Characterizing Uncertainty
Linear Model

\[ y_i \sim \beta_0 + \beta(x_i) + \varepsilon_i \]
Linear Model

Classic Assumptions

- Error in Y is measurement error
- Normally distributed error
- Observations are independent

\[ y_i \sim \beta_0 + \beta(x_i) + \varepsilon_i \]
Linear Model

Classic Assumptions

- Error in Y is measurement error
- Normally distributed error
- Observations are independent
- Homoskedasticity

\[ y_i \sim \beta_0 + \beta(x_i) + \epsilon_i \]
Linear Model

Classic Assumptions

- Error in Y is measurement error
- Normally distributed error
- Observations are independent
- Homoskedasticity
- No error in X variables

\[ y_i \sim \beta_0 + \beta(x_i) + \varepsilon_i \]
Linear Model

\[ y_i \sim \beta_0 + \beta(x_i) + \epsilon_i \]

Data Model

\[ Y \]

Process Model

\[ \beta_0 + \beta_1(x_i) \]

Parameters

\[ \beta \]

response

quality
Linear Model – Graph Notation

\[ y_i \sim \beta_0 + \beta(x_i) + \varepsilon_i \]
Linear Model –

Data Model

Process Model

Parameter Model

\[ X \rightarrow Y \]

\[ \beta , \varepsilon \]

\[ b1, b.v \quad T \]
Data & Distributions

\( y \sim N(\mu, \varepsilon) \)
Expected relationship between data samples and the range of all possible data is described by a probability distribution.

\[ y \sim N(\mu, \epsilon) \]
Data & Distributions

![Frequency Chart]

- Frequency scale from 0 to 8
- Y-axis represents values
- Bar chart showing frequency distribution
Data & Distributions

\[ \exp \left[ -\frac{(y-\mu)^2}{2\sigma^2} \right] \]
Data & Distributions

\[ \exp \left[ -\frac{(y-\mu)^2}{2\sigma^2} \right] \]

\[ \frac{\lambda^y}{y!} e^{-\lambda} \]
Uncertainty & Variance
Uncertainty & Variance
Uncertainty & Variance

\[ \varepsilon_i \sim \mathcal{N}(0, \tau) \]

Data Model

\[ X \rightarrow Y \]

Process Model

\[ \beta, \varepsilon \]

Parameter Model

\[ b_1, b, v \]
Uncertainty & Variance: Heteroskedasticity

\[ y \sim N(\beta_1 + \beta_2 x, (\alpha_1 + \alpha_2 x)^2) \]
Uncertainty & Variance: Heteroskedasticity

\[ y \sim N(\beta_1 + \beta_2 x, s^2) \]

\[ y \sim N(\beta_1 + \beta_2 x, (\alpha_1 + \alpha_2 x)^2) \]
Regression model usually assumes all error is in $Y$ (measurement of response)
Uncertainty & Variance

But we’re not always great at measuring x either...
Uncertainty & Variance: Errors in Variables

Model $x$ as random variable

$$x^0 \sim N(x, \delta)$$

$$y \sim N(\beta_1 + \beta_2 x, \varepsilon)$$
Latent Variables

- State not directly observed
  - Missing data
Latent Variables

Missing Data

![Graph showing missing data over time](image-url)
Latent Variables

\[ X \rightarrow Y \]

\[ X_{\text{mis}} \rightarrow \beta, \varepsilon \]

\[ X_0, X_v \rightarrow B_0, B_v, \tau \]

Missing Data
Latent Variables

Missing Data

- Update the regression model conditioned on the current values of the missing data
- Update the missing data based on the current regression model and covariate values
Latent Variables

Missing Data

data is missing at random
Latent Variables

$X \rightarrow Y$

$X_{\text{mis}} \rightarrow \beta, \varepsilon$

$X_0, X_v \rightarrow B_0, B_v, \tau$

Missing Data

data is missing at random
Latent Variables

• State not directly observed
  – Missing data
  – Proxy measures
Latent Variables

- State not directly observed
  - Missing data
  - Proxy measures

Will rising atmospheric CO$_2$ increase tree fecundity?
Latent Variables

Ignoring variable latency can lead to incorrect or falsely overconfident conclusions (and wonky bad forecasts)
Latent Variables

Will rising atmospheric CO$_2$ increase tree fecundity?

Duke Free Air Carbon Enrichment (FACE)
Latent Variables

Will rising atmospheric CO\(_2\) increase tree fecundity?

Fecundity \(\sim f(CO_2)\)
Latent Variables

Will rising atmospheric CO$_2$ increase tree fecundity?

Seeds/m$^2$ $\sim$ f(CO$_2$)
Latent Variables

Will rising atmospheric CO$_2$ increase tree fecundity?

Seeds/m$^2$ $\sim f$(CO$_2$)

Seed production $\sim f$ (tree size)

Seeds observed $\sim f$(trees/m$^2$, dispersal)
Latent Variables

Will rising atmospheric CO\textsubscript{2} increase tree fecundity?

\[
\text{Fecundity} \sim f(\text{CO}_2 + \text{tree size}) \& f(\text{Obs})
\]
Hierarchical Bayes
Hierarchical Bayes

Independent

Shared
Hierarchical Bayes

Independent

Shared
Hierarchical Bayes

Independent

Shared
Hierarchical Bayes

Independent

Hierarchical

Shared
Hierarchical Bayes

Model variability in the parameters
Partition variability explicitly into multiple terms
Borrow strength across data sets
Hierarchical Bayes

Predicting $X_4$?

Independent

Hierarchical

Shared
Hierarchical Bayes

Model variability in the parameters
Partition variability explicitly into multiple terms
Borrow strength across data sets
Hierarchical with respect to parameters
Pay attention to subscripts!
Hierarchical Bayes
Hierarchical Bayes

\[ Y_k \sim N(u_k, \sigma^2) \]
Hierarchical Bayes

\[ Y_k \sim N(u_k, \sigma^2) \]
\[ u_k \sim N(u, \tau^2) \]
\[ \sigma^2 \sim IG(s_1, s_2) \]
Hierarchical Bayes

\[ Y_k \sim N(u_k, \sigma^2) \]
\[ u_k \sim N(u, \tau^2) \]
\[ \sigma^2 \sim IG(s_1, s_2) \]
\[ u \sim N(u_0, \nu_m) \]
\[ \tau^2 \sim IG(t_1, t_2) \]
Hierarchical Bayes – random effects model

\[ Y_k \sim N(u_k, \sigma^2) \]

\[ u_k \sim N(u, \tau^2) \]

\[ \sigma^2 \sim IG(s_1, s_2) \]

\[ Y_k \sim N(\mu_g + \alpha_k, \sigma^2) \]

\[ u \sim N(\mu_0, \nu_m) \]

\[ \tau^2 \sim IG(t_1, t_2) \]
Hierarchical Bayes – random effects

Random effects attempt to account for the unexplained variance associated with some group (plot, year, etc.) due to all the things that were not measured.

May point to scales that need additional explanation.

Adding covariates may explain some portion of this variance, but there's always something you didn't measure.
Hierarchical Bayes – random effects model

\[ Y_k \sim N(u_k, \sigma^2) \]
\[ u_k \sim N(u, \tau^2) \]
\[ \sigma^2 \sim IG(s_1, s_2) \]
\[ u \sim N(u_0, \nu_m) \]
\[ \tau^2 \sim IG(t_1, t_2) \]

\[ Y_k \sim N(u_g + \alpha_k, \sigma^2) \]
\[ \alpha_k \sim N(0, \tau^2) \]
\[ \sigma^2 \sim IG(s_1, s_2) \]
\[ u_g \sim N(u_0, \nu_m) \]
\[ \tau^2 \sim IG(t_1, t_2) \]
Hierarchical Bayes – random effects model

\[ Y_k \sim N(u_g + \alpha_k , \sigma^2) \]
\[ \alpha_k \sim N(0, \tau^2) \]
\[ \sigma^2 \sim IG(s_1, s_2) \]
\[ u_g \sim N(u0, \nu_m) \]
\[ \tau^2 \sim IG(t1, t2) \]
Hierarchical Bayes – random effects model

\[ Y_k \sim N(u_g + \alpha_k, \sigma^2) \]

REs have mean 0

\[ \alpha_k \sim N(0, \tau^2) \]

\[ \sigma^2 \sim IG(s_1, s_2) \]

\[ u_g \sim N(u_0, \nu_m) \]

\[ \tau^2 \sim IG(t1, t2) \]
Hierarchical Bayes – random effects model

\[ Y_k \sim N(u_g + \alpha_k, \sigma^2) \]

REs have mean 0

\[ \alpha_k \sim N(0, \tau^2) \]

\[ \sigma^2 \sim IG(s_1, s_2) \]

REs variance attributes portion of uncertainty to specific source

\[ u_g \sim N(u0, \nu_m) \]

\[ \tau^2 \sim IG(t1, t2) \]
Hierarchical Bayes – random effects model

\[ Y_k \sim N(\mu_g + \alpha_k, \sigma^2) \]
\[ \alpha_k \sim N(0, \tau^2) \]
\[ \sigma^2 \sim IG(s_1, s_2) \]
\[ \mu_g \sim N(\mu_0, V_\mu) \]
\[ \tau^2 \sim IG(t_1, t_2) \]

(a) Large individual effects
(b) Large process “error”

\[ \text{tau}^2 = 0.09, \text{sig}^2 = 0.01 \]
\[ \text{tau}^2 = 0.01, \text{sig}^2 = 0.09 \]
Hierarchical Bayes – random effects model

RE’s traditionally aspects of the study that would likely change if experiment/observation protocol was replicated

• *e.g.*, *Plots, Years*

• used to account for lack of independence among reps
Hierarchical Bayes – random effects model

RE’s traditionally aspects of the study that would likely change if experiment/observation protocol was replicated

- *e.g.*, *Plots, Years*
- used to account for lack of independence among reps

Replication is important for identifiability (i.e., partitioning variance between process error and REs)
Hierarchical Bayes – random effects model

RE’s traditionally aspects of the study that would likely change if experiment/observation protocol was replicated

- *e.g.*, *Plots, Years*

- used to account for lack of independence among reps

Replication is important for identifiability (i.e., partitioning variance between process error and REs)

Treatments/covariates of interest are treated as fixed effects (mixed effects model)
Global Mean = 

model{
    mu ~ dnorm(0, 0.001)
    tau ~ dgamma(0.001, 0.001)

    for(i in 1:nblocks){
        for(t in 1:ndates){
            X[i, t] ~ dnorm(mu, tau)
        }
    }
}
Global Mean =
model{
    mu ~ dnorm(0, 0.001)
    tau ~ dgamma(0.001, 0.001)

    for(i in 1:nblocks){
        for(t in 1:ndates){
            X[i, t] ~ dnorm(mu, tau)
        }
    }
}
Global Mean = 
model{
    mu ~ dnorm(0, 0.001)
    tau ~ dgamma(0.001, 0.001)

    for(i in 1:nbblocks) {
        for(t in 1:ndates) {
            X[i, t] ~ dnorm(mu, tau)
        }
    }
}

Space = 
model{
    mu ~ dnorm(0, 0.001)
    tau ~ dgamma(0.001, 0.001)
    tau.sp ~ dgamma(0.001, 0.001)

    for(i in 1:nbblocks) {
        alpha.sp[i] ~ dnorm(0, tau.sp)
        Emu[i] = mu + alpha.sp[i]
    }

    for(t in 1:ndates) {
        X[i, t] ~ dnorm(Emu[i], tau)
    }
}
Global Mean = 
\[
\text{model}\
\begin{align*}
\mu & \sim \text{dnorm}(0, 0.001) \\
\tau & \sim \text{dgamma}(0.001, 0.001)\
\end{align*}
\]

for (i in 1:nbblocks){
  for (t in 1:ndates){
    X[i,t] ~ dnorm(\mu, \tau)
  }\}

Space = 
\[
\text{model}\
\begin{align*}
\mu & \sim \text{dnorm}(0, 0.001) \\
\tau & \sim \text{dgamma}(0.001, 0.001) \\
\tau_{sp} & \sim \text{dgamma}(0.001, 0.001)\
\end{align*}
\]

for (i in 1:nbblocks){
  alpha_{sp}[i] ~ \text{dnorm}(0, \tau_{sp})
  Emu[i] = \mu + alpha_{sp}[i]
}

for (t in 1:ndates){
  X[i,t] ~ dnorm(Emu[i], \tau)
}\}
Global Mean:
model{
    mu ~ dnorm(0,0.001)
    tau ~ dgamma(0.001,0.001)
}

for(i in 1:nblocks){
    for(t in 1:ndates){
        X[i,t] ~ dnorm(mu, tau)
    }
}

Space:
model{
    mu ~ dnorm(0,0.001)
    tau ~ dgamma(0.001,0.001)
    tau.sp~dgamma(0.001,0.001)
    for(i in 1:nblocks){
        alpha.sp[i]~dnorm(0,tau.sp)
        Emu[i]=mu + alpha.sp[i]
    }
    for(t in 1:ndates){
        X[i,t] ~ dnorm(Emu[i], tau)
    }
}

Parameter Model
(Priors)
Hyperprior

Process Model

Data Model
model{
  mu ~ dnorm(0,0.001)
  tau ~ dgamma(0.001,0.001)
  tau.sp~dgamma(0.001,0.001)
  for(i in 1:nblocks){
    alpha.sp[i]~dnorm(0,tau.sp)
    Emu[i]=mu + alpha.sp[i]
  }
  for(t in 1:ndates){
    X[i,t] ~ dnorm(Ex[i], tau)
  }
}
model{
  mu ~ dnorm(0, 0.001)
  tau ~ dgamma(0.001, 0.001)
  tau.sp ~ dgamma(0.001, 0.001)
  for(i in 1:nblocks){
    alpha.sp[i] ~ dnorm(0, tau.sp)
    Emu[i] = mu + alpha.sp[i]
  }
  for(t in 1:ndates){
    X[i, t] ~ dnorm(Ex[i], tau)
  }
}

"
<table>
<thead>
<tr>
<th>Model</th>
<th>mu (95% ci)</th>
<th>tau (sigma)</th>
<th>tau.sp (sigma)</th>
<th>Deviance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global</td>
<td>2.99 (2.98, 3.13)</td>
<td>0.85 (1.18)</td>
<td>NA</td>
<td>698</td>
</tr>
<tr>
<td>Spatial RE</td>
<td>2.97 (2.66, 3.28)</td>
<td>1.04 (0.96)</td>
<td>4.98 (0.20)</td>
<td>654</td>
</tr>
</tbody>
</table>
defines a distribution that can be used to infer mean at out-of-sample site. Tighter REs (smaller variance) = more precise prediction

<table>
<thead>
<tr>
<th>Model</th>
<th>mu (95% ci)</th>
<th>tau (sigma)</th>
<th>tau.sp (sigma)</th>
<th>Deviance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global</td>
<td>2.99 (2.98, 3.13)</td>
<td>0.85 (1.18)</td>
<td>NA</td>
<td>698</td>
</tr>
<tr>
<td>Spatial RE</td>
<td>2.97 (2.66, 3.28)</td>
<td>1.04 (0.96)</td>
<td>4.98 (0.20)</td>
<td>654</td>
</tr>
</tbody>
</table>
Hierarchical Bayes

Hierarchical model allows predictions about the unobserved

Out-of-sample predictions integrate over the random effects variance in known sites/years – will include more uncertainty than in-sample estimates.

Hierarchical

Shared
Hierarchical Bayes

Ecology is complex. You cannot measure everything.

Hierarchical models can help identify important drivers and account for non-independence in sampling scheme.

Hierarchical model allows for borrowed strength from data-rich to data-poor:

- predictions about the unobserved
- integrate over the random effects variance = more real reflection of uncertainty in out-of-sample predictions
Ecological Forecasting is about characterizing uncertainty
Ecological Forecasting is about characterizing uncertainty

What will mosquito abundance be next month?

When and where should city invest in control?