Wavelet-Based Modeling of Hierarchical Functions in Colon Carcinogenesis

Jeffrey S. Morris

UT MD Anderson Cancer Center

Marina Vannucci and Raymond J. Carroll

Texas A&M University

Philip J. Brown

University of Kent, Canterbury

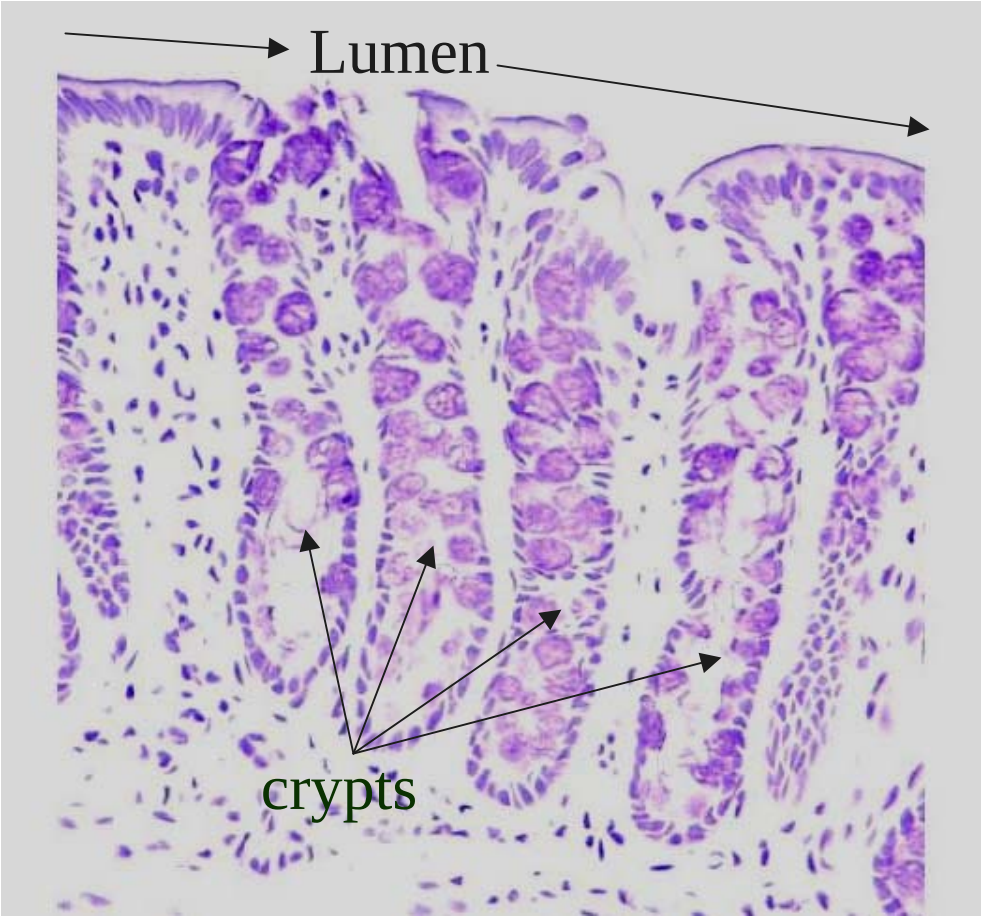
Goals

- **Biological Goal**: Gain insights into biological mechanisms of early colon carcinogenesis
- **Statistical Goal**: Develop methodology to nonparametrically model and perform inference on irregular functional data arising from a nested design.

Outline

- Introduction
  - Scientific Background/Experiment
  - Hierarchical Functional Model
- Methodology
  - Wavelet Space Modeling
  - Regularization by Wavelet Shrinkage
  - Model Fitting and Inference
- Application
- Conclusions

Architecture of Colon

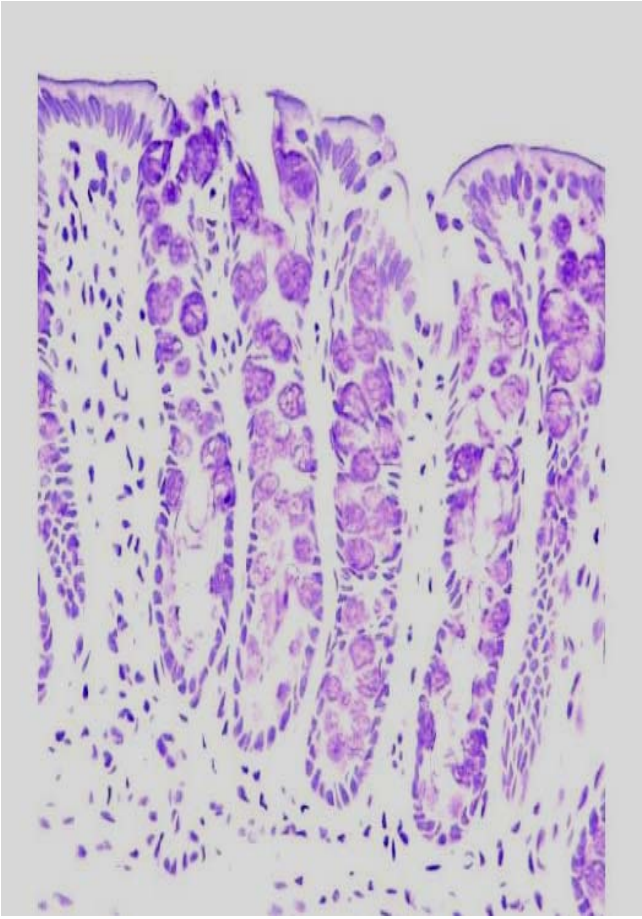


- ***Stem Cells***: Mother cells near crypt base
- Depth in crypt ~ age of cells
- ***Relative Cell Position***

  (Depth w/in crypt):

  $\boldsymbol{t} \in (0,1)$

Colon Carcinogenesis Studies



- Rats given different treatments, exposed to carcinogen, then euthanized.
- DNA adducts, DNA repair, apoptosis quantified as staining intensities along crypt wall
- Data has hierarchical structure
  - Treatment groups - rats - crypts - cells
  - Hierarchical Functional Data

Details of Experiment

- Fish oil diet protective compared with corn oil diet
- Study:
  - 2 diets (fish/corn) x 5 times (0h,3h,6h,9h,12h)
  - 3 rats per diet x time combination
  - ~ 25 crypts per rat,
  - Responses quantified on equally-spaced, fine grid
  - Interpolation used to get common grid
- Response: DNA Repair Enzyme (MGMT)
- Questions of interest:
  - At what depths within crypts is MGMT expressed?
  - Diet effect on MGMT? Depend on time/depth?
  - Does MGMT vary more from crypt-crypt or rat-rat?

Hierarchical Functional Model

- **Hierarchical Functional Model:**

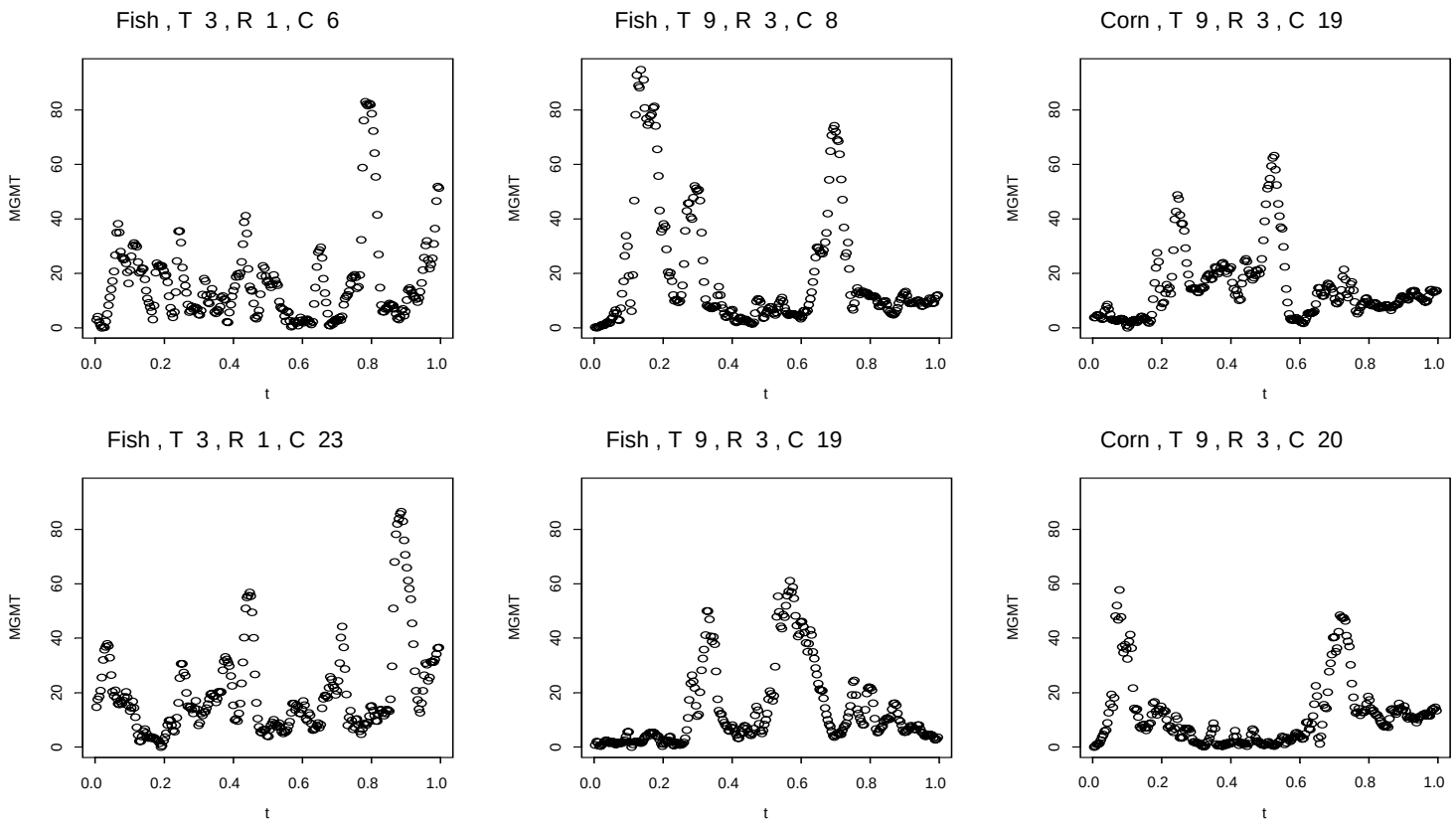
$$\boxed{\begin{aligned} Y_{abc}(t) &= g_{abc}(t) + e_{abc}(t), \\ g_{abc}(t) &= g_{ab}(t) + \eta_{abc}(t) \\ g_{ab}(t) &= g_a(t) + \xi_{ab}(t) \end{aligned}}$$

where $\mathbf{e}_{abc} \sim MVN(\mathbf{0}, \sigma_e^2 \mathbf{I})$,

$\eta_{abc}(\bullet)$ and $\xi_{\mathrm{ab}}(\bullet)$: mean 0 random processes with covariance surfaces $\Sigma_1(t_1, t_2)$ and $\Sigma_2(t_1, t_2)$.

DNA Repair Data

DNA Repair Enzyme for Selected Crypts



Functional Mixed Models

- Papers in literature for nonparametrically modeling nested functional data:
- Brumback/Rice 1998, Wang 1998, Guo 2002
- Each paper only applies methodology to simple 1-level hierarchy?
- All use spline bases.
- Splines may not be best choice for spiky data, since smoothness assumptions may lead to attenuation of important local features.

Introduction to Wavelets

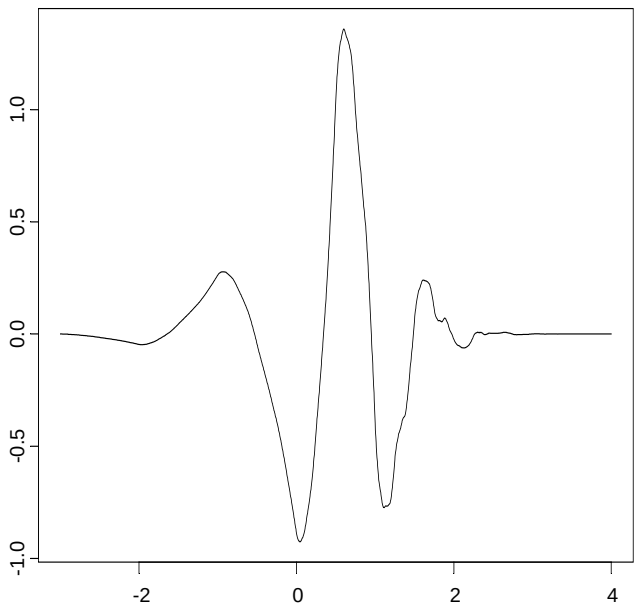
- **Wavelets:** families of orthonormal basis functions

$$g(x) = \sum_{j,k \in \Im} d^{j,k} \psi^{j,k}$$

$$\psi^{j,k} = 2^{-j/2} \psi(2^{-j/2} x - k)$$

$$d^{j,k} = \int g(x) \psi^{j,k}(x) dx$$

Daubechies Basis Function



Wavelet Space Model

- $\mathbf{d}_{abc} = \{ d_{abc}^{j,k} \} = \mathrm{W}\, \mathbf{y}_{abc}$

Discrete Wavelet Transform

$$d_{abc}^{j,k} \sim N(\theta_{abc}^{j,k}, \sigma_e^2)$$

$$\theta_{abc}^{j,k} \sim N(\theta_{ab}^{j,k}, \sigma_{1,j}^2)$$

$$\theta_{ab}^{j,k} \sim N(\theta_{a}^{j,k}, \sigma_{2,j}^2)$$

- IDWT can be used to transform wavelet space quantities back to data space
- Yields estimates of
  - $g_a(t)$, $g_{ab}(t)$, $g_{abc}(t)$
  - $\Sigma_1(t_1,t_2)$, $\Sigma_2(t_1,t_2)$

Regularization by Shrinkage Prior

- Prior on $\theta_a^{j,k} \Rightarrow$ Normal-0 mixture

$$\theta_a^{j,k} = \gamma_a^{j,k} N(0, \tau_j^2) + (1 - \gamma_a^{j,k}) \mathrm{I}_0$$

$$\gamma_a^{j,k} \sim \text{Bernoulli}(p_j)$$

- Nonlinear shrinkage – adaptive regularization
- p_j and τ_{j}^2 - regularization parameters (scale dependent regularization)

Mixture Prior and Nonlinear Shrinkage

Shrinkage Factor:

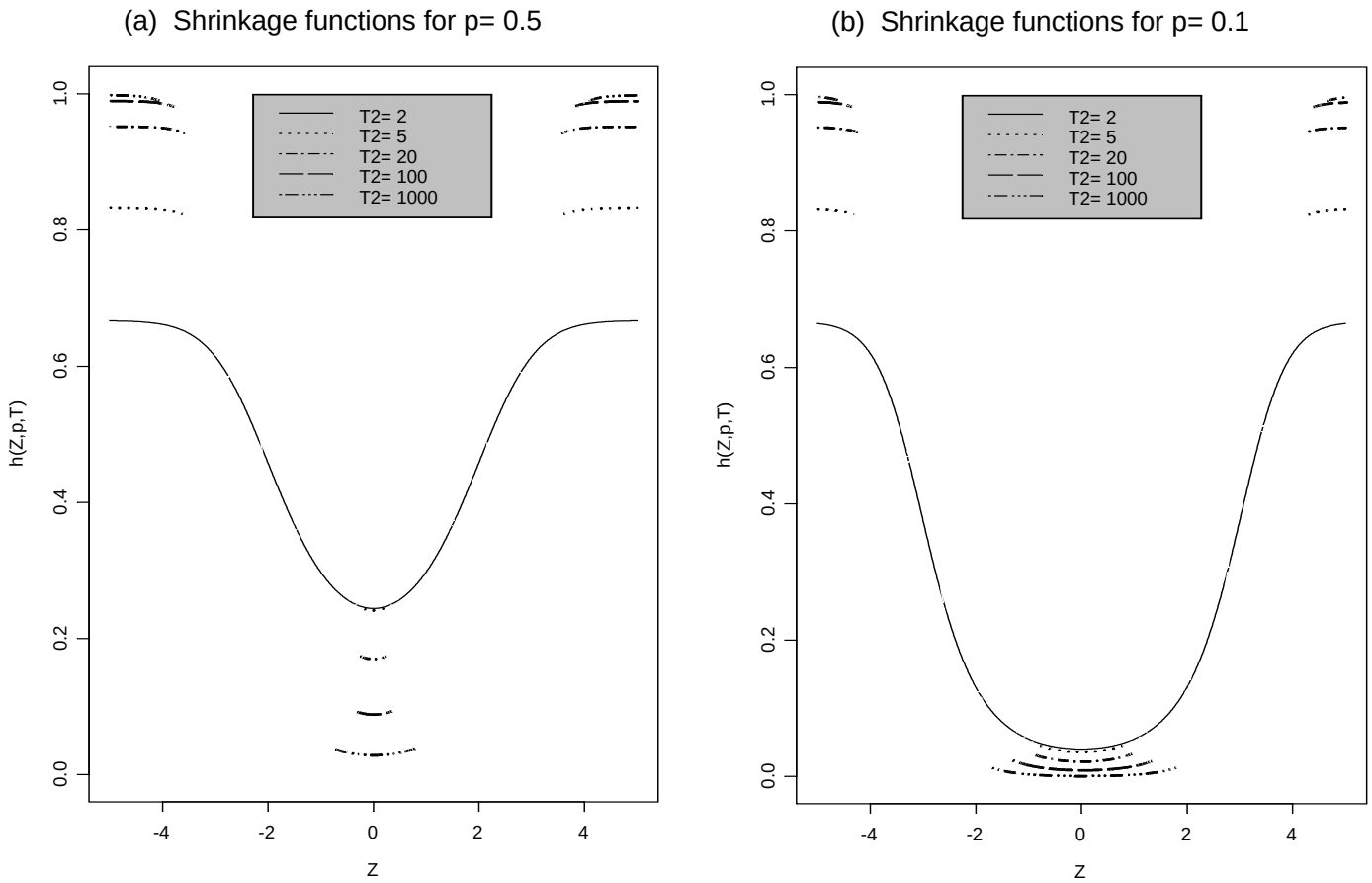
$$h(Z, T_j^2, p_j) = \underbrace{\left(\frac{T_j^2}{T_j^2+1}\right)}_{\text{Linear Shrinkage}} \underbrace{\Pr\left\{\gamma_a^{j,k} = 1 \mid \mathbf{d}_a^{j,k}\right\}}_{\text{Nonlinear Shrinkage}}$$

$$\Pr\left(\gamma_a^{j,k} = 1 \mid \mathbf{d}_a^{j,k}\right) = \frac{O^{j,k}}{O^{j,k}+1}, \qquad O^{j,k} = \text{Posterior Odds}$$

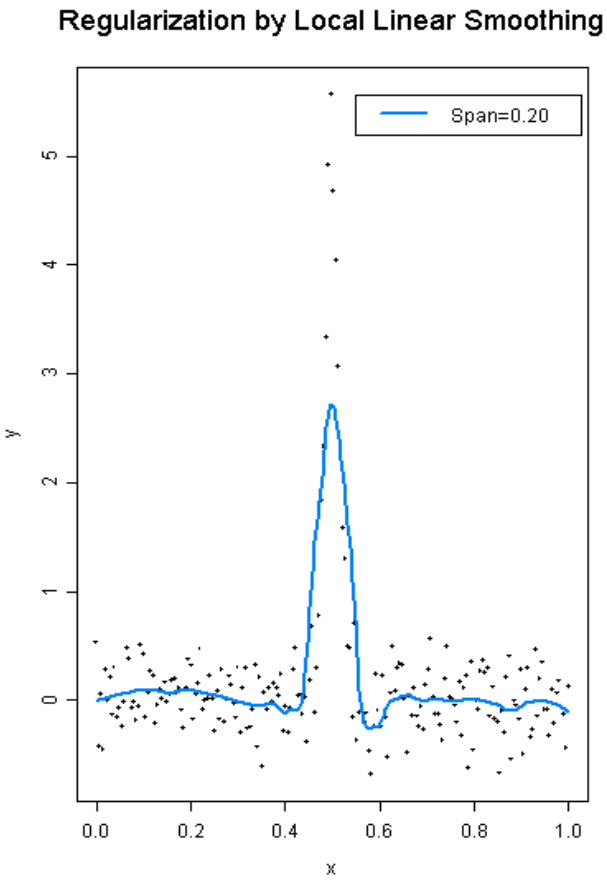
$$\underbrace{O^{j,k}}_{\text{Posterior Odds}} = \underbrace{\left(\frac{p_j}{1-p_j}\right)}_{\text{Prior Odds}} \underbrace{\left(1+T_j^2\right)^{-1} \exp\left\{\frac{(Z^{j,k})^2}{2}\left(\frac{T_j^2}{T_j^2+1}\right)\right\}}_{\text{Bayes Factor}}$$

$$Z^{j,k} = \hat{\theta}_{a,MLE}^{j,k} / \sqrt{\mathrm{Var}(\hat{\theta}_{a,MLE}^{j,k})} \qquad T_j^2 = \tau_j^2 / \mathrm{Var}(\hat{\theta}_{a,MLE}^{j,k})$$

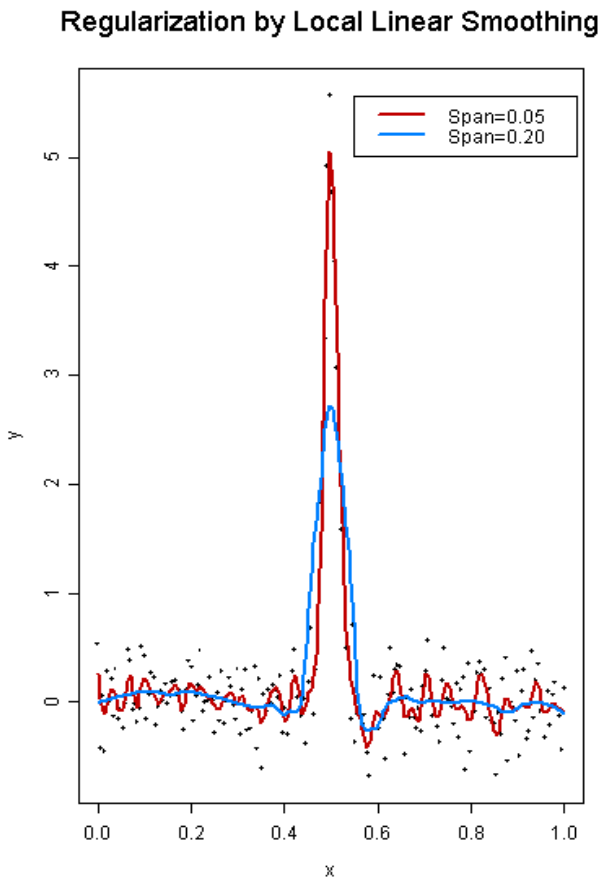
Plotting the Shrinkage Factors



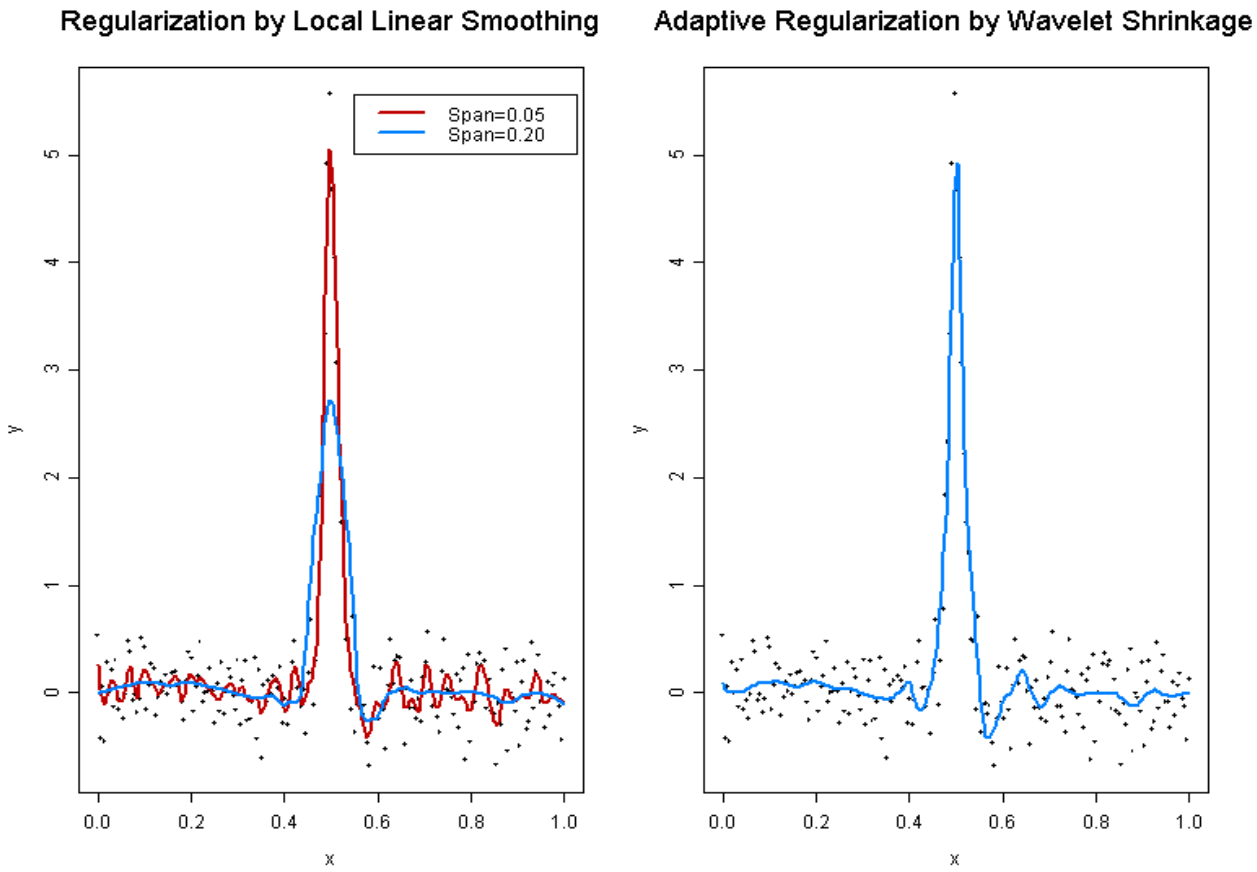
Adaptive Regularization



Adaptive Regularization



Adaptive Regularization



Fitting the Model

- Hierarchical model fit using MCMC
  - Marginal likelihood used.
  - Steps: (Ω^j=VCs at wavelet level j)

1. Sample $\theta_a^{j,k}$ from $(\theta_a^{j,k} \mid \mathbf{d}_a^{j,k}, \Omega^j)$ for each j, k, a
   - Mixture of point mass at 0 and Normal
2. Sample Ω^j from $(\Omega^j \mid \mathbf{d}^{j,k}, \theta_a^{j,k})$ for each j
   - Metropolis step used

- IDWT applied to samples of $\{\theta_a^{j,k}\}$ yield posterior samples of $g_a(t)$.

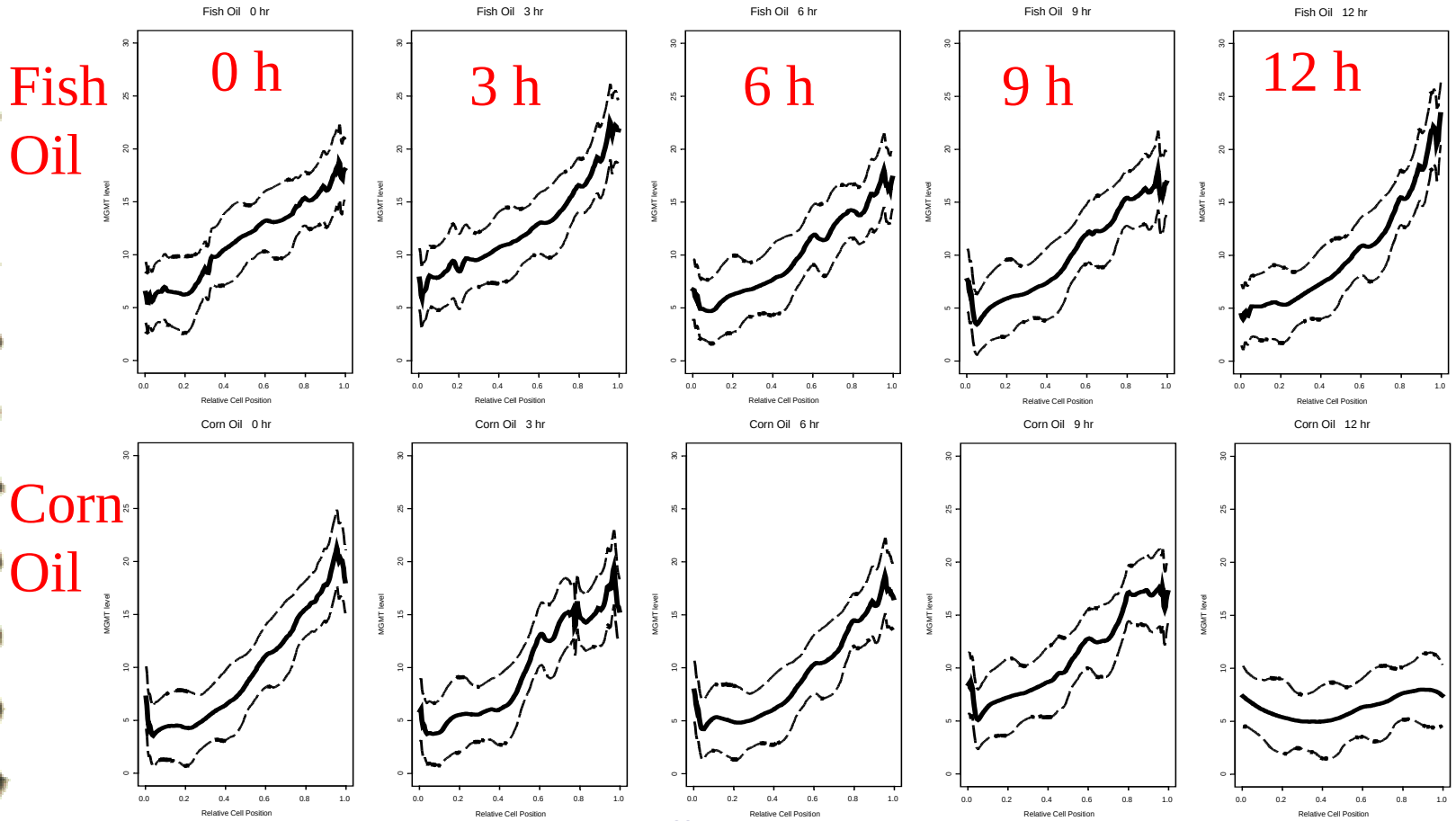
Fitting the Model

- Relative variability at hierarchical levels measured by trace of covariance matrices at each level (Σ_1, Σ_2 and $\sigma^2 I$)
- Easily computed from wavelet space VC
- Can add steps to MCMC if posterior samples for rat- and crypt-level curves are desired.

$$\boxed{\begin{array}{l} 1.\ \text{Sample from } (\theta_{ab}^{j,k} \mid \theta_a^{j,k}, \Omega^j, \mathbf{d}_{ab}^{j,k}) \text{ for each } a, b, j, k \\ 2.\ \text{Sample from } (\theta_{abc}^{j,k} \mid \theta_{ab}^{j,k}, \Omega^j, \mathbf{d}_{abc}^{j,k}) \text{ for each } a, b, c, j, k \end{array}}$$

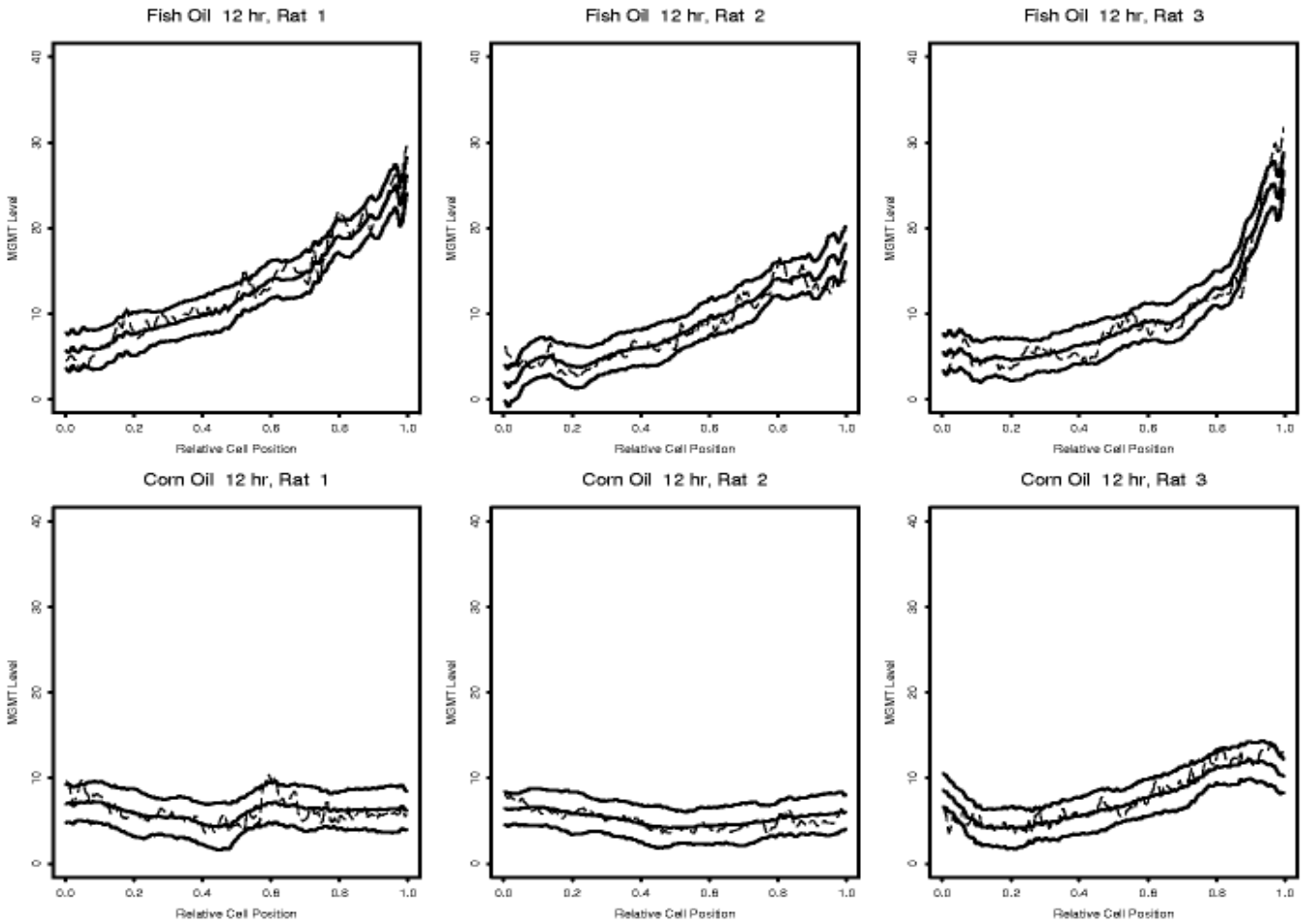
Results: DNA Repair Enzyme

Estimates & 90% posterior bounds by diet/time



Results: DNA Repair Enzyme

Estimates of individual rat profiles for T12 rats



Results: DNA Repair Enzyme

- More repair enzyme at lumenal surface
- Possible diet effect at 12 hr time point
- **Relative variability:**
  - 79% between crypts
    20% between rats
    1% within crypts
  - Important to sample lots of crypts per rat

Conclusions

- Bayesian method for hierarchical functional data
  - Nonparametrically estimate curves at mean, individual and subsampling levels using wavelet shrinkage.
  - Covariance parameter estimates
- Posterior samples from MCMC allow inference to be done
- All done with unified, ‘nonparametric’ model that appropriately adjusts for imbalance & correlation.
- Wavelet space modeling allow use of simpler covariance structures and adaptively regularizes irregular functions
- Methodology applicable beyond case study presented.

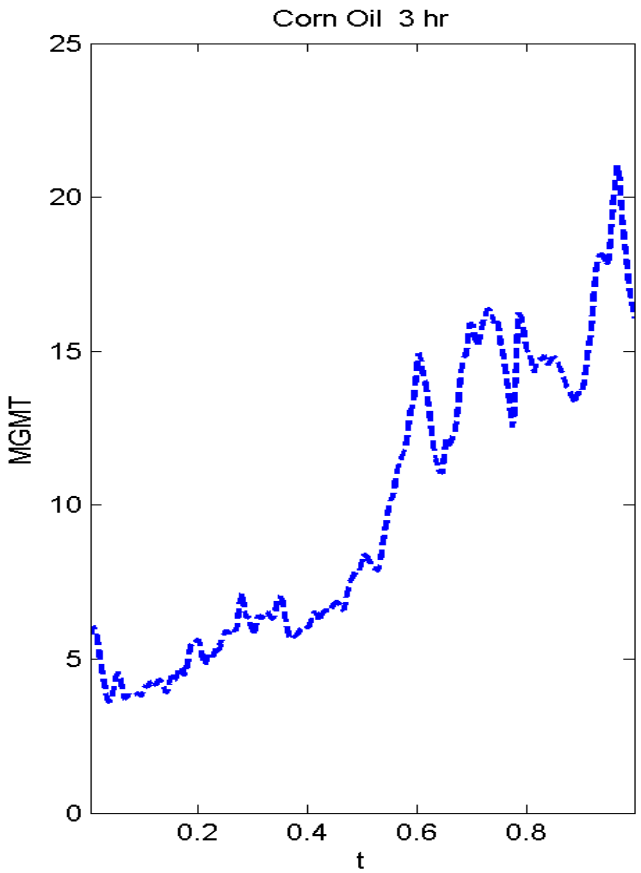
Acknowledgements

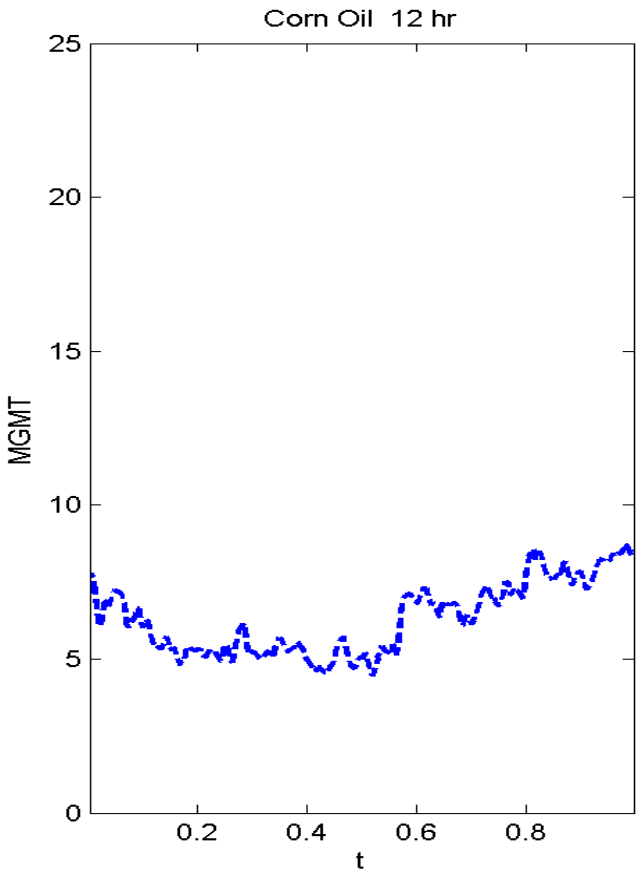
- Nutrition Collaborators:
  - Joanne Lupton
  - Rob Chapkin
  - Nancy Turner
  - MeeYoung Hong

Response to Clyde/George

- Many interesting observations/suggestions
- Focus on two here:
  1. Allow different regularization parameters for each treatment curve.
  2. Propose Empirical Bayes estimation of regularization parameters

Response to Clyde/George





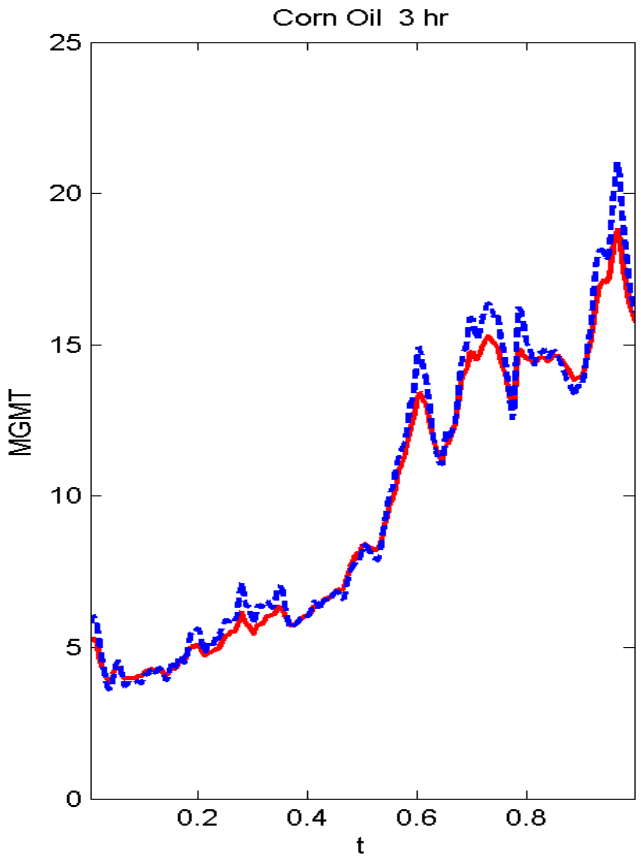
Response to Clyde/George

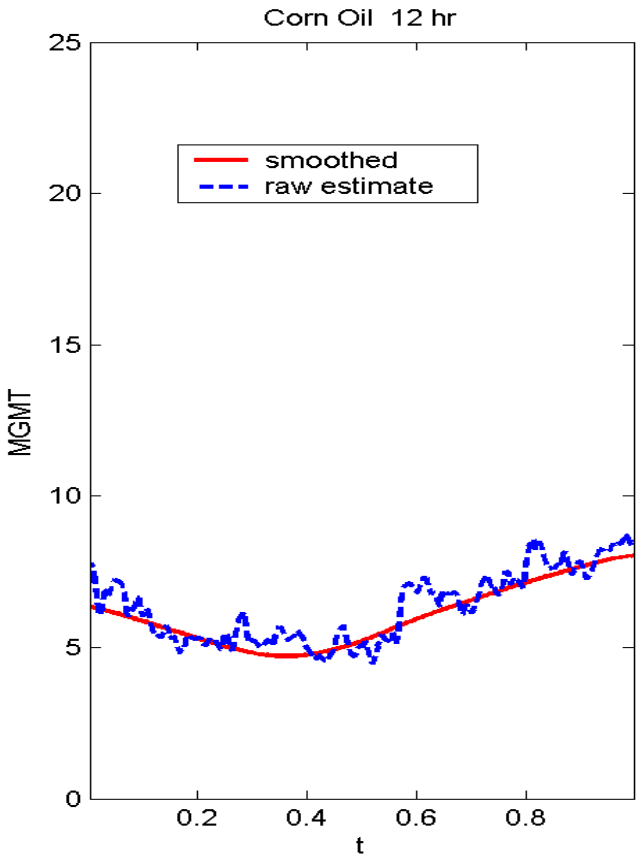
Corn/3h	Scale	6	5	4	3	2	1
p_j	1.00	0.87	1.00	0.86	0.69	0.32	0.78
T^2_j	31.7	0.89	4.23	1.49	1.74	0.41	1.43

Corn/12h

p_j	1.00	0.05	0.06	0.04	0.12	0.10	0.18
T^2_j	17.3	0.00	0.00	0.00	0.00	0.00	0.00

Response to Clyde/George



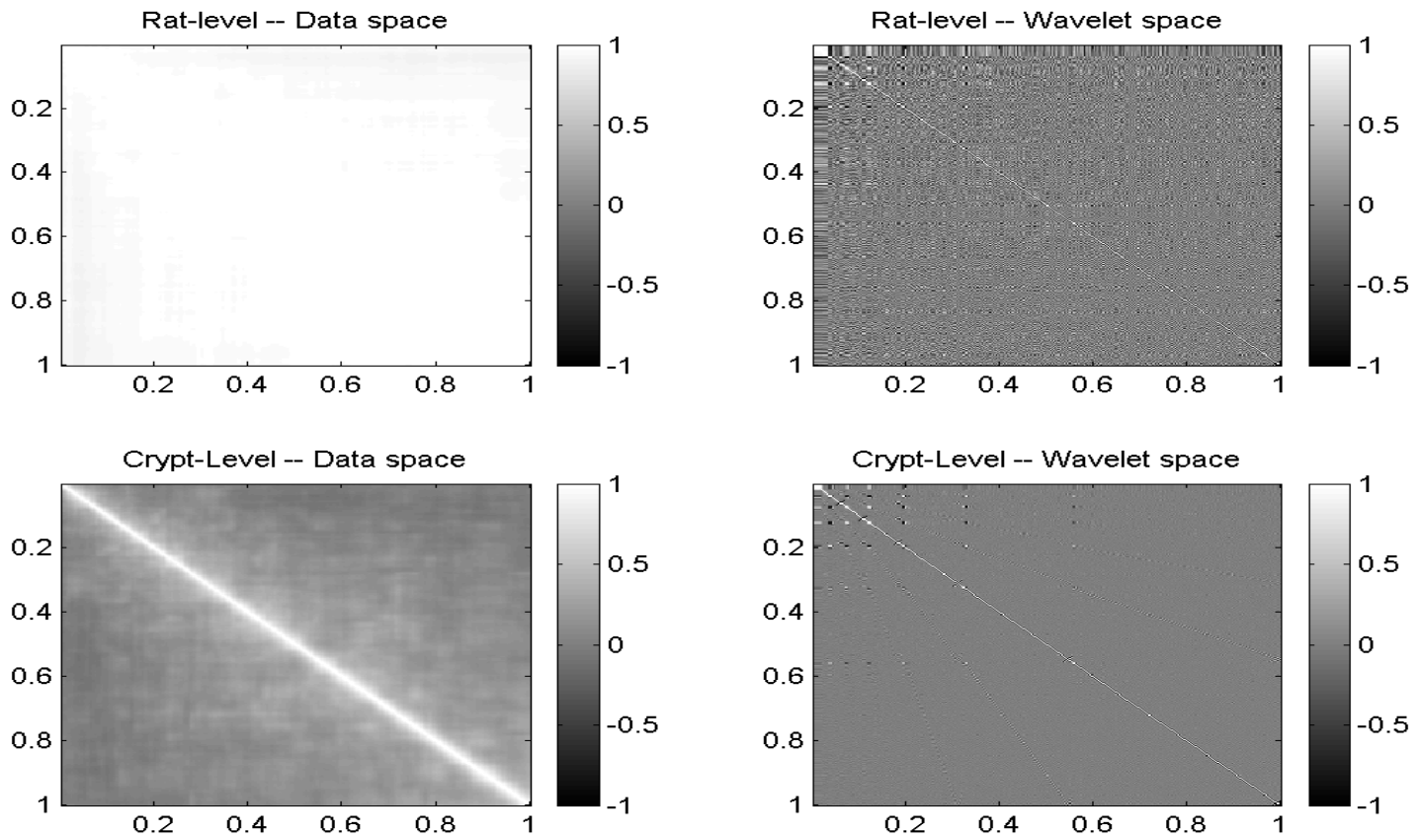


Response to Wu/Huang

- Interesting alternative modeling approach
  - Handles wavelet basis in different manner
- Practical issues to resolve:

1. Selection of basis functions
2. Modeling covariance of basis functions
3. Model-fitting
4. Performing inference

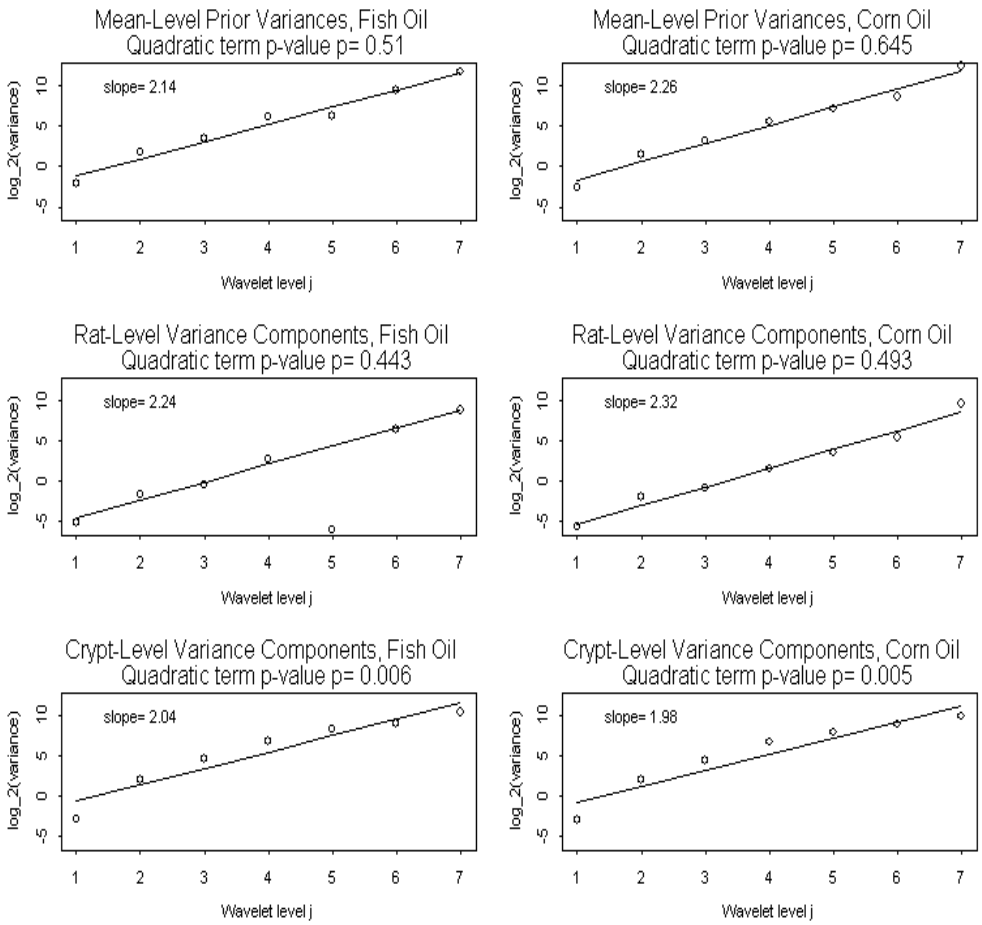
Response to Wu/Huang



Response to Kolaczyk

- Kolaczyk's theorem: The Besov space of the crypt-level functions g_{abc} depends on an interaction among the regularization parameters and the variance components at the various hierarchical levels, with the dominating variance component playing the major role.
- The data themselves contribute to the determination of the underlying space of the functions.

Response to Kolaczyk



- Crypt-level var. dominates
- ‘Broken boundary’
- Curvature in crypt-level plots
- Besov space of rat- and treatment-level functions?

Response to Kolaczyk

Simple Hierarchical Model:

$$Y_{ij} \sim N(\mu_i, \sigma^2)$$
$$\mu_i \sim N(\theta, \xi^2)$$

$$E(\mu_i \mid Y) = (\bar{Y}_i \xi^2 + \theta\sigma^2 / n) / (\xi^2 + \sigma^2 / n)$$

- Little shrinkage to mean in estimating θ_{abc}
- Considerable shrinkage to mean in estimating θ_{ab}
- θ_a shrunken differentially towards prior mean 0.
- NOTE: g_{ab} and g_{abc} are indirectly smoothed.
  - Amount of smoothing determined by relative variance components

Sensitivity Analysis