

Supplementary Figures for "Using Wavelet-Based Functional Mixed Models to Characterize Population Heterogeneity in Accelerometer Profiles: A Case Study"

List of Figures

1	<i>Summary of missingness in the data.</i> The top curve summarizes the proportion of daily profiles observed (i.e., nonmissing) as a function of the time of day. The bottom is a histogram summarizing the proportion of time between 9am and 8pm that is observed for each profile.	2
2	<i>Posterior mean and 90% pointwise band for child-to-child standard deviation function, $f(t) = \sqrt{Q(t,t)}$</i>	3
3	<i>Posterior mean and 90% pointwise band for day-to-day standard deviation function, $f(t) = \sqrt{S(t,t)}$</i>	4
4	<i>Posterior mean and 90% pointwise band for proportion of variability from day-to-day as function of time, $f(t) = S(t,t)/\{Q(t,t) + S(t,t)\}$</i>	5
5	<i>Within-Scale Decorrelation.</i> Heat maps of empirical estimates of the wavelet-space correlation matrices corresponding to Q^* and S^* for wavelets at the three finest scales, $j=1,2$, and 3 for the complete case data. Note how these plots are dominated by the diagonal, so it appears that the DWT did a reasonable job of decorrelation within the wavelet scales for our example.	6
6	<i>Correlation Matrix for White Noise and AR(1).</i> Plot of empirically estimated correlation matrix for white noise and and AR(1) process with $\rho = 0.8$, with sample sizes equivalent to the plots in Figure 5. The purpose of this plot is to serve as a comparison for Figure 5.	7
7	<i>Simulated Data.</i> We randomly generated 200 realizations from a Gaussian process with mean $\mu(t)$ and covariance $S(t_1, t_2)$ on an equally-spaced grid of length 256 on $(0,1)$. From top to bottom, column (a) contains the true mean function $\mu(t)$, the true variance function $v(t) = \text{diag}(S)$, and the true autocorrelation surface $\rho_S(t_1, t_2) = v^{-1/2}Sv^{-1/2}$. Columns (b) and (c) contain the posterior mean estimates of these quantities using wavelet-based methods. Both assume independence across wavelet coefficients, but (b) allows the wavelet-space variance components to vary across scale j and location k as in Morris and Carroll (2004), and (c) only allows them to vary across j , as assumed in Morris, et al. (2003a) and other work involving wavelet regression. Note that the framework used in (b) is sufficiently flexible to pick up on the nonstationary features of S , while (c) is not. Specifically, it is able to model the increasing variance in t , the extra variance near the peak at 0.5, the different degrees of smoothness in the region $(0,0.4)$ and $(0.6,1)$, and the extra autocorrelation from the peak at 0.5. Also note it appears to have done a marginally better job of denoising the estimate of the mean function. These same principles apply to the covariance across random effect functions.	8

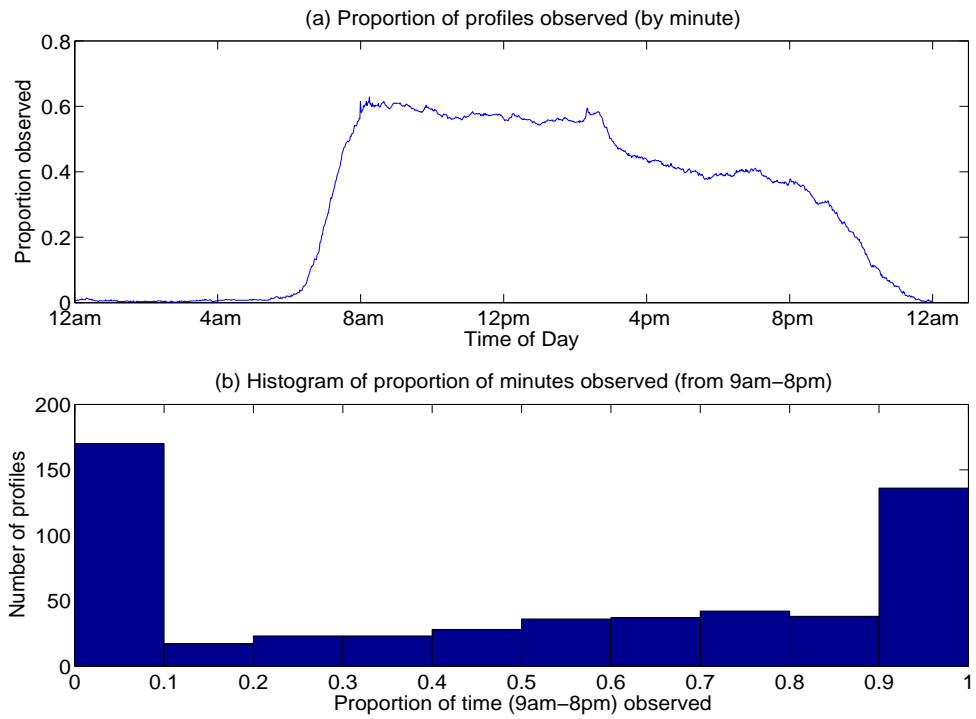


Figure 1: *Summary of missingness in the data.* The top curve summarizes the proportion of daily profiles observed (i.e., nonmissing) as a function of the time of day. The bottom is a histogram summarizing the proportion of time between 9am and 8pm that is observed for each profile.

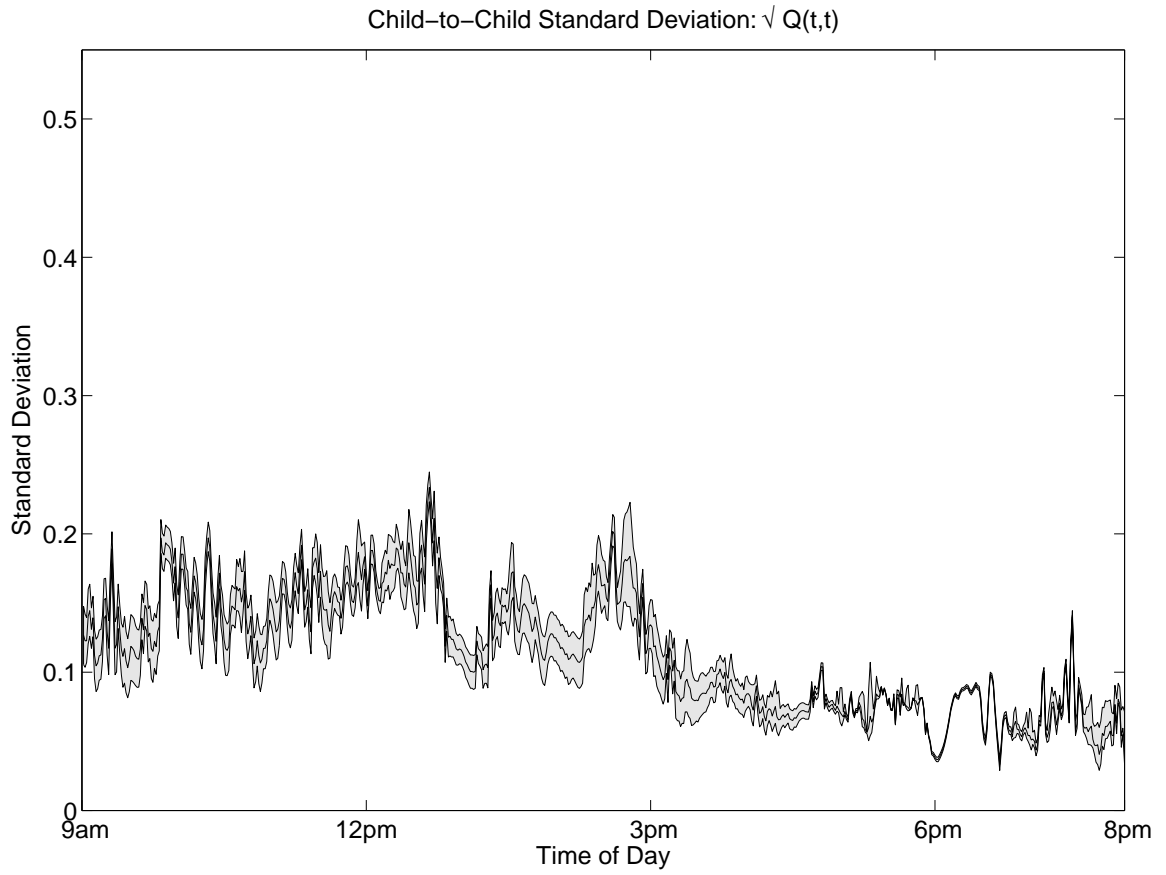


Figure 2: *Posterior mean and 90% pointwise band for child-to-child standard deviation function, $f(t) = \sqrt{Q(t,t)}$*

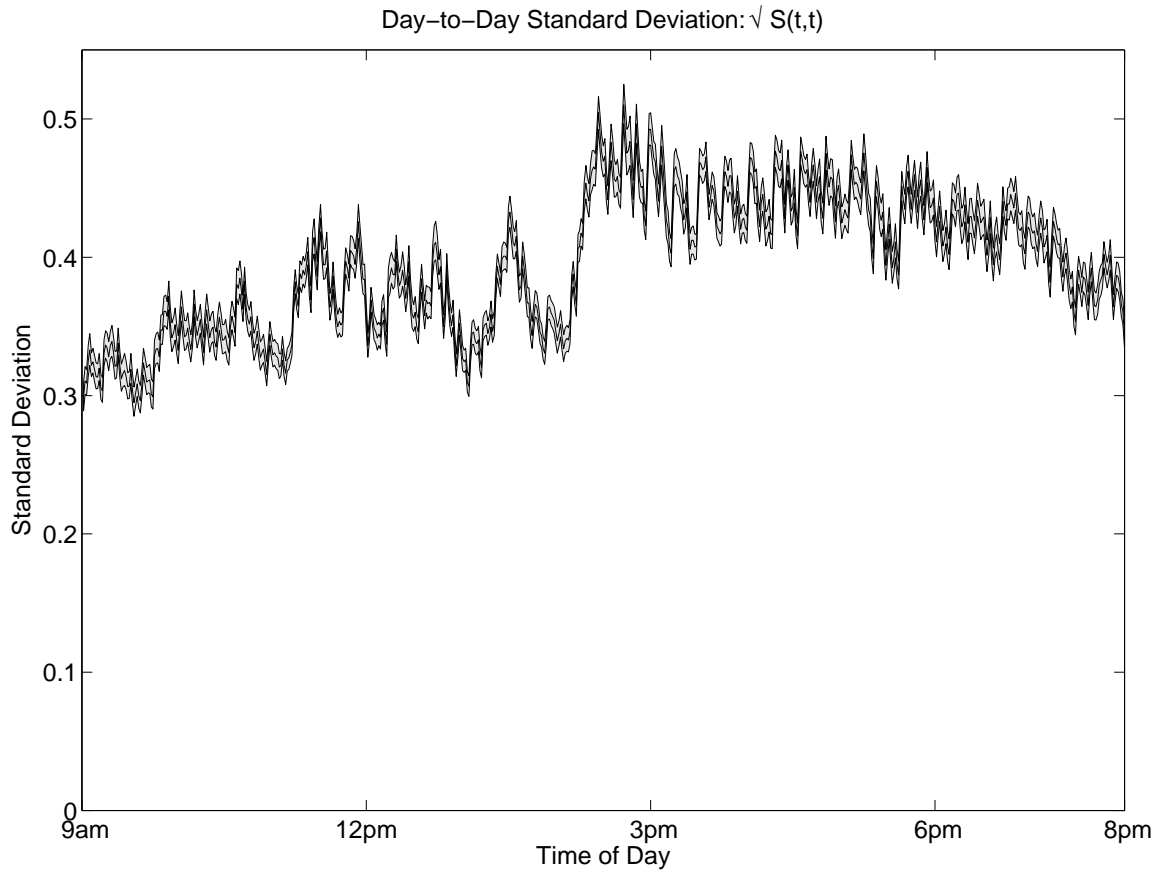


Figure 3: *Posterior mean and 90% pointwise band for day-to-day standard deviation function, $f(t) = \sqrt{S(t,t)}$*

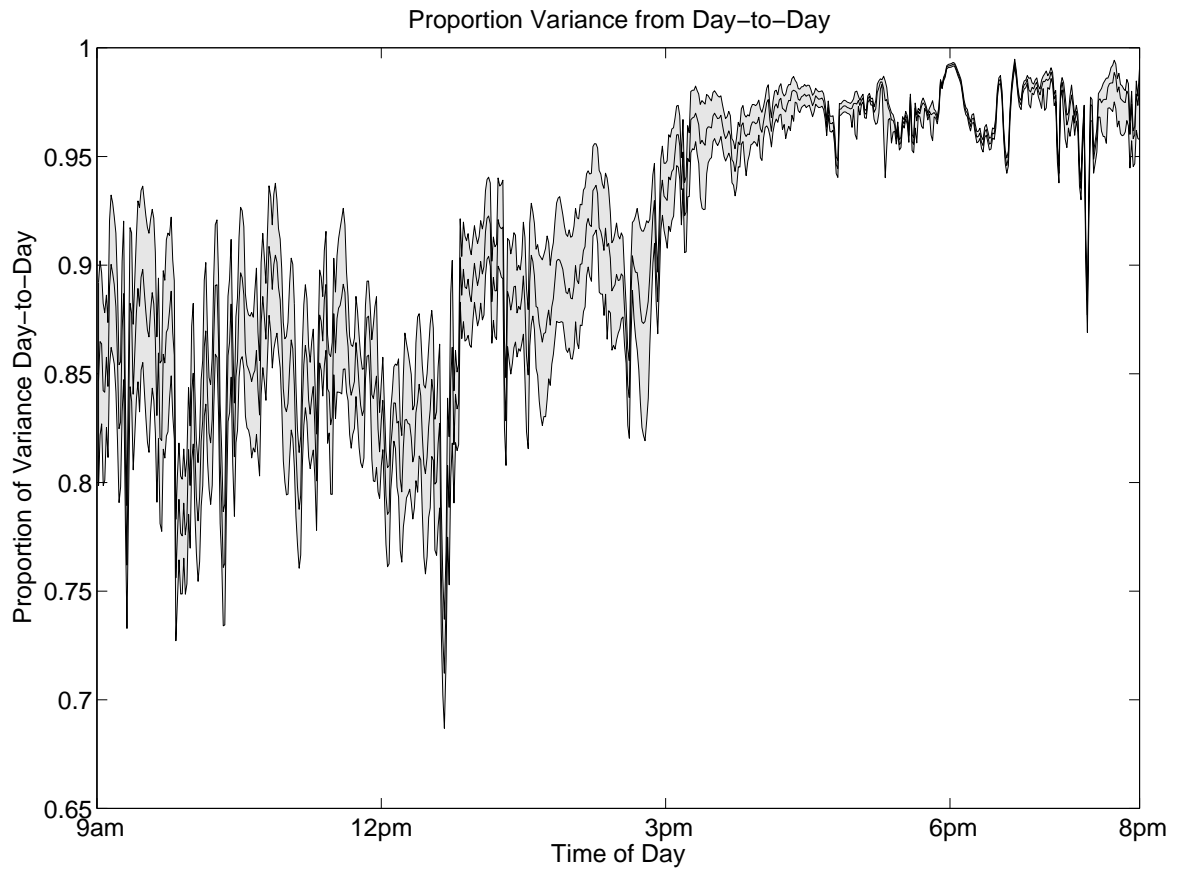


Figure 4: *Posterior mean and 90% pointwise band for proportion of variability from day-to-day as function of time, $f(t) = S(t,t)/\{Q(t,t) + S(t,t)\}$*

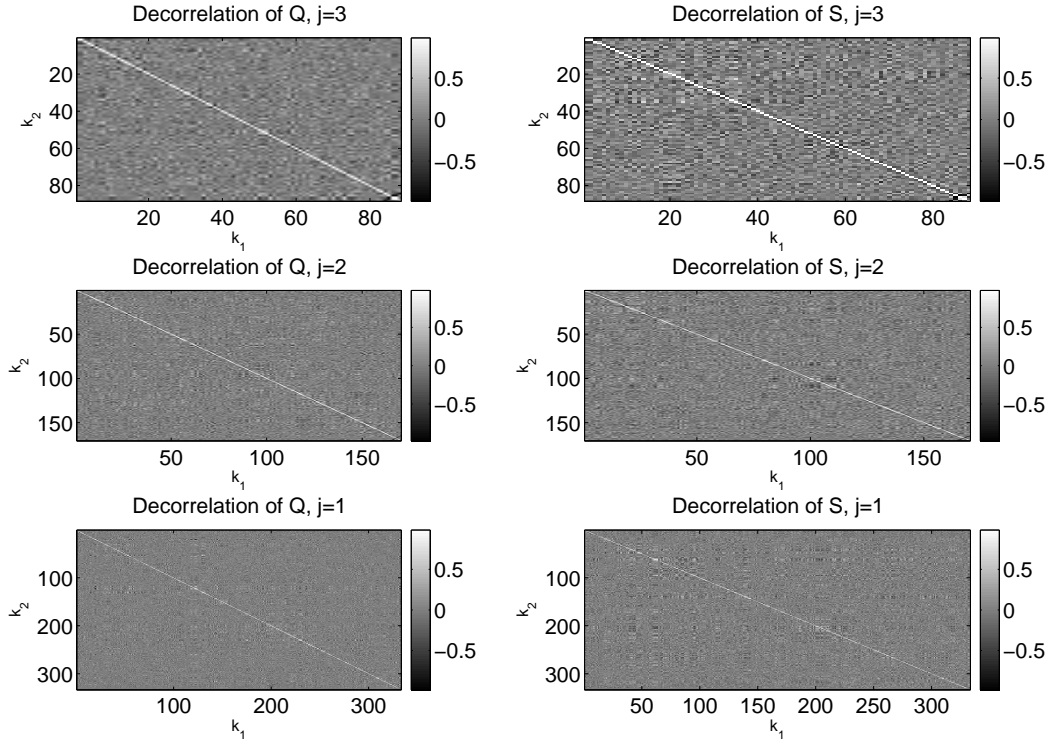


Figure 5: *Within-Scale Decorrelation*. Heat maps of empirical estimates of the wavelet-space correlation matrices corresponding to Q^* and S^* for wavelets at the three finest scales, $j=1,2$, and 3 for the complete case data. Note how these plots are dominated by the diagonal, so it appears that the DWT did a reasonable job of decorrelation within the wavelet scales for our example.

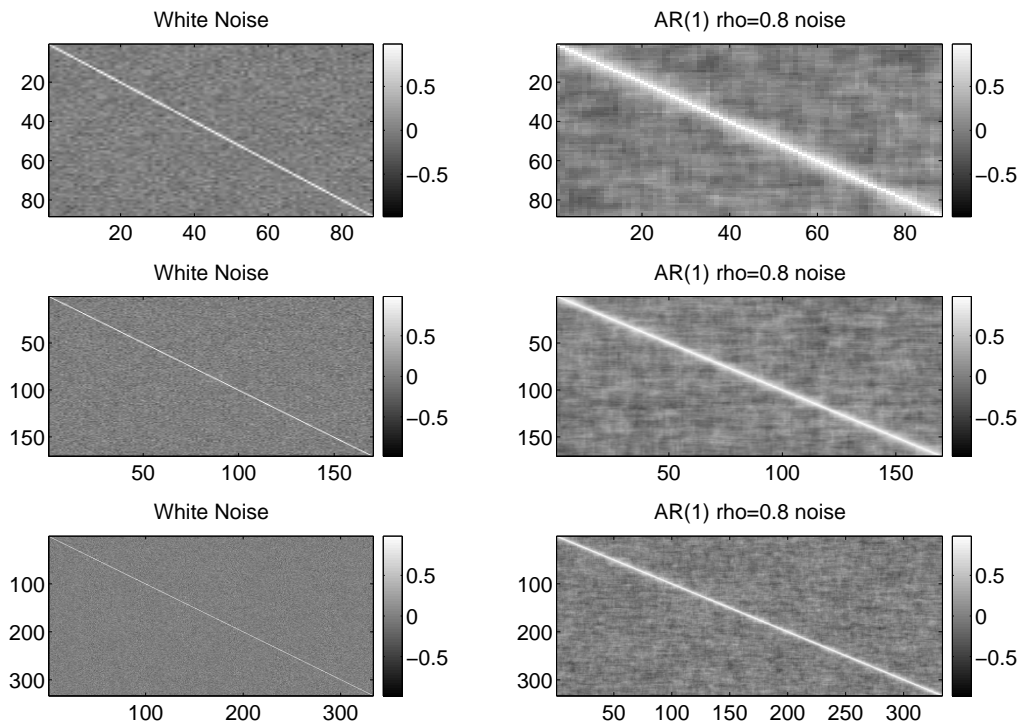


Figure 6: *Correlation Matrix for White Noise and AR(1)*. Plot of empirically estimated correlation matrix for white noise and and AR(1) process with $\rho = 0.8$, with sample sizes equivalent to the plots in Figure 5. The purpose of this plot is to serve as a comparison for Figure 5.

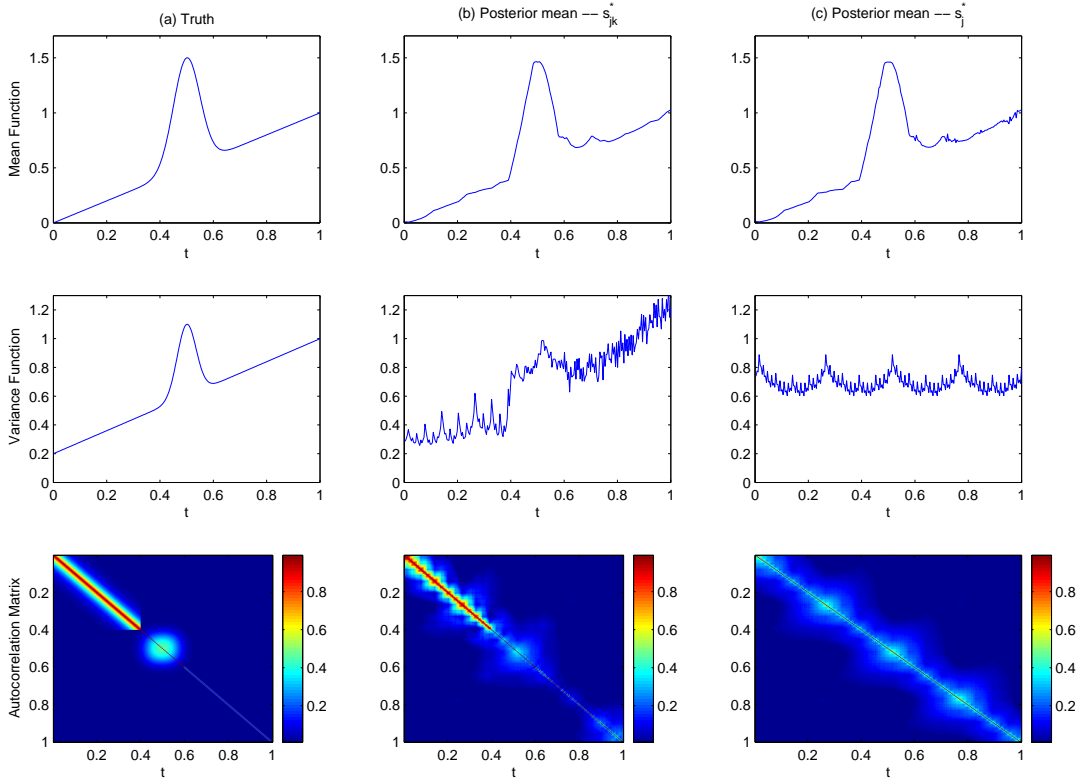


Figure 7: *Simulated Data*. We randomly generated 200 realizations from a Gaussian process with mean $\mu(t)$ and covariance $S(t_1, t_2)$ on an equally-spaced grid of length 256 on $(0, 1)$. From top to bottom, column (a) contains the true mean function $\mu(t)$, the true variance function $v(t) = \text{diag}(S)$, and the true autocorrelation surface $\rho_S(t_1, t_2) = v^{-1/2} S v^{-1/2}$. Columns (b) and (c) contain the posterior mean estimates of these quantities using wavelet-based methods. Both assume independence across wavelet coefficients, but (b) allows the wavelet-space variance components to vary across scale j and location k as in Morris and Carroll (2004), and (c) only allows them to vary across j , as assumed in Morris, et al. (2003a) and other work involving wavelet regression. Note that the framework used in (b) is sufficiently flexible to pick up on the nonstationary features of S , while (c) is not. Specifically, it is able to model the increasing variance in t , the extra variance near the peak at 0.5, the different degrees of smoothness in the region $(0, 0.4)$ and $(0.6, 1)$, and the extra autocorrelation from the peak at 0.5. Also note it appears to have done a marginally better job of denoising the estimate of the mean function. These same principles apply to the covariance across random effect functions.