Diversity Example: Wadden Sea

Bob O'Hara

An Example

Observation Model

The data are counts of each species, and we model both presence/absence and counts. We have multiple replicate observations per site, so we assume the counts, n_{kj} on replicate k for species j, follow a zero inflated Poisson lognormal distribution with mean λ_{ij} (i.e. constant acriss replicates within a site). If the detection probability is constant within a species (but differing between species) this will change the $\beta_j^{(\lambda)}$ term. We have a Gaussian overdispersion term for excess variation within a replicate that represents sampling variation.

Downloading and formatting the data

The data are abundances (counts) on each location on a transect. There are 3 islands, each with 3 transects, each with 3 stations. Each station was sampled in 2 seasons, and has 3 replicates.

The steps are: 1. import the data

2. Format the data for JAGS.

3. Write the model in JAGS code

4. Run the JAGS code

Diversity Output

All of our diversity metrics are based on decomposing the distributions into:

- species richness
- composition

And looking at the means and variances of the parameters associated with these.

Alpha Diversity

Species Richness

We calculate species richness on site i, S_i, as $\sum_j z_{ij}$

In Figure 1 we plot the estimated species richness for each site for the data, with colours for islands and open/closed circles for summer/spring respectively.

Unevenness

Our estimate of (site specific) unevenness is

$$\mathrm{Var}_j(\eta_{ij}|i) = \sigma_\beta^2 + \sigma_\gamma^2 x_i + \sigma_{\delta,i}^2$$

We plot it for our sites in Figure 2. The unevennesses look very high, apparently because of rare species. We can look at the distribution of posterior means for a couple of sites (grey is the total distribution of present species, red is for species that are observed to be present) in Figure 3.

Most species that are modelled to be present are really rare.

Beta Diversity

Species Identity

Nestedness

We can calculate nestedness, and also look at the proportion of variance due to the gradient (Figure 4).

Most (59%) of the variation in nestedness is due to temperature. The effect has a posterior mean of -0.29 and a 95% credible interval of -0.51 to -0.05. The distribution of the proporyion looks horrible here.
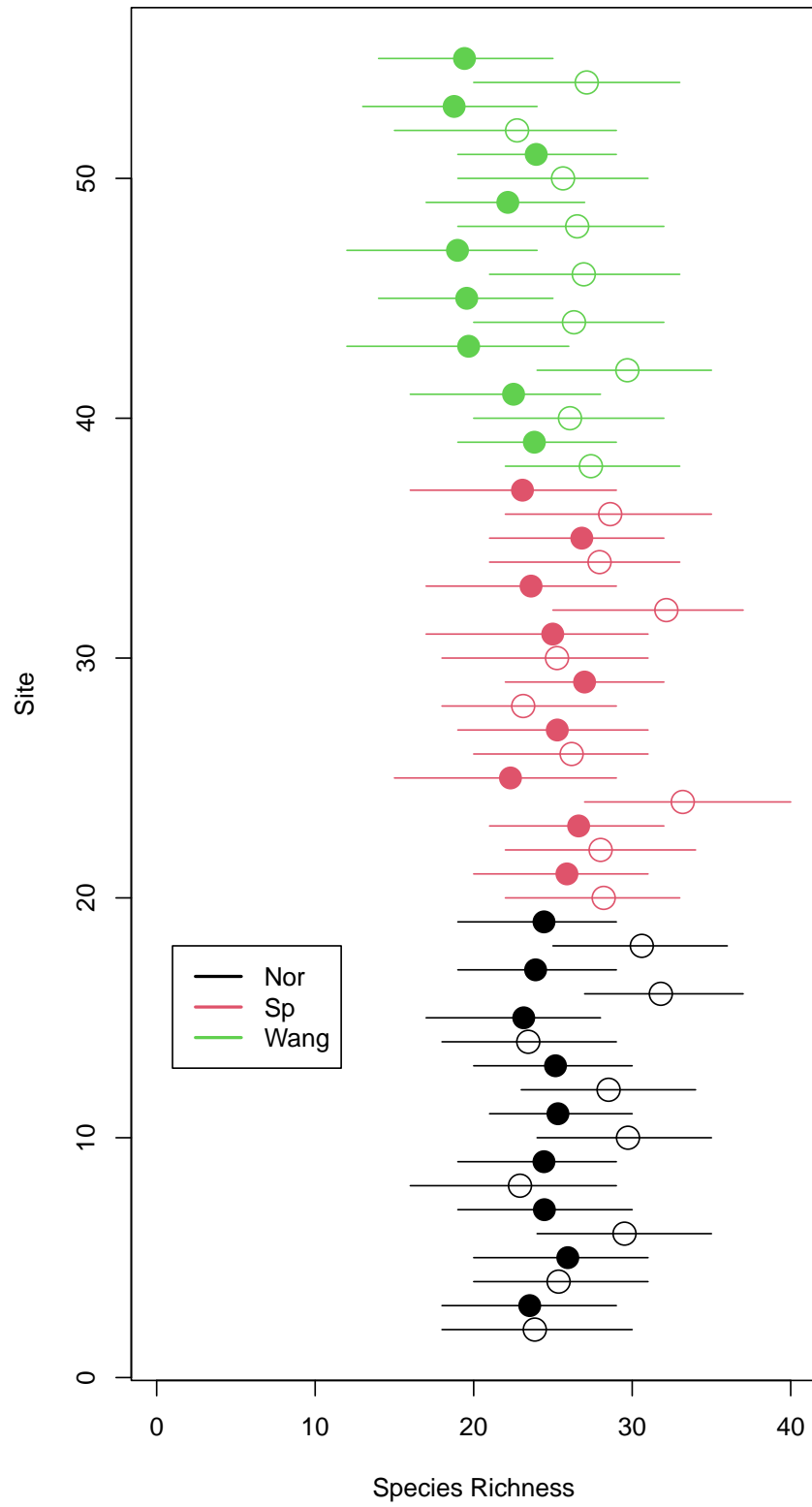
Figure 1: Posterior estimates of species richness. Filled circles: spring, open circles: summer.
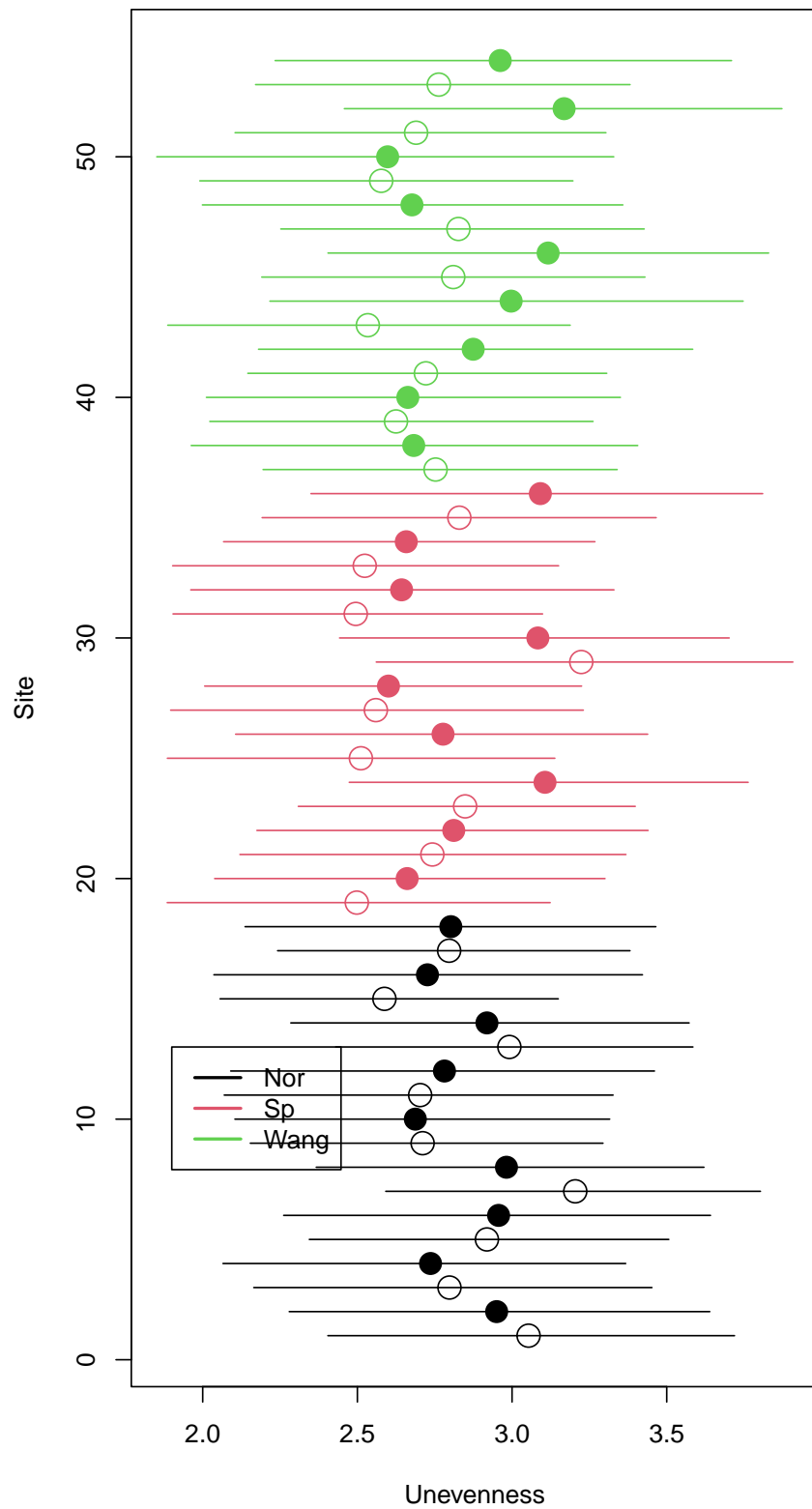
Figure 2: Posterior estimates of unevenness for all sites. Filled circles: spring, open circles: summer.
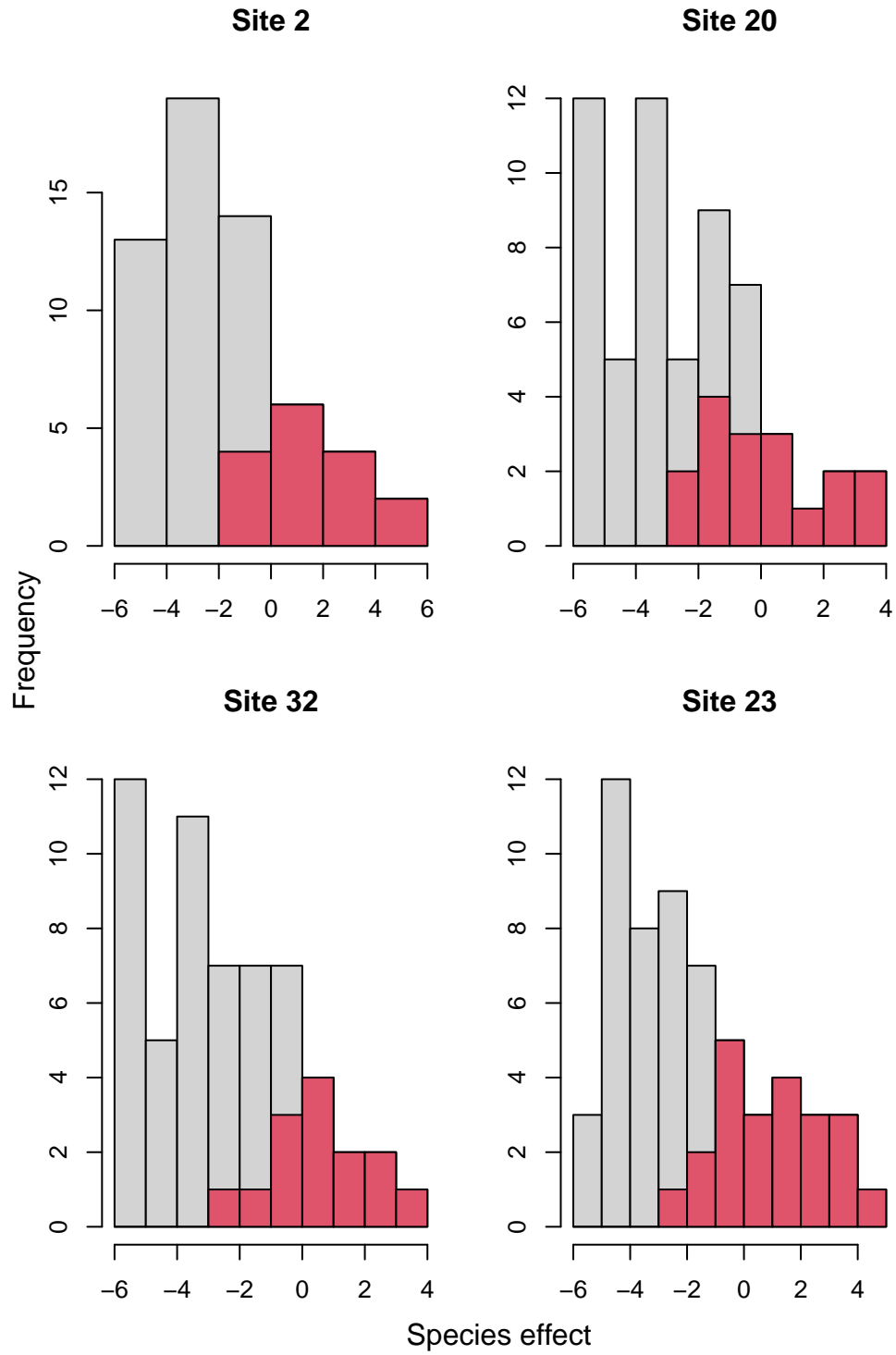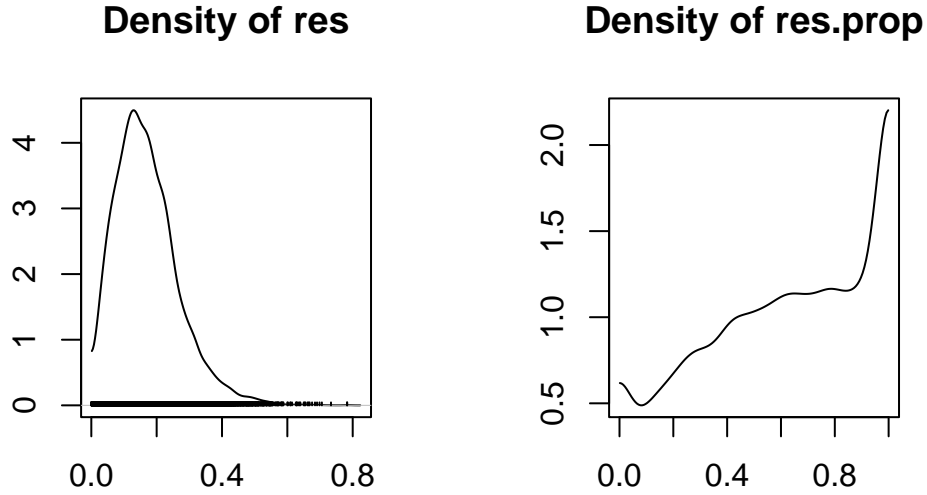
Figure 3: Estimates of abundance distributions for four sites: estimates for species which were observed are in red, and unobserved species (i.e. those with abundance 0 on these sites, but where z=1) in grey.

Figure 4: Nestedness. Variance on left, proportion due to the environment on the right. The horror that is the density on the right is real, and not due to the MCMC behaving poorly.

Replacement

Replacement is $E_j\left(\mathrm{Var}_i(\eta_{ij}^{(z)})\right) = \sigma_\alpha^2 + \left(\sigma_{\gamma^{(z)}}^2 + (\bar{\gamma}^{(z)})^2\right)\mathrm{Var}_i(x_i) + \pi^2/3$, and the density is plotted in Figure 5.

Replacement is a number. Of it, about 15% of the variation is due to temperature, so most of this variation cannot be explained by the gradient.

Composition

Our compositional beta diversity statistic is $\sigma_\alpha^2 + \left(\sigma_{\gamma^{(\lambda)}}^2 + (\bar{\gamma}^{(\lambda)})^2\right)\mathrm{Var}_i(x_i) + \sigma_\delta^2$

We can see that we have beta diversity. For GDM fans, only about 7% of the variation is driven by the temperature gradient.

Gamma Diversity

Total Variation between sites & species. Does anyone care by this point? In case you do, it's in Figure 7.

Density of res

Density of res.prop

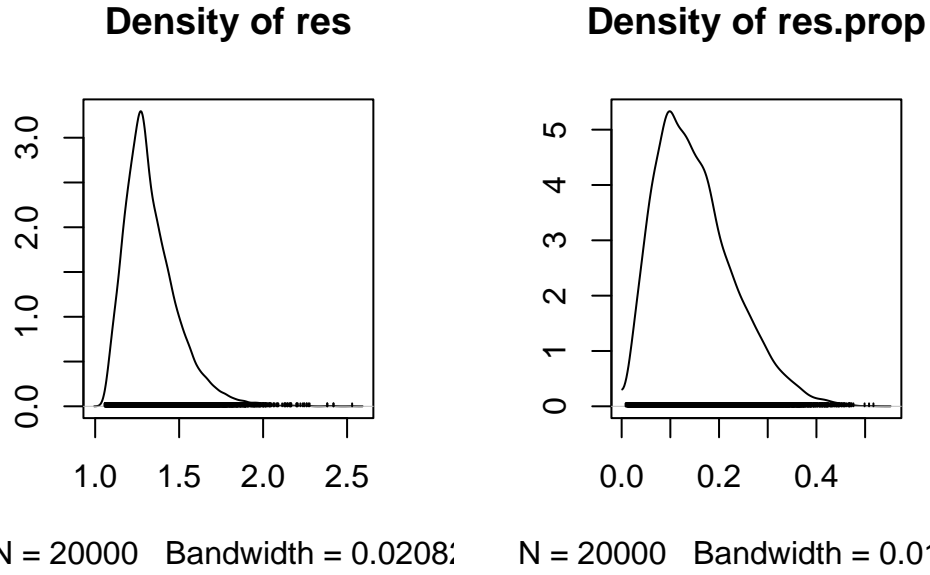N = 20000   Bandwidth = 0.0208   N = 20000   Bandwidth = 0.0117

Figure 5: Densities of replacement: variance on left, proportion due to covariates on the right.

Density of res

Density of res.prop

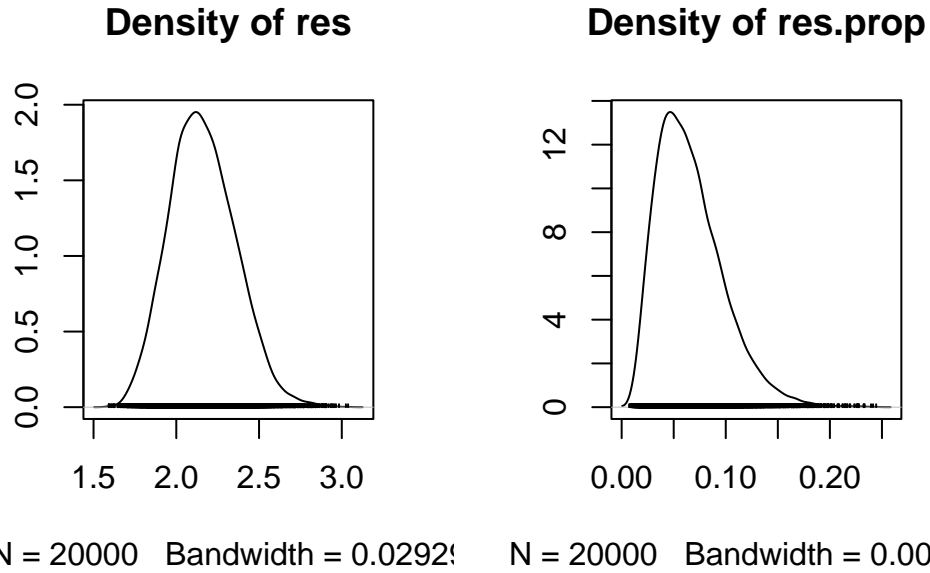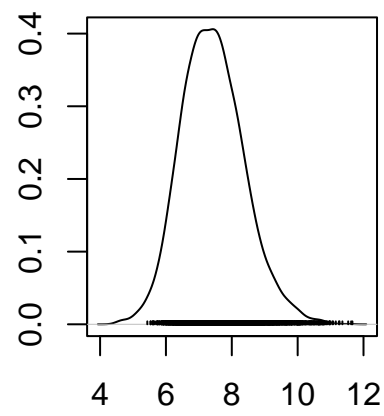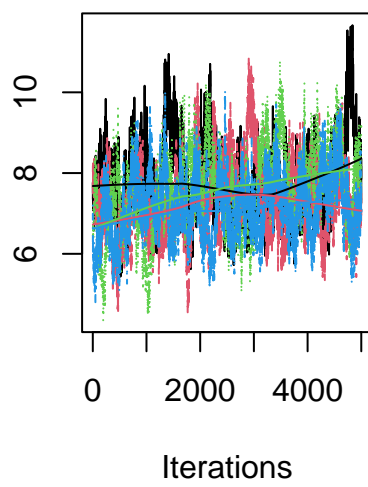N = 20000   Bandwidth = 0.0292   N = 20000   Bandwidth = 0.00462

Figure 6: Densities of compositional diversity: variance on left, proportion due to the environment on right. The proportion due to the environment is equivalent to a statistic for GDM.

7

Figure 7: Summary of posterior distribution of Gamma diversity (history of MCMC draws on left, density on right).