GE 509 – Applied Env Statistics
Model Description

DUE Friday March 6th, 11:59pm

1-2 pages, double spaced

Synopsis: A brief description of *how* the data will be analyzed. Should include a mathematical specification of the process model(s), the data model, and the parameter model and a figure of how these relate to one another (e.g. a model graph).

In this, the second step of your independent project, you will build upon the ideas presented in your project proposal. The goal of this assignment is to move from a *qualitative* statement of hypotheses to a *mathematical* statement of a model or set of models that will be fit to data. Include a brief justification/rational for why you are proposing the model(s) that you are proposing (e.g. logical/mechanistic reasons for that functional form, biological interpretation of the parameters, or visual assessment during exploratory data analysis)

The description of the model should include at least two parts, the process model and the data model. If you are conducting a Bayesian analysis it should also include the parameter model.
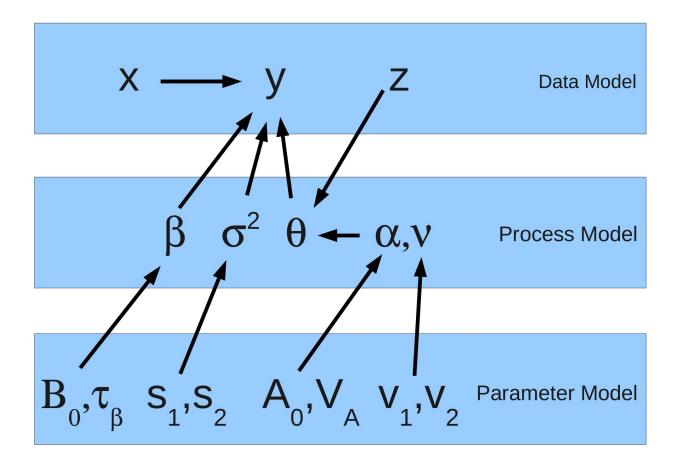
The process model(s) need to be written out explicitly, for example

$$g(x) = \beta_0 + \beta_1 \frac{x}{\theta + x}$$

For the data model and the parameter models, probability distributions do NOT need to be written out explicitly, but any other mathematic relationships do need to be explicit. Distributions can be named either by a standard abbreviation or by their R/BUGS name. For example

$y \sim N(g(x), \sigma^2)$		$y \sim dnorm(g(x), \sigma^2)$
$\vec{\beta} \sim N(0, \tau_\beta)$		$\vec{\beta} \sim dnorm(0, \tau_\beta)$
$\sigma^2 \sim IG(s_1, s_2)$		$\sigma^2 \sim dinvgamma(s_1, s_2)$
$\theta \sim \text{TruncatedNormal}(h(x), \nu)$	VS.	$\theta \sim dtnorm(h(x), \nu)$
$h(x) = \alpha_0 + \alpha_1 z$		$h(x) = \alpha_0 + \alpha_1 z$
$\vec{\alpha} \sim N(A_0, V_\alpha)$		$\vec{\alpha} \sim dnorm(A_0, V_\alpha)$
$\nu \sim IG(v_1, v_2)$		$\nu \sim dinvgamma(v_1, v_2)$

The model description should also include a graph of the model

Data Model

$X \longrightarrow y \qquad z$

Process Model

$\beta \quad \sigma^2 \quad \theta \leftarrow \alpha, \nu$

Parameter Model

$B_0, \tau_\beta \quad s_1, s_2 \quad A_0, V_A \quad v_1, v_2$

Finally, one does not need to go into detail on how the analysis will be performed, but give a brief statement or outline of what you are thinking of doing and relate this back to the hypotheses you are testing. For example:

Models will be fit using a combination of Gibbs sampling and Metropolis-Hastings. Fit will be assessed by standard MCMC diagnostics. Model A vs Model B will be compared based on DIC and if Model A is a better fit I will interpret this as XXXX while if Model B is a better fit we will interpret this as YYYYY. Credible intervals will be constructed for all parameters and credible and predictive intervals will be constructed for the whole model. If the 95% CI for parameter b0 is different from ZZZ I will interpret this as support for of the hypothesis that *****. I will use classical variance decomposition to determine what parameters in the model contribute most to model uncertainty, which will help me in the decision to measure more of UUUU vs VVVV. Finally, I will use numerical simulation to estimate the power of my current analysis by and to estimate the sample size of UUUU needed to constrain b1 to a SE of 5%.

(Note, this last example is probably more involved than your analysis needs to be...remember, I expect you to have output from this analysis by April 10th)