid of length *T*.**
- **Linear Representation:** $\mathbf{d = y\, W'}$
  - ***W'* =*T*-by-*T* orthogonal projection matrix**
- **Inverse DWT (IDWT):** $\mathbf{y = d\, W}$

# Wavelet Regression

- **Wavelet Regression – 3 step process**
  1. **Project data into wavelet space**
  2. **Threshold/shrink coefficients**
  3. **Project back to data space**
- **Yields *adaptively regularized* (plot) nonparametric estimates of function**
- **Morris, et al. (2003) extended to nested functional model (Bayesian)**
- **Morris and Carroll (2004) extended to general functional mixed model framework (wavelet-based FMM)**

# Adaptive Regularization




Return

# Wavelet-Based FMM:

## General Approach

1. **Project observed functions Y into wavelet space.**
2. **Fit FMM in wavelet space.**

   **(Use MCMC to get posterior samples)**
3. **Project wavelet-space estimates (posterior samples) back to data space.**

# Wavelet-Based FMM:

## General Approach

1. **Project observed functions Y into wavelet space.**

2. **Fit FMM in wavelet space**

   **(Use MCMC to get posterior samples)**

3. **Project wavelet-space estimates (posterior samples) back to data space.**

# Wavelet-Based FMM

1. **Project observed functions Y to wavelet space**

- **Apply DWT to rows of Y to get wavelet coefficients corresponding to each observed function**

$$\boxed{\underbrace{D}_{N\times T} = \underbrace{Y}_{N\times T}\ \underbrace{W'}_{T\times T}}$$

- **Projects the observed curves into the space spanned by the wavelet bases.**

# Wavelet-Based FMM:

## General Approach

1. **Project observed functions Y into wavelet space.**
2. **Fit FMM in wavelet space**

   **(Use MCMC to get posterior samples)**
3. **Project wavelet-space estimates (posterior samples) back to data space.**

# Projecting FMM to Wavelet Space

$$\underbrace{Y}_{N\times T} = \overbrace{X}^{N\times p} \underbrace{B}_{p\times T} + \overbrace{Z}^{N\times m} \underbrace{U}_{m\times T} + \underbrace{E}_{N\times T}$$

$$U_i \sim MVN(0, Q)$$

$$E_i \sim MVN(0, S)$$

# Projecting FMM to Wavelet Space

$$\underbrace{Y}_{N\times \mathrm{T}}\ \overbrace{\mathbf{W'}}^{T\times T} = \overbrace{X}^{N\times p}\ \underbrace{B}_{p\times \mathrm{T}} + \overbrace{Z}^{N\times m}\ \underbrace{U}_{m\times \mathrm{T}} + \underbrace{E}_{N\times \mathrm{T}}$$

$$U_i \sim MVN(0,Q)$$

$$E_i \sim MVN(0,S)$$

# Projecting FMM to Wavelet Space

$$\underbrace{Y}_{N\times \mathrm{T}}\overbrace{\mathbf{W}'}^{T\times T} = \overbrace{X}^{N\times p}\underbrace{B}_{p\times \mathrm{T}}\overbrace{\mathbf{W}'}^{T\times T} + \overbrace{Z}^{N\times m}\underbrace{U}_{m\times \mathrm{T}}\overbrace{\mathbf{W}'}^{T\times T} + \underbrace{E}_{N\times \mathrm{T}}\overbrace{\mathbf{W}'}^{T\times T}$$

$$U_i \sim MVN(0, Q)$$

$$E_i \sim MVN(0, S)$$

# Projecting FMM to Wavelet Space

$$\underbrace{Y}_{N\times \mathrm{T}} \overbrace{\mathbf{W'}}^{\mathbf{T}\times\mathbf{T}} = \overbrace{X}^{N\times p} \underbrace{B}_{p\times \mathrm{T}} \overbrace{\mathbf{W'}}^{T\times T} + \overbrace{Z}^{N\times m} \underbrace{U}_{m\times \mathrm{T}} \overbrace{\mathbf{W'}}^{T\times T} + \underbrace{E}_{N\times \mathrm{T}} \overbrace{\mathbf{W'}}^{T\times T}$$

$$U_i \mathbf{W'} \sim MVN(0, \mathbf{W}Q\mathbf{W'})$$

$$E_i \mathbf{W'} \sim MVN(0, \mathbf{W}S\mathbf{W'})$$

# Projecting FMM to Wavelet Space

$$\underbrace{\mathbf{D}}_{N\times T} = \overbrace{X}^{N\times p}\underbrace{\mathbf{B}^{*}}_{p\times T} + \overbrace{Z}^{N\times m}\underbrace{\mathbf{U}^{*}}_{m\times T} + \underbrace{\mathbf{E}^{*}}_{N\times T}$$

$$\mathbf{U}_i^{*} \sim MVN(0,\mathbf{Q}^{*})$$

$$\mathbf{E}_i^{*} \sim MVN(0,\mathbf{S}^{*})$$

# Adaptive Regularization via Shrinkage Prior

## Mixture prior on $B_{ijk}^*$:

$$B_{ijk}^* = \gamma_{ijk}^* N(0, \tau_{ij}) + (1 - \gamma_{ijk}^*)\delta_0$$

$$\gamma_{ijk}^* = \text{Bernoulli}(\pi_{ij})$$

- **Nonlinearly shrinks $B_{ijk}^*$ towards 0, leading to adaptively regularized estimates of $B_i$.**
- **$\tau_{ij}$ & $\pi_{ij}$ are regularization parameters**
  - **Can be estimated from the data using empirical Bayes**
  - **Extend Clyde&George (1999) to functional mixed model**

# Model Fitting

- Let $\Omega$ be a vectorcontaining ALL covariance parameters (in $Q^*$ and $S^*$).

## MCMC Steps

1. **Sample from $f(B^*|D,\Omega)$:**

   **Mixture of normals and point masses at 0 for each $i,j,k$.**

2. **Sample from $f(\Omega|D,B^*)$:**

   **Metropolis-Hastings steps for each $j,k$**

3. **If desired, sample from $f(U^*|D,B^*,\Omega)$:**

   **Multivariate normals**

# Wavelet-Based FMM:

## General Approach

1. **Project** **observed functions Y** **into wavelet space.**
2. **Fit FMM** **in wavelet space**
   **(Use MCMC to get posterior samples)**
3. **Project** **wavelet-space estimates (posterior samples)** **back to data space**.

# Wavelet-Based FMM

3. **Project wavelet-space estimates (posterior samples) back to data space.**

- **Apply IDWT to posterior samples of $B^*$ to get posterior samples of fixed effect functions $B_i$ for $i=1,\ldots, p$, on grid $t$.**
  - **B=B*W**
- **Posterior samples of $U$, $Q$, and $S$ are also available, if desired.**
- **Can be used for any desired Bayesian inference**

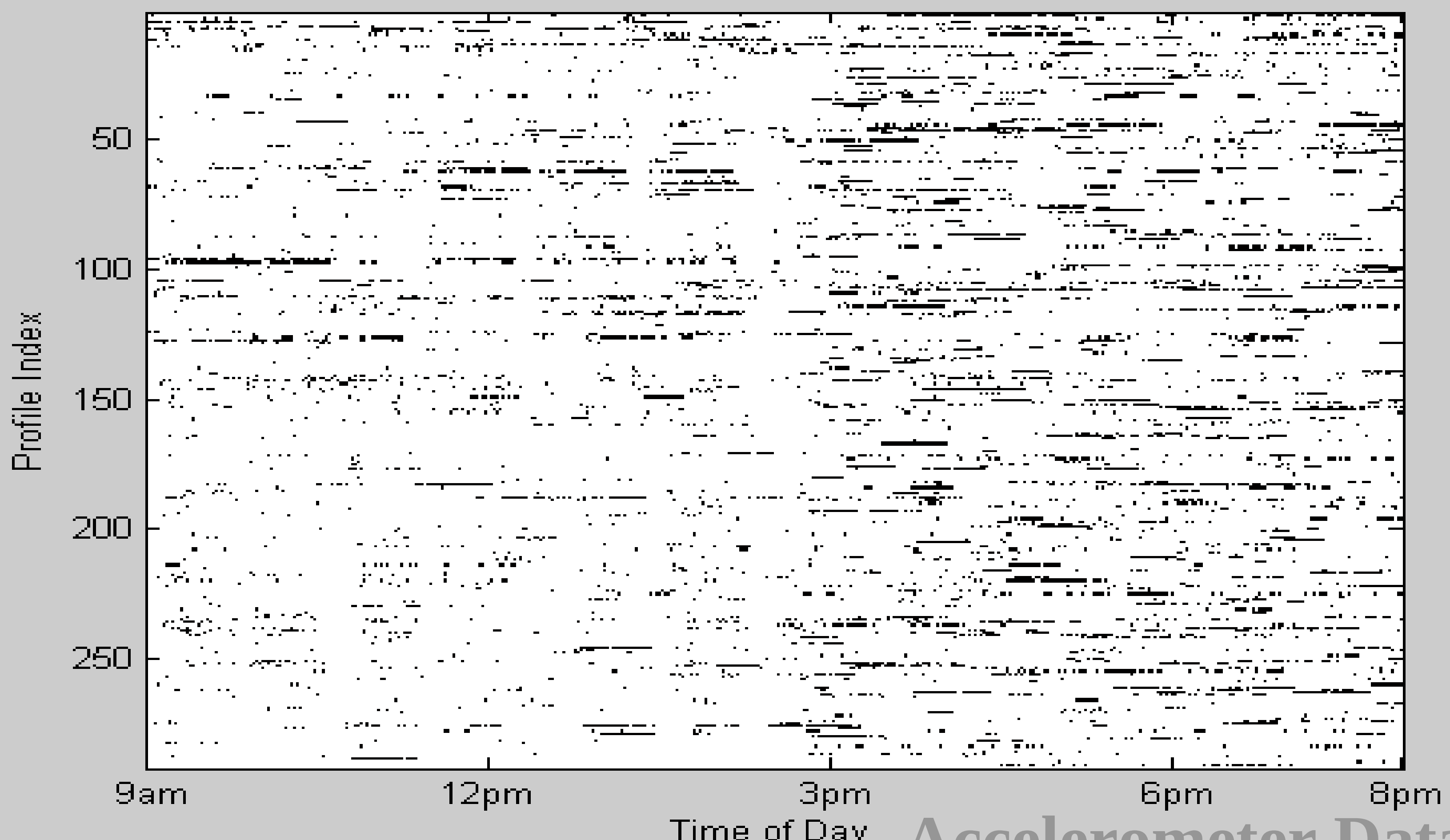
# Incomplete Profiles

- **Lots of missing data** **[(Missing Data)]**

  **Example of *incomplete profile***

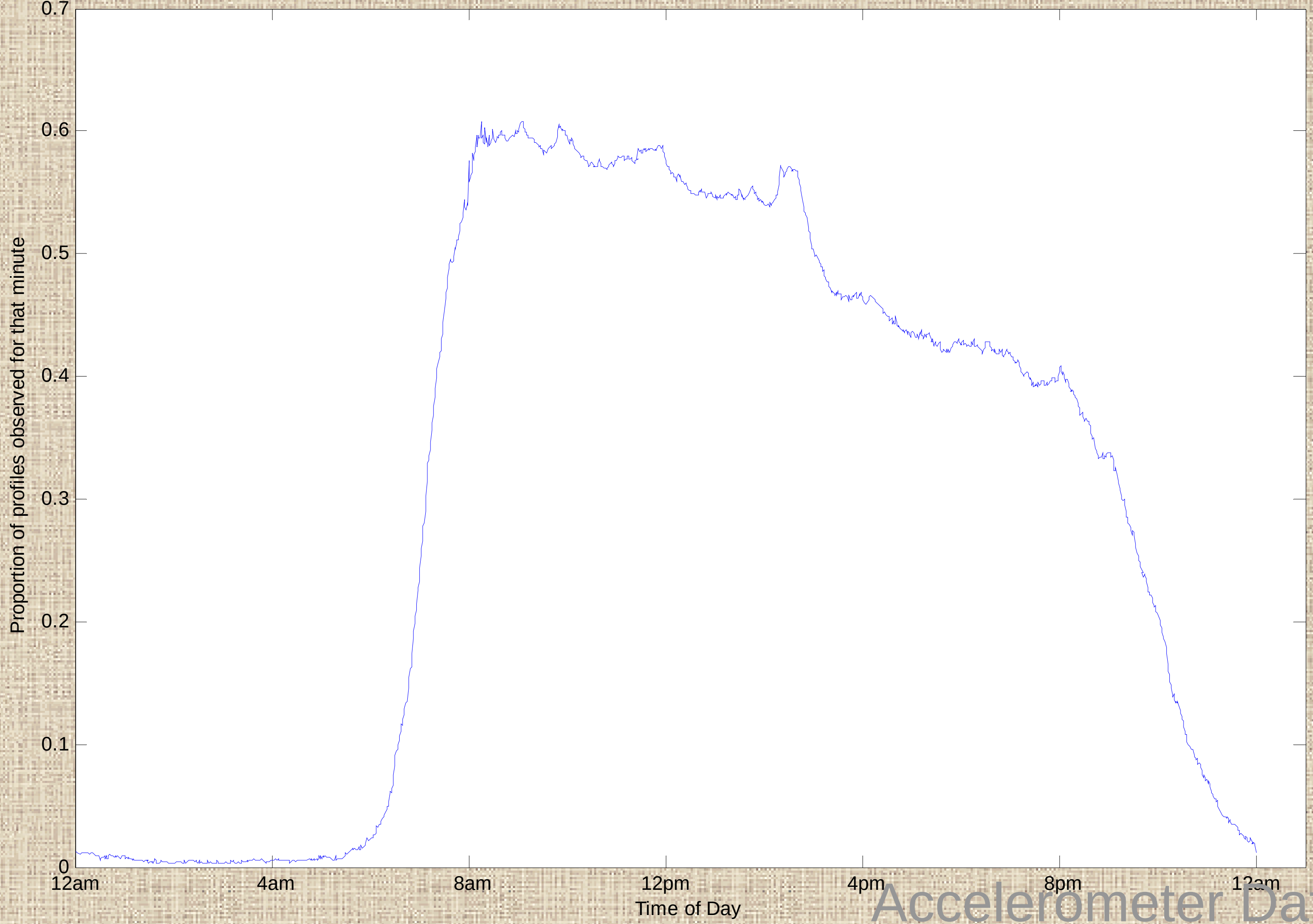
- **WFMM can only be applied to *complete profiles* (with no missing regions)**
  - **95 of the 292 profiles *complete* from 9am-8pm**
- **How do we incorporate information from other 197 *incomplete profiles* ?**

# Heatmap of Missingness
## (Black=missing)




Accelerometer Data

# Missingness vs. Time-of-Day

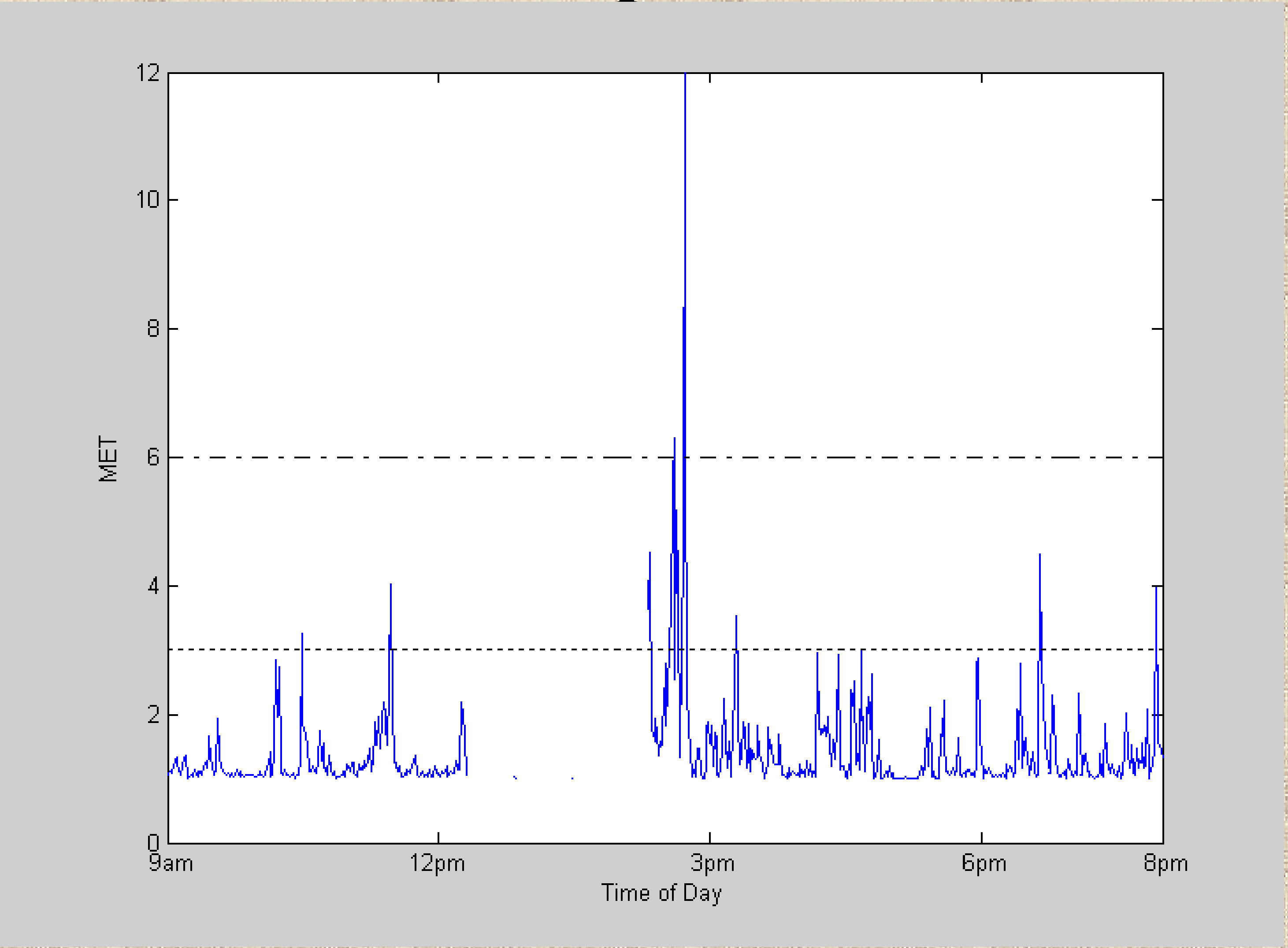



Accelerometer Data

# Approach: Incomplete Profiles

1. First fit model to ***complete profiles***, get posterior distribution samples for model parameters.
2. Use these to estimate ***predictive distributions*** for the the incomplete profiles (fig)
   - Borrow information about what the curves in these regions look like.
   - Account for child-specific and day-specific covariates.
3. Regress missing data on the observed data to obtain ***imputation distribution*** for missing regions (fig)
   - Borrow information from nearby times in incomplete profiles.
   - Makes predictions for missing regions "connected" with observed.
4. Supplement WFMM with step to **stochastically impute** values for missing data.
   - Inference appropriately accounts for uncertainty in imputation

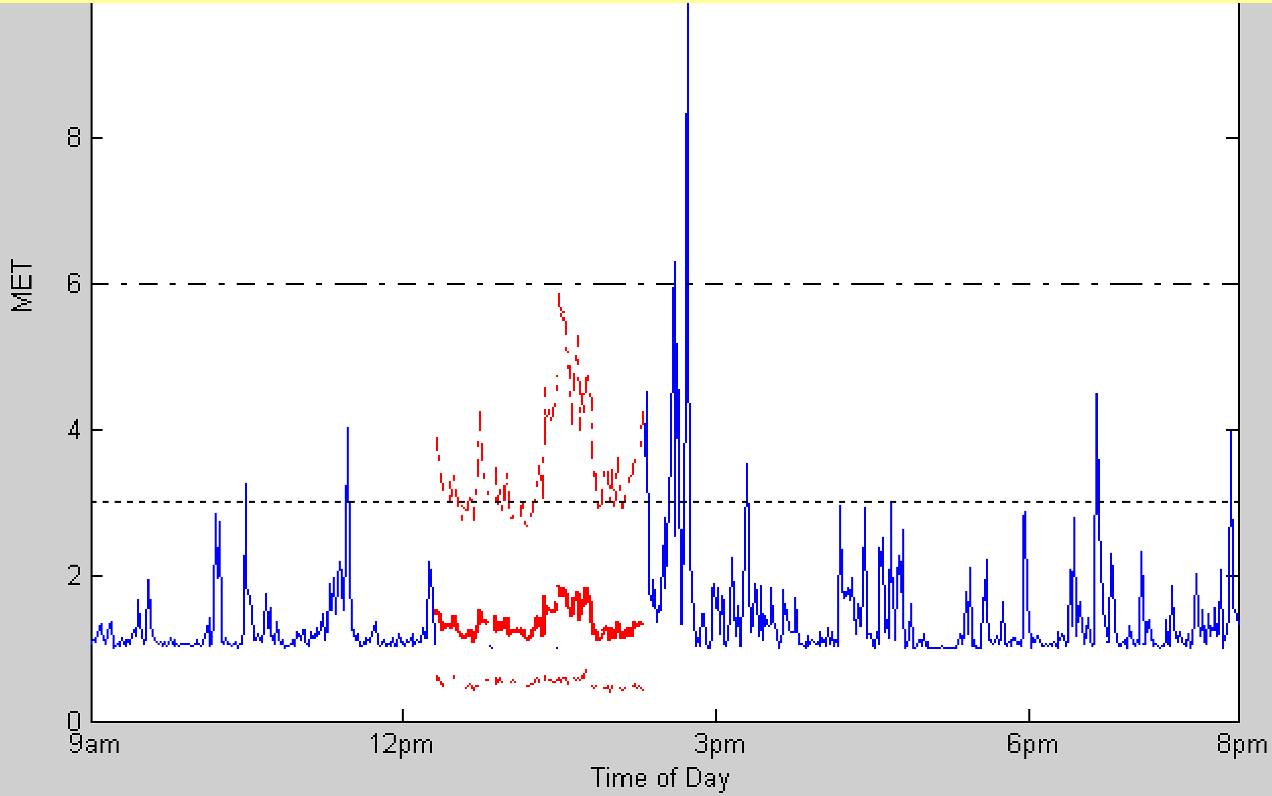
# Incomplete Profile



# Predictive Distribution

$$\mu_i(t) = E\{Y_i(t) \mid Y^C\} = \int Y_i(t) f\{Y_i(t) \mid X, Z, \Theta\} f(\Theta \mid Y^C) d\Theta$$

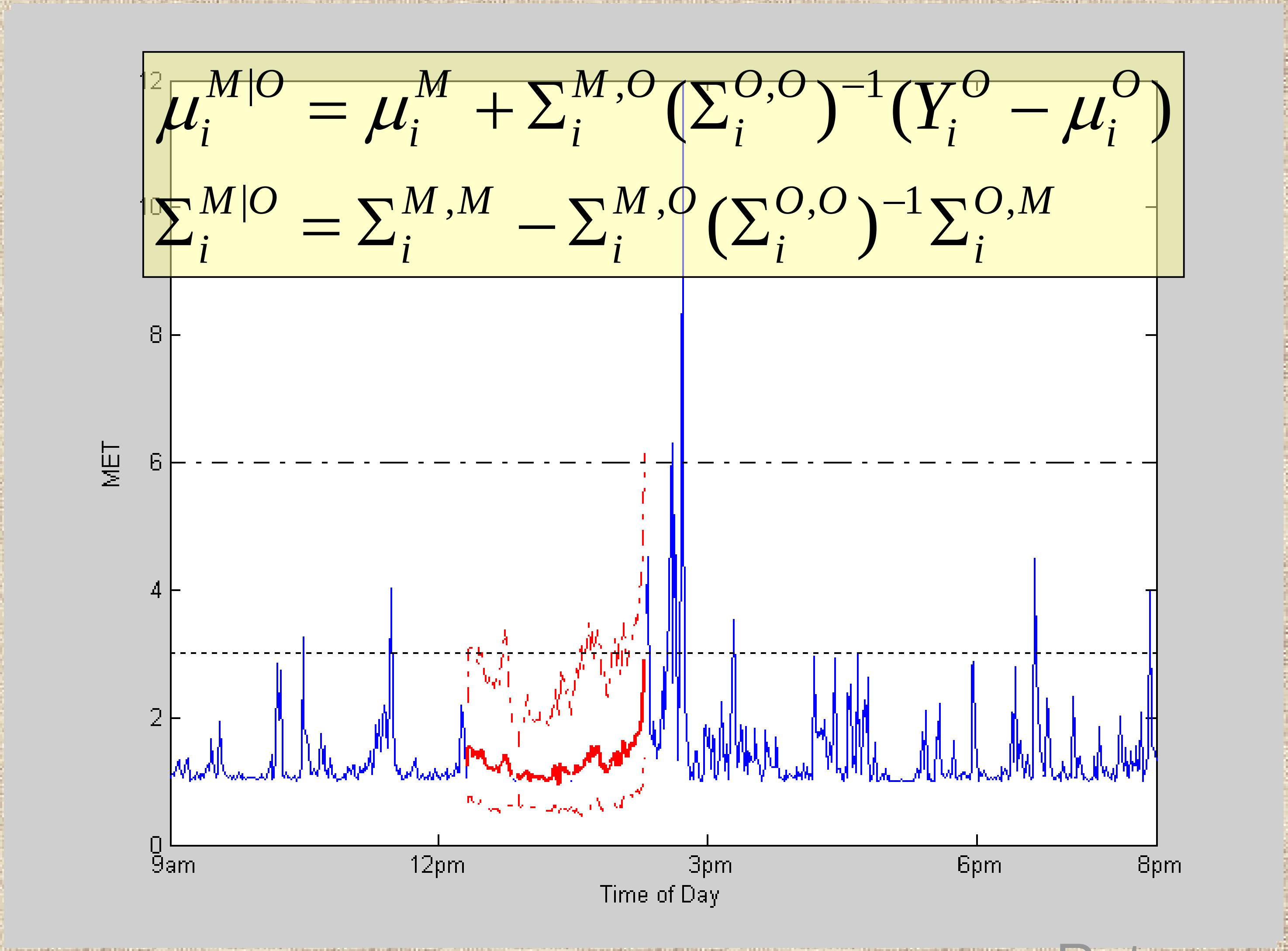
$$\Sigma_i(t_1, t_2) = COV\{Y_i(t_1), Y_i(t_2) \mid Y^C\}$$



# Imputation distribution

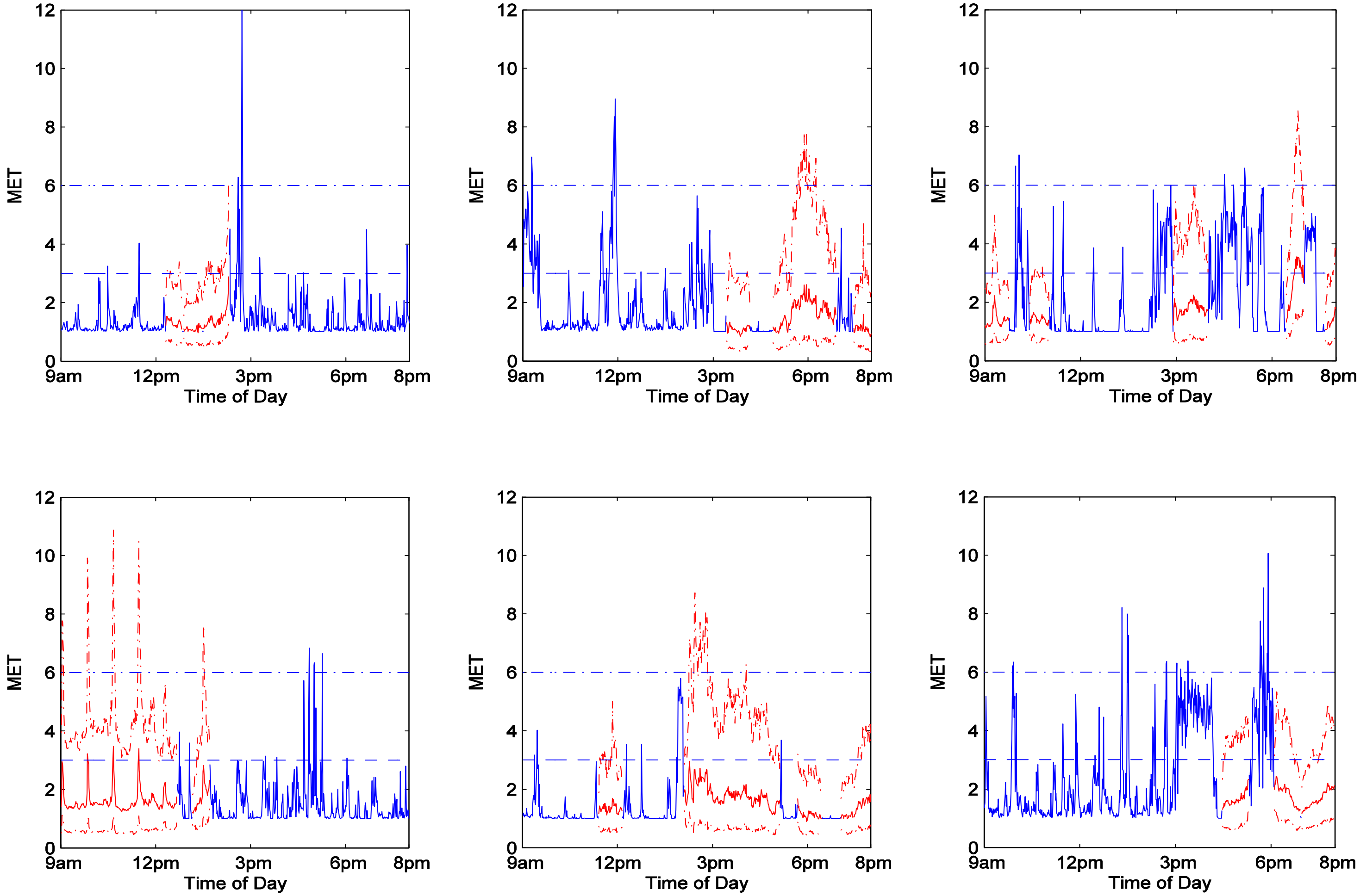
$$\mu_i^{M|O} = \mu_i^M + \Sigma_i^{M,O}(\Sigma_i^{O,O})^{-1}(Y_i^O - \mu_i^O)$$

$$\Sigma_i^{M|O} = \Sigma_i^{M,M} - \Sigma_i^{M,O}(\Sigma_i^{O,O})^{-1}\Sigma_i^{O,M}$$




Return

# Incomplete Profiles



# Missing Data in the WFMM

- **Problem:** Imputation distribution in data space, modeling done in wavelet space
- **Solution:** Project imputation distributions into wavelet space

$$M_i(t) = \begin{cases} Y_i(t) \text{ if } t \text{ observed} \\ \mu_i^{M|O}(t) \text{ otherwise} \end{cases}$$

$$V_i(t_1, t_2) = \begin{cases} 0 \text{ if either } t_1 \text{ or } t_2 \text{ obs.} \\ \Sigma_i^{M|O}(t_1, t_2) \text{ otherwise} \end{cases}$$

$$M_i^* = M_i W'$$

$$V_i^* = W V_i W'$$

- Add step to MCMC whereby "missing" wavelet coefficients $D_{ijk} \sim N(M^*_{ijk}, V^*_{ijk})$

# Model

**Let *Y* be 292×660 matrix containing *292* accelerometer profiles for each minute from 9am-8pm.**

$$\boxed{Y = XB + ZU + E}$$

- **X = 292 × 14 matrix of covariates**
  - **School effects (5), gender, triceps skinfold, BMI, day-of-week (4), daylight savings time, avg TV hrs/day**
- **B = 14 × 660 matrix of fixed effects functions**
  - **$B_{ij}$ is effect of covariate *i* at time $t_j$**
- **Z = 292 × 106 matrix indicating child for each profile**
  **U = 106 × 660 matrix of random effect functions (1/child)**
- **E = 292 × 660 matrix of residual errors**

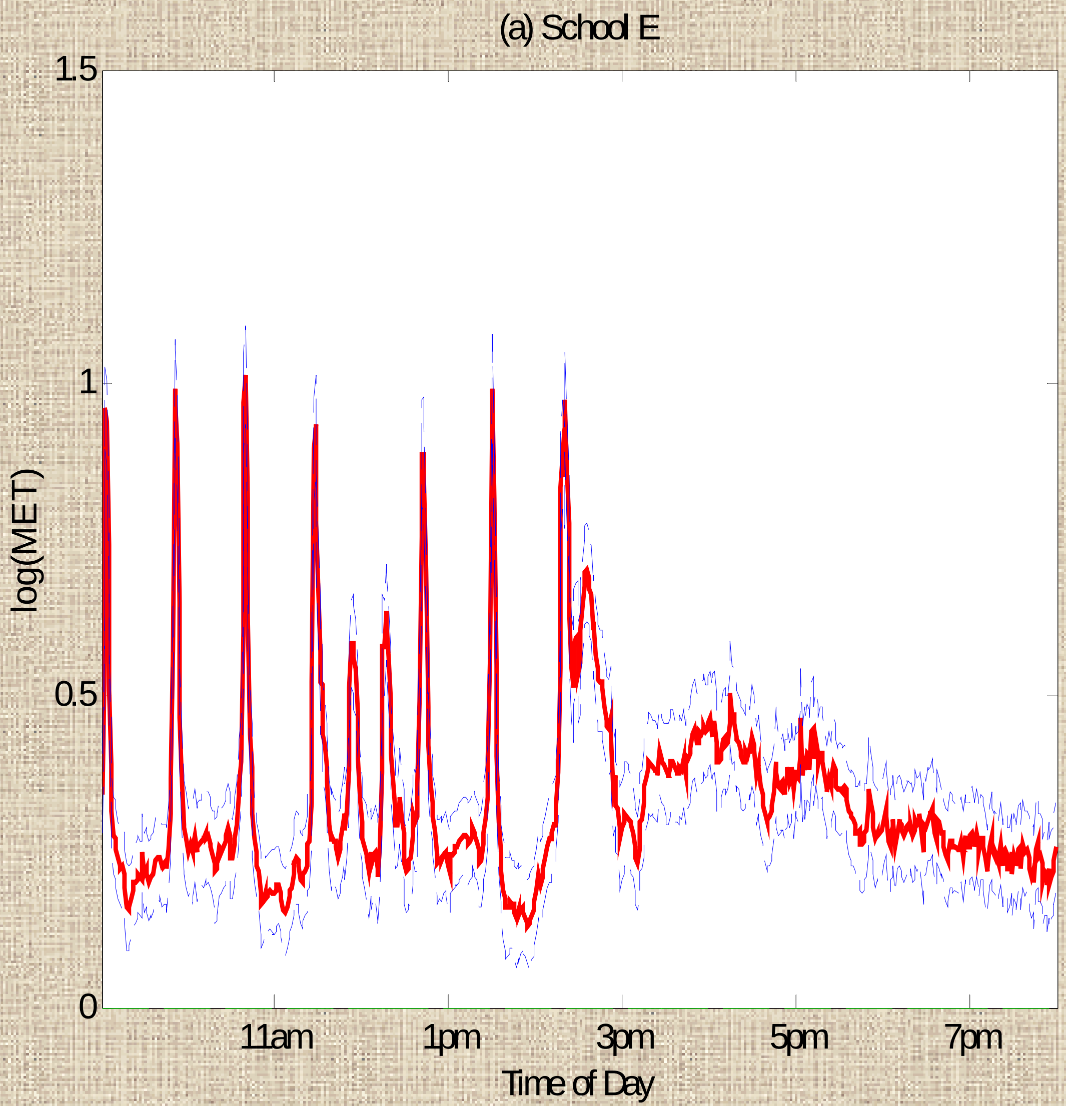
(short results)
(Bayesian Inference)

# Statistical Inference in FMM

- Prior distributions made "uninformative"
  - Exception: *smoothing parameters*, which can be estimated from the data using an "empirical Bayes" approach
- We use MCMC to draw samples from posterior distributions of $B_i(t)$ functions and covariances.
- For each covariate effect, we compute 90% pointwise credible intervals at each value of $t$
- We also computed 90% credible intervals, aggregating data within following time intervals:

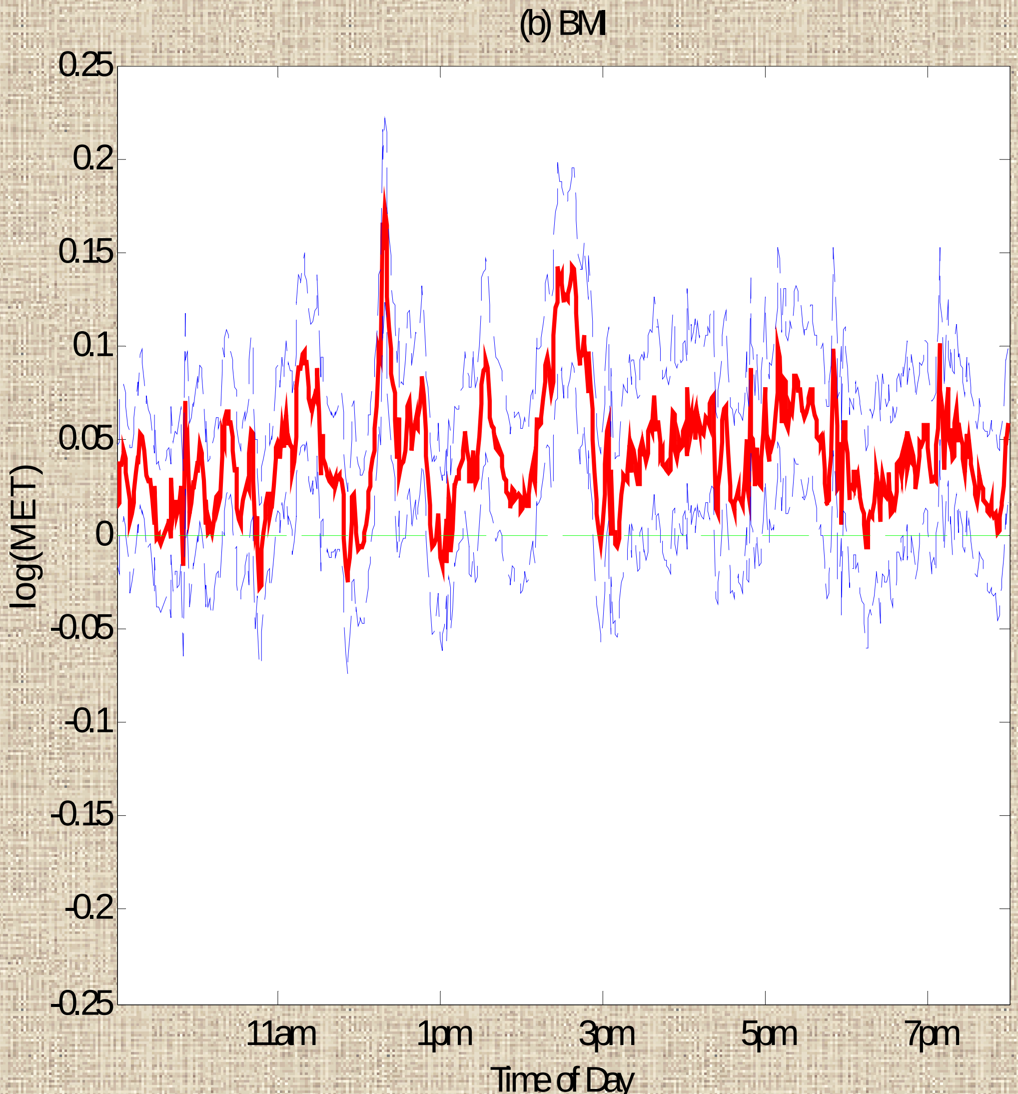
1. All day (9am-8pm)
2. Morning (9am-11:30am)
3. Lunch (11:30am-12:30pm)
4. Afternoon (12:30-2:15pm)
5. Going home (2:15pm-3pm)
6. After school (3pm-5:30pm)
7. Early evening (5:30pm-7pm)
8. Late evening (7pm-8pm)

# Selected Results: School Effects



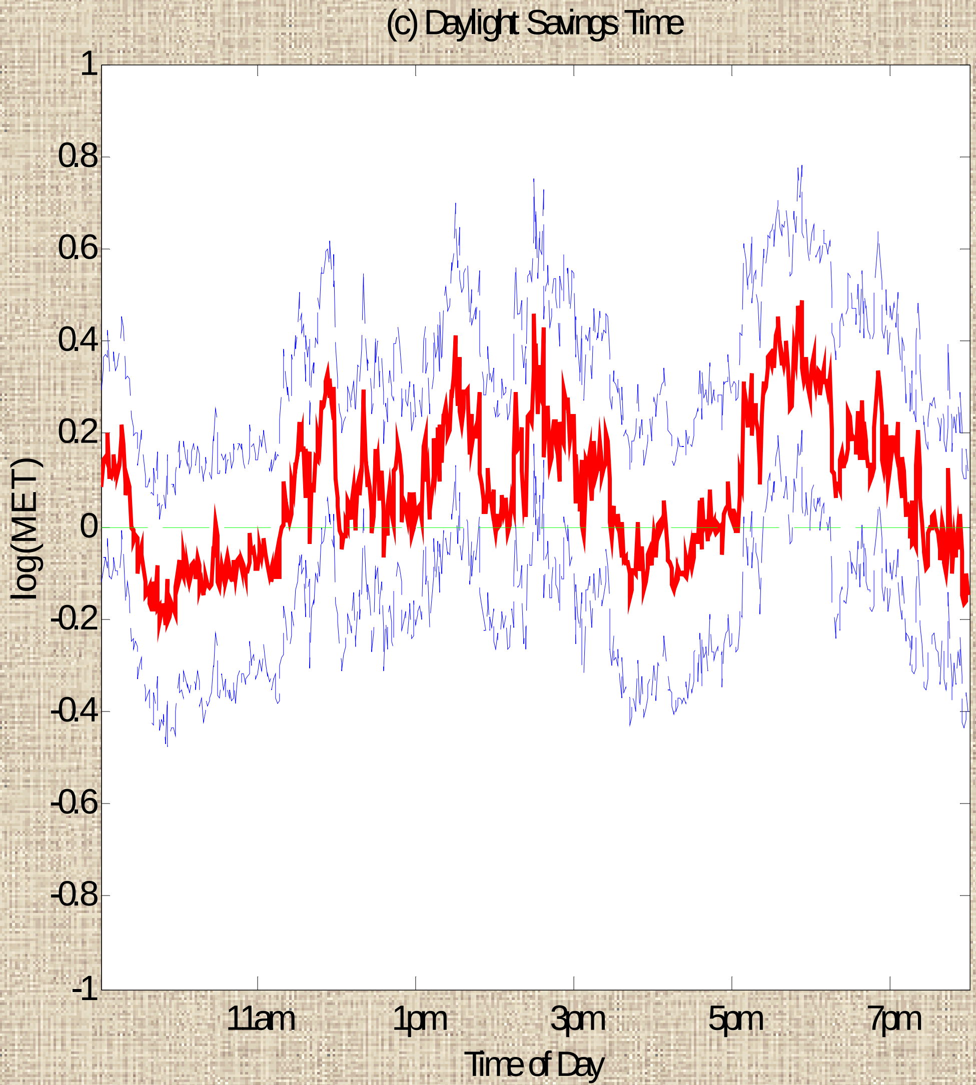

- **School schedules evident in effects**
  - Spikes every 48min (changing classes)
  - 3 lunch periods
  - School out at 2:15pm
- **Not so evident in individual curves**

# Selected Results: BMI Effect



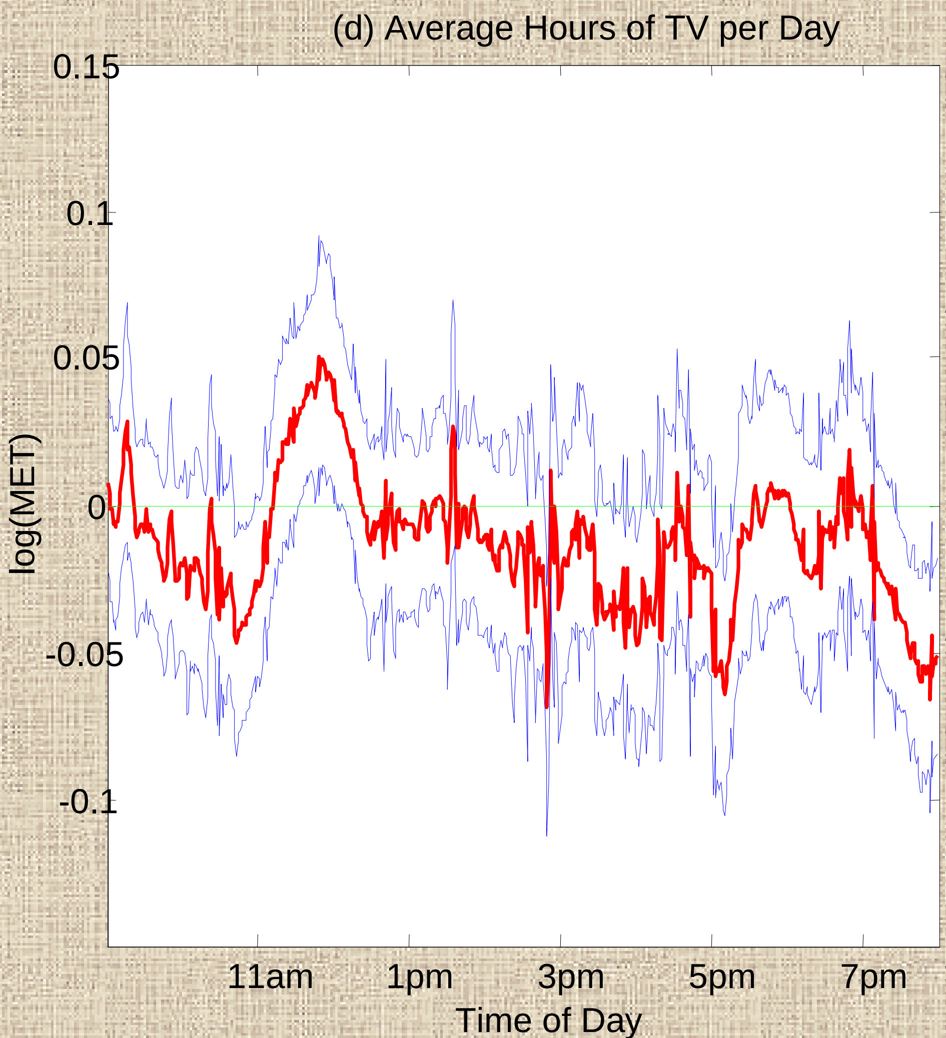

- **BMI coded as continuous factor (mean-centered)**
- **BMI effect positive (*post prob<0.0005*)**
  - **Higher BMI, more energy expenditure**
- **Not present in raw activity levels**
  - **Artifact of conversion from acceleration to energy scale?**

# Selected Results: DST Effect



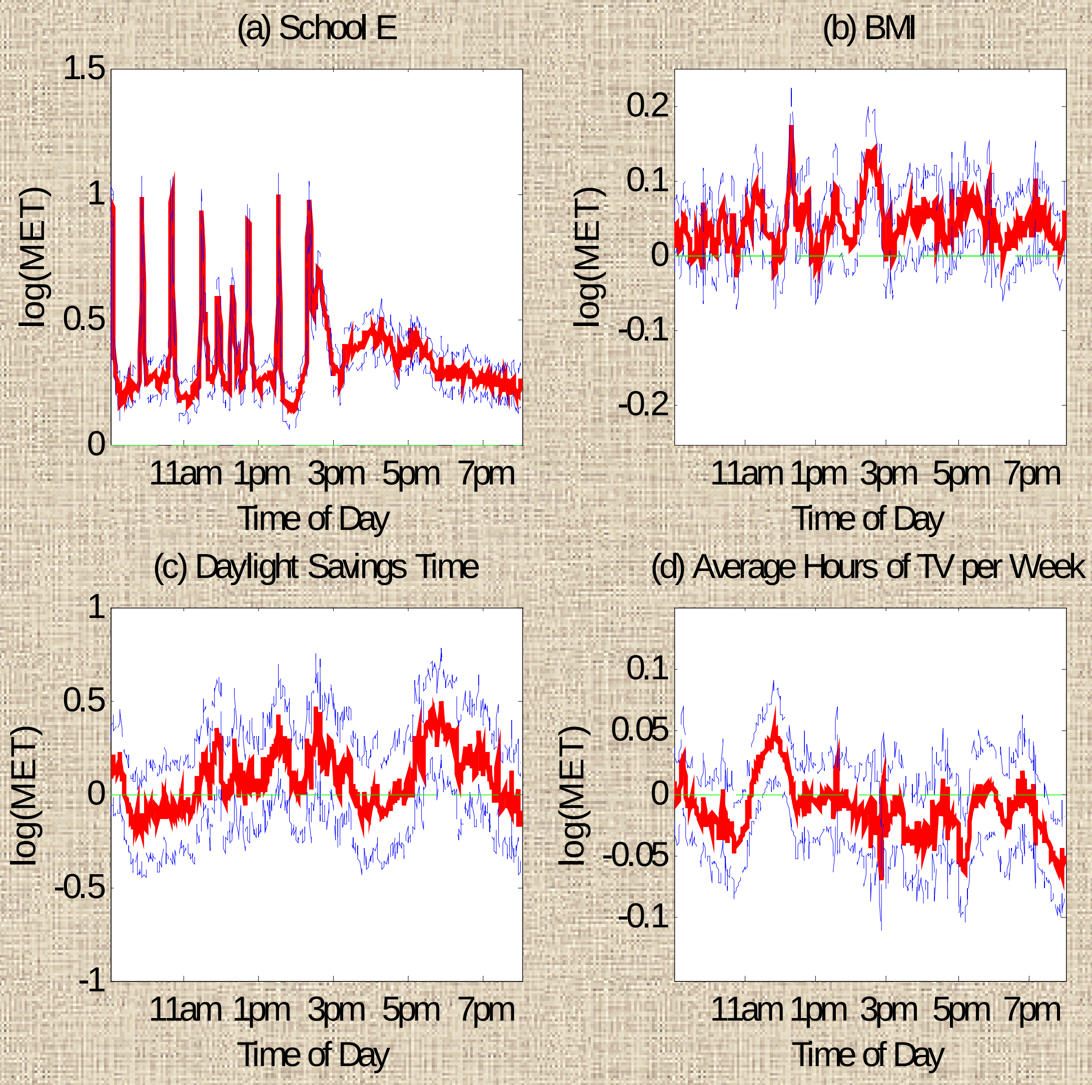

- **DST – April 6th**
- **More active after DST (overall 8%, *p=0.062*)**
- **Especially strong:**
  - **As school is letting out (2:15-3:00, 25%, *p=0.03*)**
  - **In early evening (5:30-7:00, 30%, *p=0.01*)**
- **Note: Sunset was**
  - **5:10-6:15 before DST**
  - **7:15-8:10 after DST**

# Selected Results: TV hours/day




- **TVhrs coded as continuous factor (standardized)**
- **TVhrs effect negative (-1.3% per sd, *p=0.03*)**
  - **More TV, less active**
  - **3:00-5:30, -2.6%, *p=0.02***
  - **7:00-8:00, -3.6%, *p=0.008***
- **Positive effect over lunch**
  - **+2.7%, *p=0.03***
  - **More TV, on average more active over lunch**

# Some Results

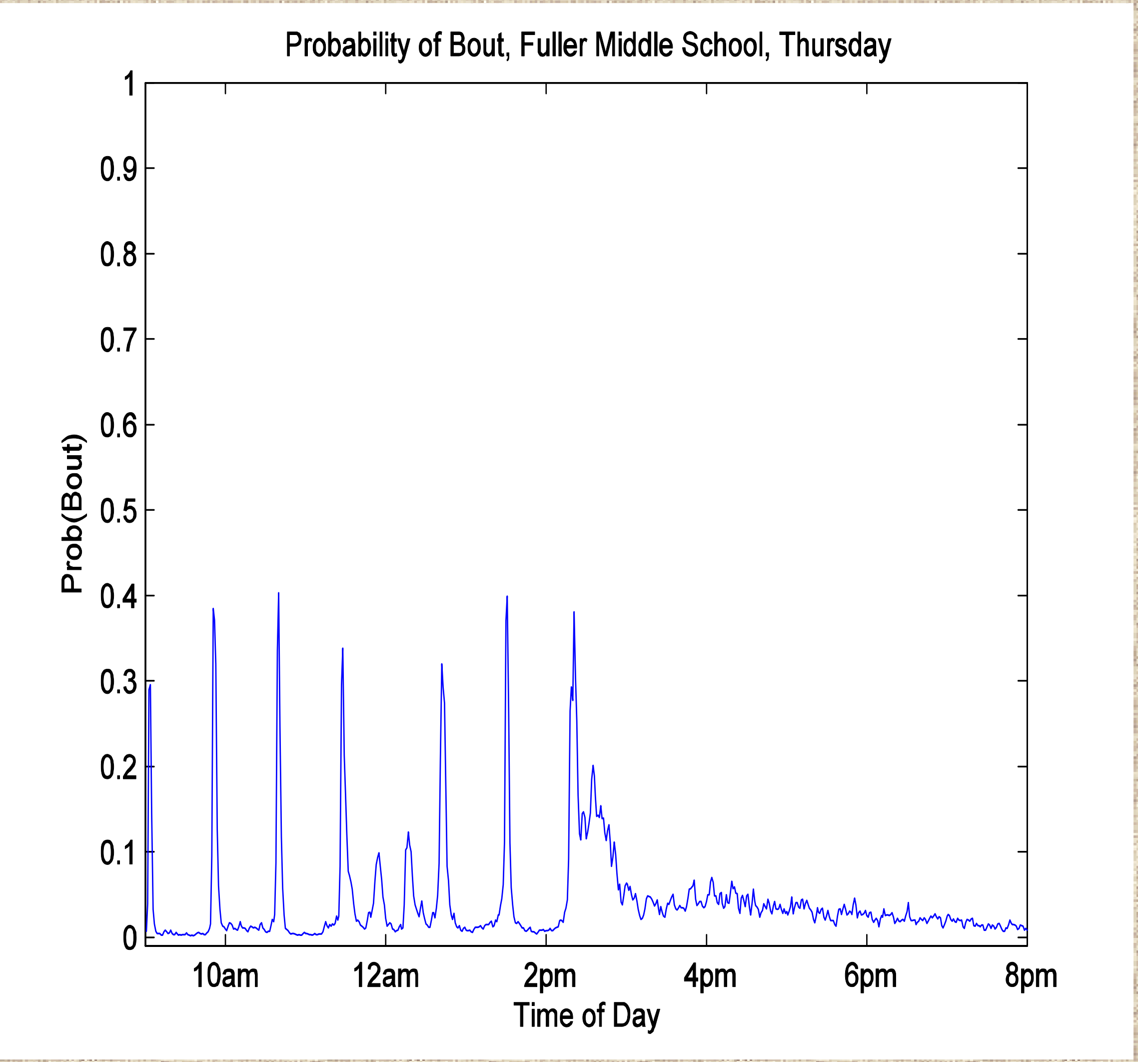



- **School #5:**
  - **Spikes every 48min (changing classes)**
  - **3 lunch periods**
  - **School out at 2:15pm**
- **BMI effect positive**
  - **More energy expenditure**
- **Daylight Savings Time**
  - **More active after DST**
  - **Especially 2-3pm, 5-7pm**
- **Avg hrs TV/day**
  - **More TV=less active**
  - **Especially 3-5pm, 7-8pm**
  - **More active at lunch**

# Results: Covariance Analysis

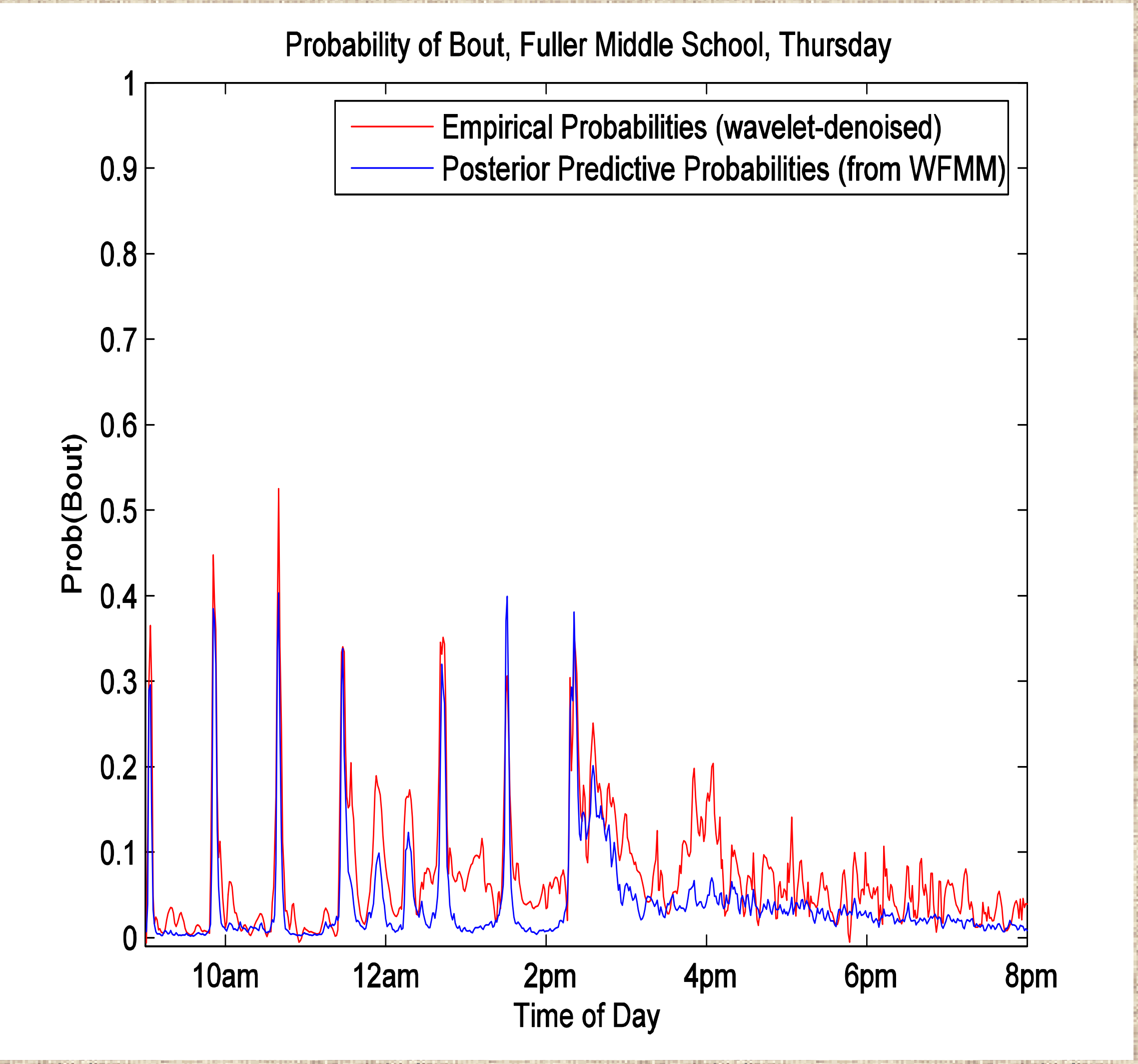
- **Variability: 91% day-to-day, 9% child-to-child**
  - **Important to have many days per child**
- **Study variability as function of *t***
  - **Child-to-child variability: school day > after school**
  - **Day-to-day variability: after school > school day**
  - **Relative day-to-day variability after school: 95-99%**
- **Less children, more days, save $$$?**
  - **But adherence a major limitation**
- **Would be nice to get more schools, too.**

# Results: Bouts




- **Can compute posterior predictive probabilities of bouts for children**

# Results: Bouts




- Can compute posterior predictive probabilities of bouts for children
- Model-based predictive probabilites not far from empirically-estimated probabilities
- May want heavier tails

# Conclusions

- Found some interesting results in Planet Health
  - School schedules, DST, Day of Week, TV hours
  - Day-to-day variability large – need lots of days/child
  - Raw vector magnitudes instead of estimated METs?
- Functional mixed models promising method for analyzing accelerometer data
  - Makes use of all of the data
  - Can detect effects that vary over time
  - Can accommodate incomplete profiles
  - After the fact, one can perform inference that averages within different time periods of interest
  - Can still look at probabilities bouts, but as function of time

# Acknowledgements


- **Statistical Collaborators on work related to WFMM: Raymond J. Carroll, Louise Ryan, Phil Brown, Marina Vannucci, Kevin Coombes, Keith Baggerly, Naisyin Wang, Brent Coull, and Cassandra Arroyo.**
- **Other Scientific Collaborators: Steve Gortmaker, Joanne Lupton, Josh Fidler, Lihong Wang, Victor Prieto, Stan Hamilton, Ryuji Kobayashi, David Hawke, John Kooman, Nancy Turner, and Meeyoung Hong.**
- **Thanks to Dick Herrick for assistance in optimizing the code for the method, and for converting the Matlab code to C++.**