**Topics by Group**

1) Group B: Background and recommendations overview paragraph.
2) Group A: Smoothing splines overview + mgcv software intro.
3) Group C: GLS model extension overview (variance and correlation models)
4) Group E: Overview of modeling/analysis approach
5) Group D: Alternate approaches (periodograms, wavelet-based methods): shortcomings/applicability

**Writing Style & Format**

Be concise. Rely on client-accessible references to support statements and *cite these liberally* in your text using the appropriate LaTeX commands. Focus your writing style and content on the client; leave all mathematical details and descriptions to the references. Again, be as concise as possible.

**Outline of Report**

1 *Background.* One paragraph (Group B).

2 *Recommendations*

Overview of recommendations/roadmap (Group B). One paragraph.

2.1 *GAMs, smoothing splines and mgcv* (Group A). Focus on the references used in your presentation.

2.1.1 BRIEF Motivation for GAMs, highlighting their applicability to the client's data and data analysis objectives.

2.1.2 High-level description of GAMs, spline bases and choices of spline bases.

2.1.3 Brief intro to the *mgcv* software package and key functions for spline specification, model fitting (gamm()) and model comparison.

2.2 *Generalized Least Squares (GLS) model extensions* (Group C). Focus on the references used in your presentation.

2.2.1 How is residual heteroskedasticity detected and characterized?

2.2.2 How are temporal patterns of correlation in residuals detected and characterized? Focus on the variogram and methods for irregularly spaced data.

2.2.3 Briefly list and describe adjustments to mixed models (nlme::lme()) and to GAMs (mgcv::gamm(); same as lme()) to account for heteroskedasticity and residual autocorrelation.

2.2.4 What is the process for evaluating competing formulations? How does the analyst arrive at a model/GAM formulation that sufficiently accounts for the temporal covariation present in a base model's residuals?

2.3 *Modeling/analysis approach* using mgcv::gamm() (Group E). Provide step-by-step guidance for constructing a model for the nitrogen fixation data. See, for example the references in Zuur et al. and Pedersen et al. (all posted on Sakai).

2.3.1 Data. Temporally aggregate the time varying variables (nitrogen fixation, and other covariates) to a common frequency on the scale of minutes. Create derived variables: time-of-day (0-24 scale), day-of-year (nominal/integer), day-of-year (continuous). Assemble all relevant variables into a data matrix.

2.3.2 Model mean structure. Describe a basic progression of models for the mean. For example, start with: daily cyclic spline, additive (unconstrained) day-of-year random effects spline (bs="re") and covariates (if any). Then consider as an alternative a smooth day-of-year spline (bs="tp"), then a joint model for daily and long-range variation.

2.3.3 Examine, then model variance structure: Examine residuals for evidence of heteroskedasticity. If present, consider variance models (details in section 2.2)

2.3.4 Examine, then model residual correlations: Examine residuals for evidence of temporal autocorrelation. If present, consider continuous correlation structure models (details in section 2.2).

2.4 A*lternate approaches: promise and shortcomings.* (Group D). Base this primarily on the Cazellas et al. (2008) and Ruf (1999) references (on Sakai).

2.4.1 Brief introduction to periodograms (including the Lomb-Scargle variant) and their shortcomings/obstacles to implementation for this analysis.

2.4.2 Wavelet-based time-scale decompositions, their utility for ecological time series and obstacles to implementation with the nitrogen fixation data.

2.4.3 Any other alternatives?