Eden: soil carbon and nitrogen

## Overall modelling approach

For all response variables we fit GAMs with:

* A smooth term for number of fires, fitted with a thin-plate regression spline that allows shrinkage to a linear function as dictated by the data. This is preferred to the previous cubic spline with fixed degrees of freedom because it gives trend lines that are easier to interpret and more stable where the data are sparse.
* An interaction between number of fires and harvesting (binary variable).
* Open vs tree micro-site.
* Random effect term for site.
* Weights assigned to observations based on proportion of samples for each number of fires.

Validation outputs for models (graphs of residuals etc) have been kept separate to keep this document uncluttered.

## Model for carbon

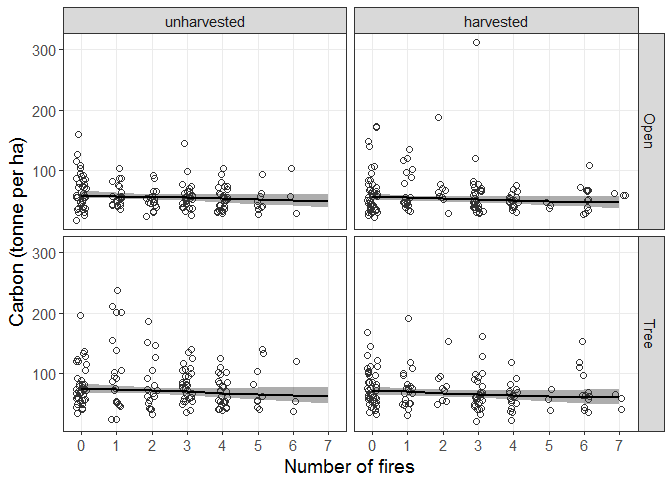
From Meaghan:

Carbon (t/ha) = Depth (cm) x Bulk density (g/cm3) x Carbon content (%)

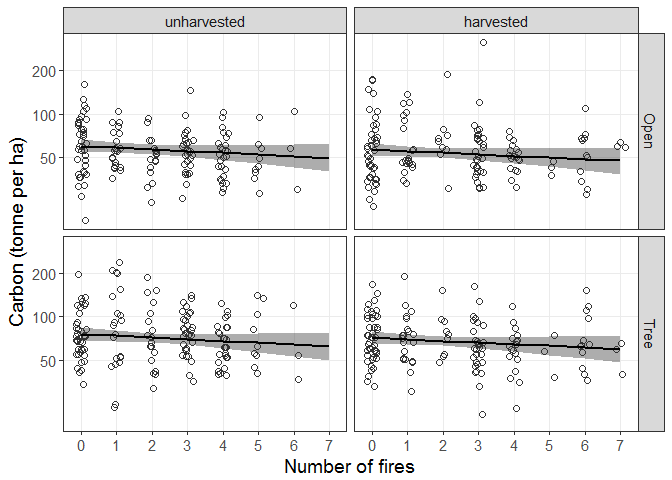
dat.model <- DAT  
dat.model$flag <- 1 # turns on site random effect  
  
mcarbon <- gam(log(carbontha) ~ s(nfires, bs = "tp", k=8, fx=FALSE, by = harvest) +  
 harvest + treeopen +  
 s(plot, bs = "re", by = flag),  
   
 data = dat.model,  
 weights = weights,  
 family = gaussian())  
  
summary(mcarbon)

##   
## Family: gaussian   
## Link function: identity   
##   
## Formula:  
## log(carbontha) ~ s(nfires, bs = "tp", k = 8, fx = FALSE, by = harvest) +   
## harvest + treeopen + s(plot, bs = "re", by = flag)  
##   
## Parametric coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.03775 0.03776 106.943 < 2e-16 \*\*\*  
## harvestharvested -0.05353 0.04945 -1.082 0.28   
## treeopenTree 0.22912 0.03094 7.405 5.3e-13 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df F p-value   
## s(nfires):harvestunharvested 1.00 1 1.861 0.173   
## s(nfires):harvestharvested 1.00 1 1.780 0.183   
## s(plot):flag 49.25 92 1.516 7.5e-14 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = 0.264 Deviance explained = 33.2%  
## GCV = 0.22836 Scale est. = 0.20685 n = 576

Graph of model predictions



Same graph with log-scaling for the y-axis.

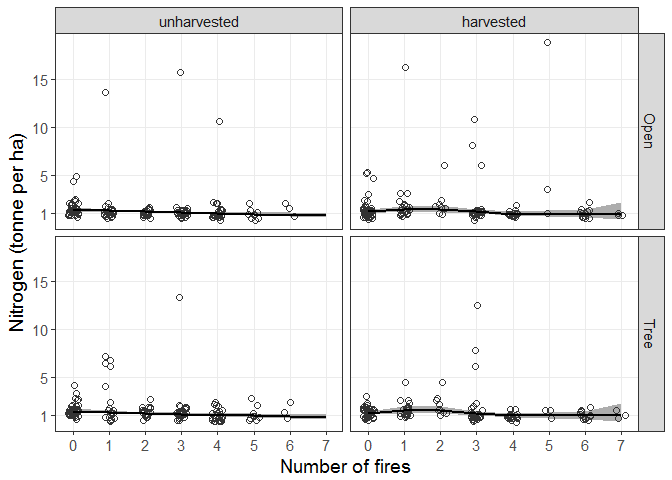


## Nitrogen tonnes per hectare

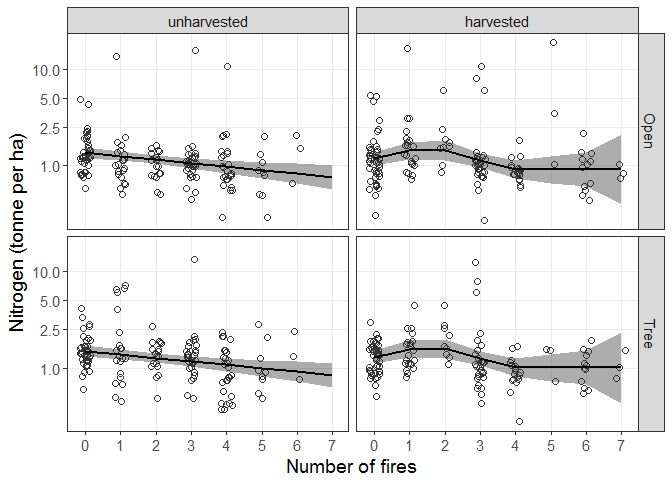
mnitrogen <- gam(log(nitrogentha) ~ s(nfires, k=8, by = harvest) +  
 harvest + treeopen +  
 s(plot, bs = "re", by = flag),  
   
 data = dat.model,  
 weights = weights,  
 family = gaussian())  
  
summary(mnitrogen)

##   
## Family: gaussian   
## Link function: identity   
##   
## Formula:  
## log(nitrogentha) ~ s(nfires, k = 8, by = harvest) + harvest +   
## treeopen + s(plot, bs = "re", by = flag)  
##   
## Parametric coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.12636 0.04985 2.535 0.0115 \*  
## harvestharvested 0.01217 0.06538 0.186 0.8524   
## treeopenTree 0.10592 0.04191 2.527 0.0118 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df F p-value   
## s(nfires):harvestunharvested 1.000 1.000 9.665 0.00198 \*\*   
## s(nfires):harvestharvested 3.434 3.853 2.530 0.04208 \*   
## s(plot):flag 45.754 92.000 1.148 5.92e-09 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = 0.204 Deviance explained = 27.7%  
## GCV = 0.41811 Scale est. = 0.3795 n = 576

Graph of model predictions



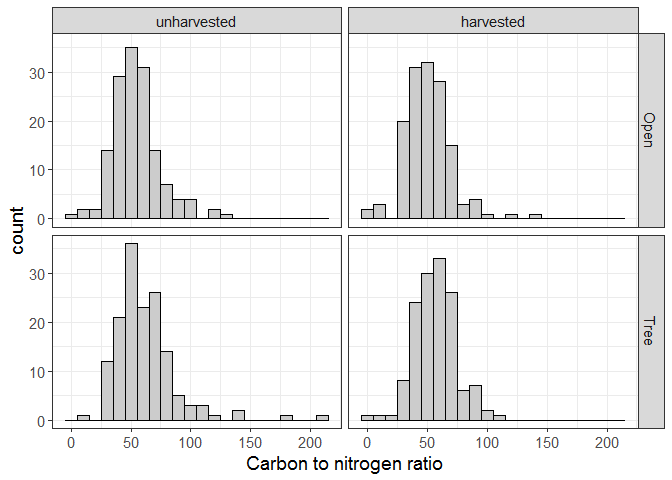
Same graph with log-scaling for the y-axis.



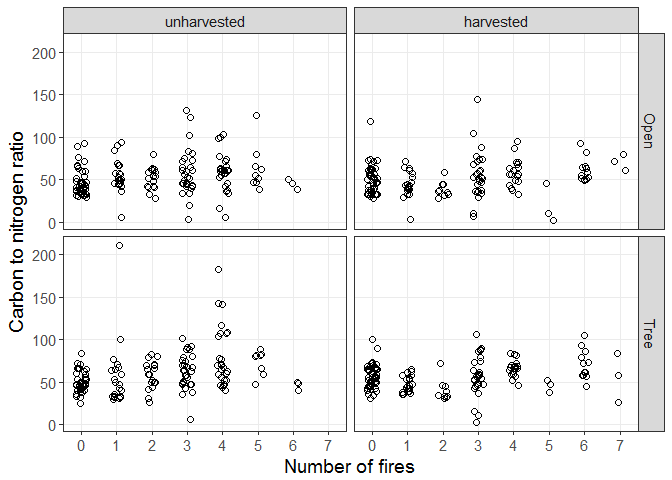
## Carbon to nitrogen ratio

This has been calculated as the ratio of the carbon t/ha to nitrogen t/ha values used for the above models.

Distribution of values within fire x harvest classes:



Ratio values versus number of fires:

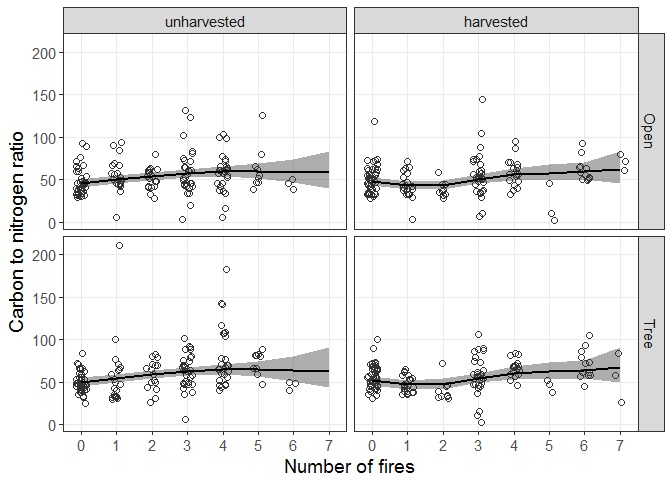


This model is fitted using a scaled t-distribution (with the argument family = scat) as a heavy-tailed alternative to a Normal distribution. This was chosen after finding that the residuals from an initial Gaussian model were very poorly distributed in both tails.

mratio <- gam(log(cnratio) ~ s(nfires, k=8, by = harvest) +  
 harvest + treeopen +  
 s(plot, bs = "re", by = flag),  
   
 data = dat.model,  
 weights = weights,  
 family = scat())  
  
summary(mratio)

##   
## Family: Scaled t(2.129,0.162)   
## Link function: identity   
##   
## Formula:  
## log(cnratio) ~ s(nfires, k = 8, by = harvest) + harvest + treeopen +   
## s(plot, bs = "re", by = flag)  
##   
## Parametric coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 3.96636 0.03101 127.903 < 2e-16 \*\*\*  
## harvestharvested -0.06812 0.04250 -1.603 0.109   
## treeopenTree 0.08453 0.01732 4.880 1.06e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df Chi.sq p-value   
## s(nfires):harvestunharvested 1.838 1.916 14.87 0.00188 \*\*   
## s(nfires):harvestharvested 3.241 3.368 13.94 0.00803 \*\*   
## s(plot):flag 73.776 92.000 428.06 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = 0.0309 Deviance explained = 35.1%  
## -REML = 201.79 Scale est. = 1 n = 576

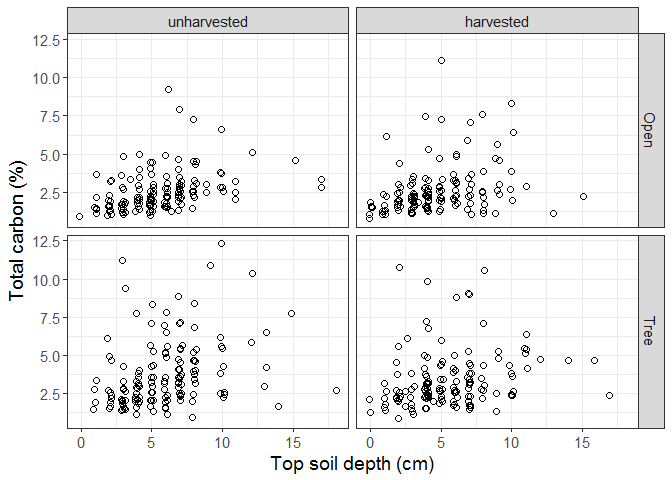
Graph of model predictions



## Extra graphs

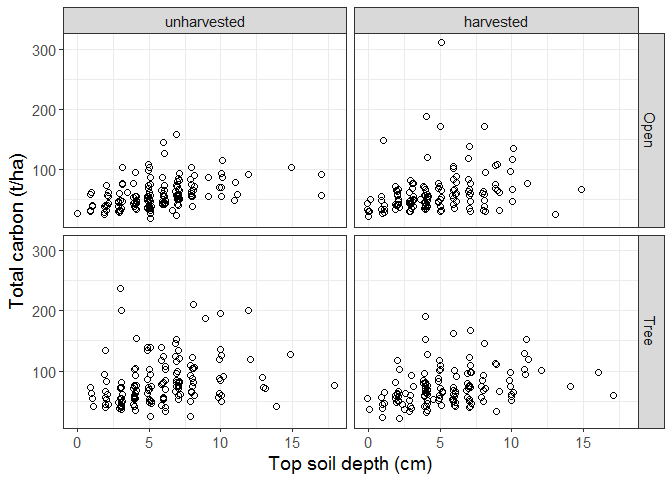
### Carbon (percentage by weight) vs top soil depth

Points have been horizontally jittered for clarity.



### Carbon (tonnes per ha) vs top soil depth

Points have been horizontally jittered for clarity.

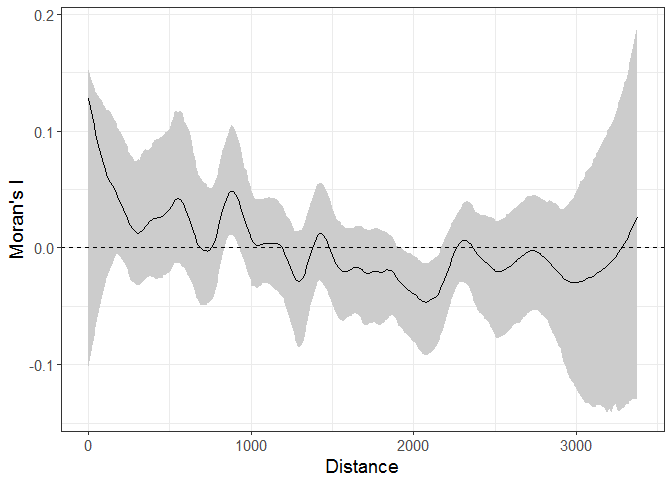


### Graphical check for spatial correlation

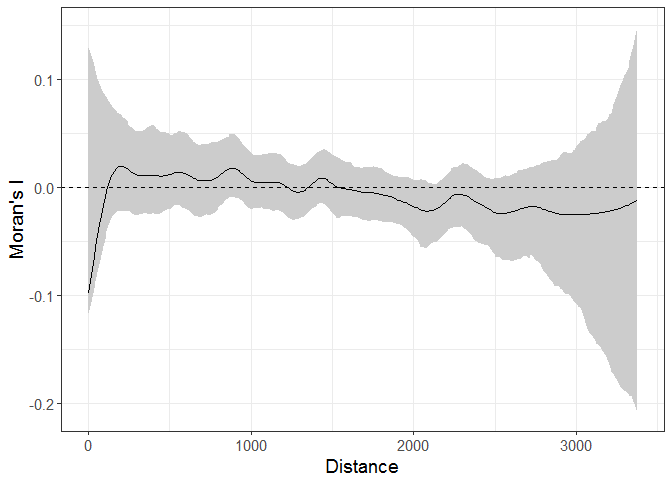
The following graphs (correlograms) show how the correlation between observations varies with distance between them. Correlation is measured as Moran's *I* with point-wise bounds derived from bootstrap resampling. We use the spline.correlog function from the ncf package to fit a smooth trend line and avoid having to define distance classes for calculation.

For each response variable we examine spatial correlation in values of the variable itself, and in the residuals of the model fitted for that variable.

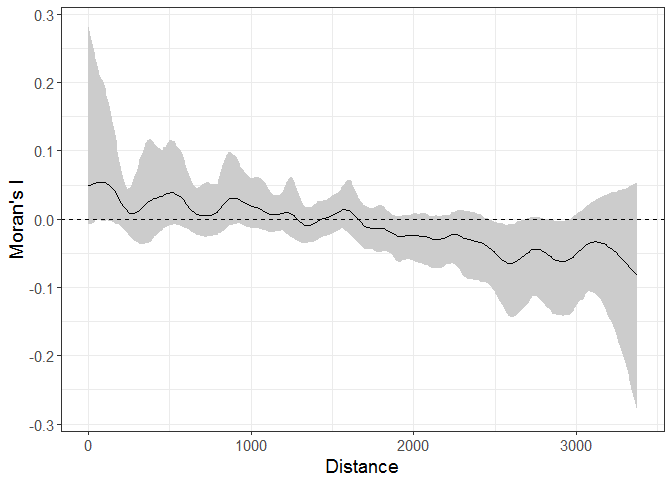
#### Carbon t/ha values



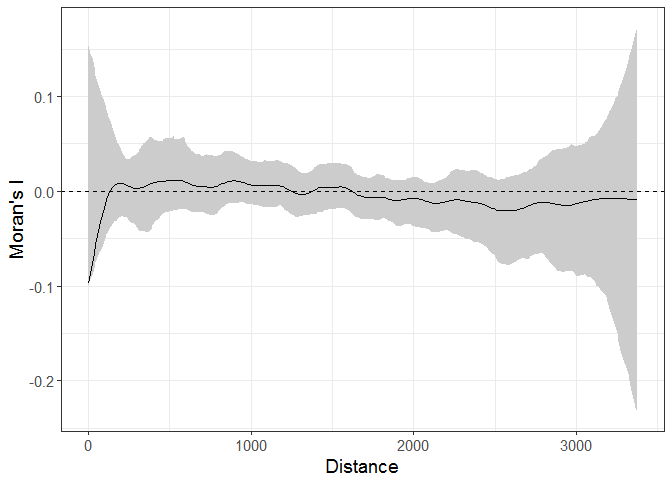
#### Carbon model residuals



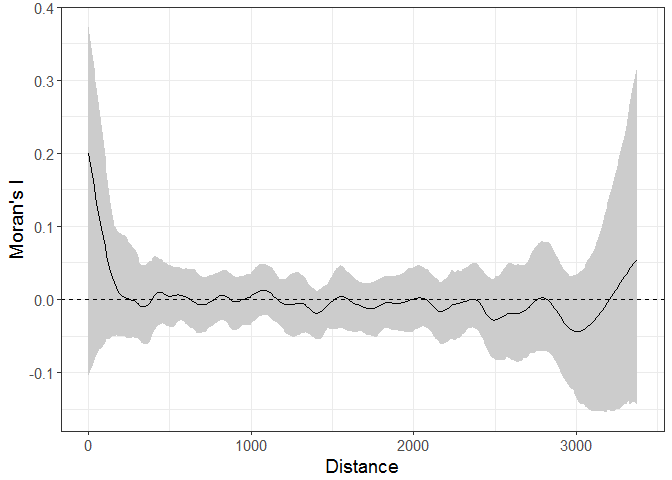
#### Nitrogen t/ha values



#### Nitrogen model residuals



#### Carbon / nitrogen ratio



#### Carbon / nitrogen ratio model residuals

