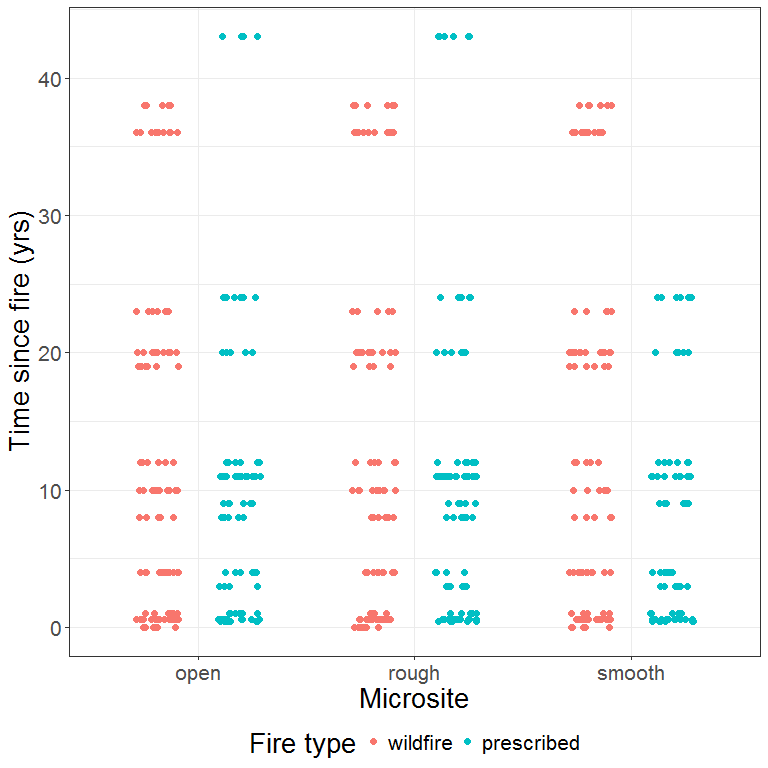
# Analysis of chrono-sequence data with generalized additive models

## Coverage of time since fire

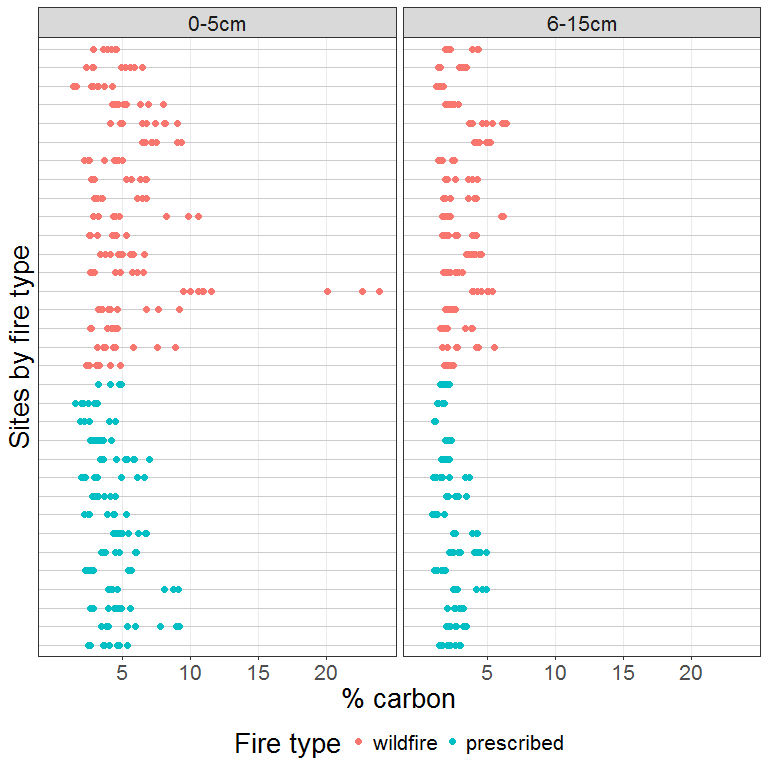
Distribution of time since fire values over design factors:



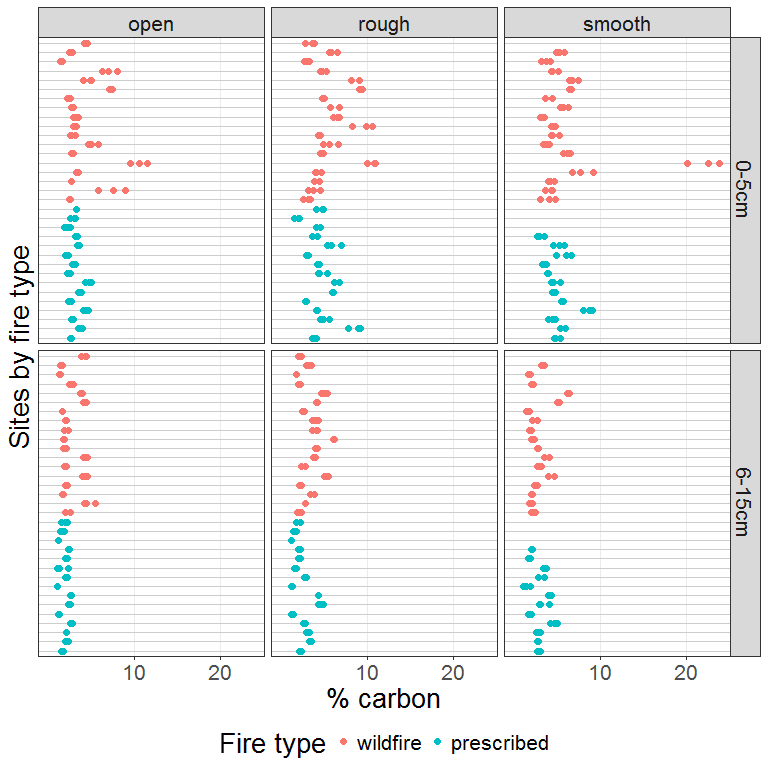
## Total carbon

### Data exploration

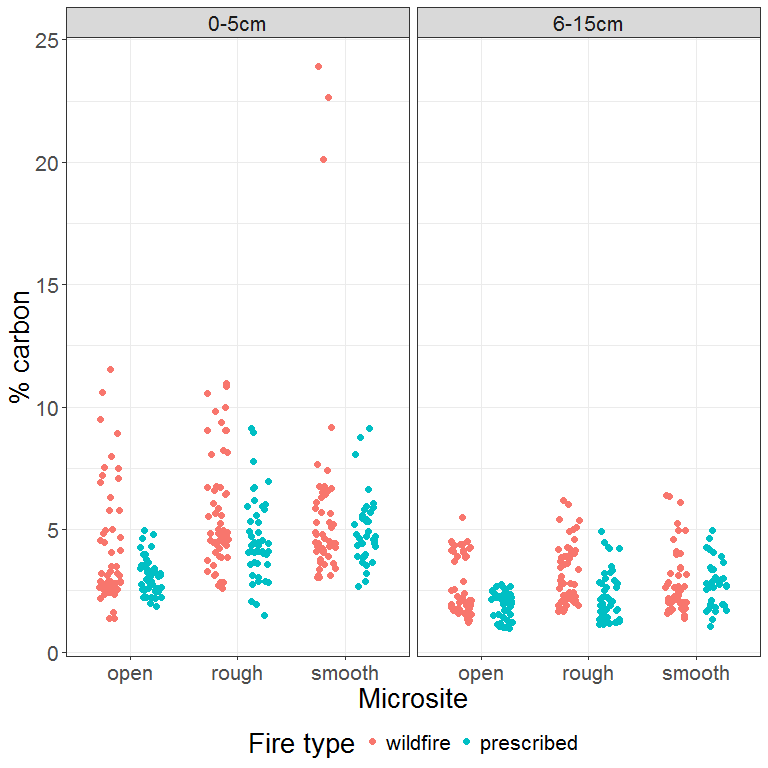
The following dot plot shows percent carbon values for all samples. Sites are ordered by fire type, then site ID value (integer: 1 - 33).



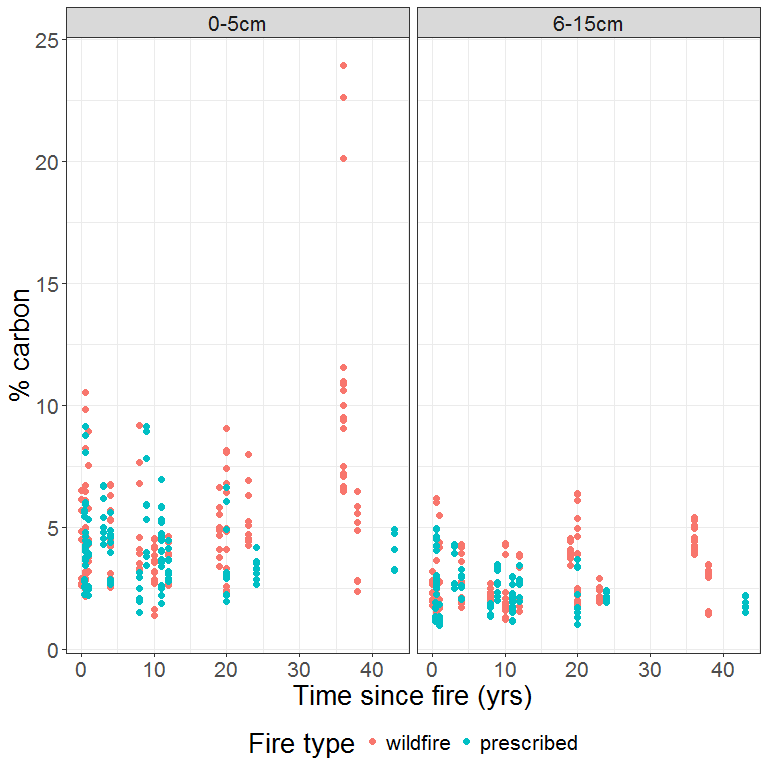
Same data, split by micro-site. Note that 'smooth' micro-site samples are missing for three prescribed fire sites and one wildfire site.

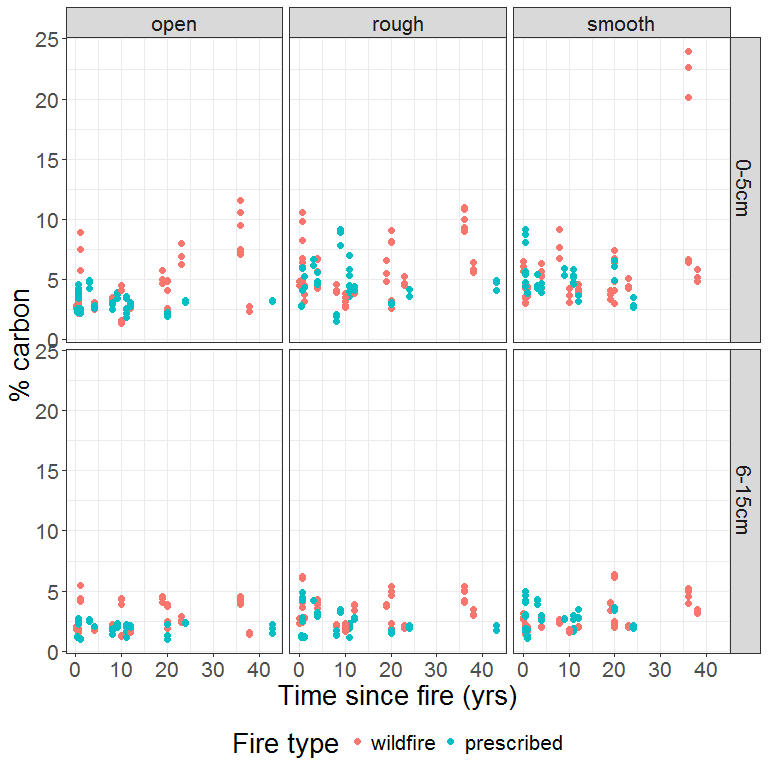


Distribution of values with respect to fire type and microsite.



Pattern of percent carbon values relative to time since fire.





### Initial model

We begin with a simple additive model, ignoring the nesting of samples within sites. Micro-site is fitted as a random effect so that we can zero it out when plotting model predictions if desired.

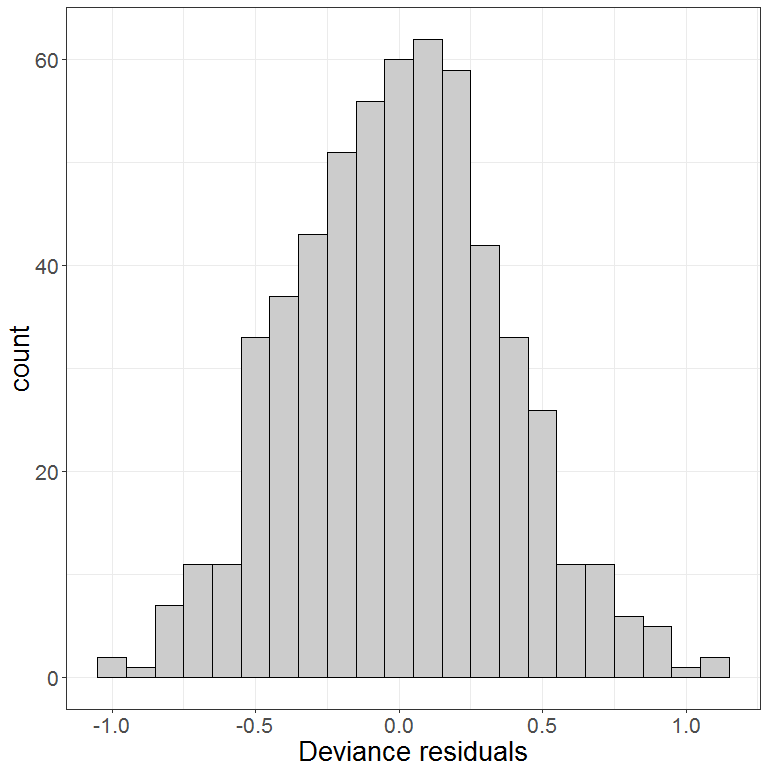
This model includes a smoothed term for time since fire to allow for non-linear effects, with an interaction between time, fire type and depth.

**Note:** We place a conservative upper limit on the number of degrees of freedom for the smooth term for time since fire. This is done because we noticed that in trial models using the default limit on degrees of freedom, the fitted smooth function for time since fire had a pronounced concave-down shape in the interval where data are absent (25-35 years since fire). This was clearly an artefact of the sparse data for longer times since fire. Adding the constraint gives a more parsimonious and credible trend.

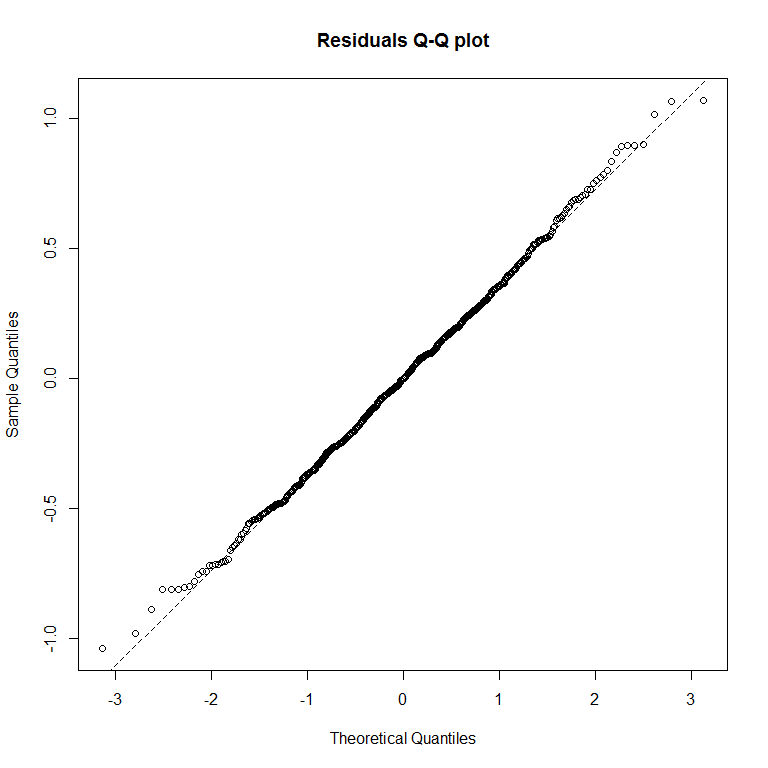
The R code to fit this model is:

dat.model <- DAT  
dat.model$flag.ms <- 1  
  
mtotalC <- gam(log(percentC) ~ s(tsf.years, by = dft, k=4) +   
 fireType + depth +   
 s(microsite, bs = "re", by = flag.ms),  
 data = dat.model,  
 method = "REML")

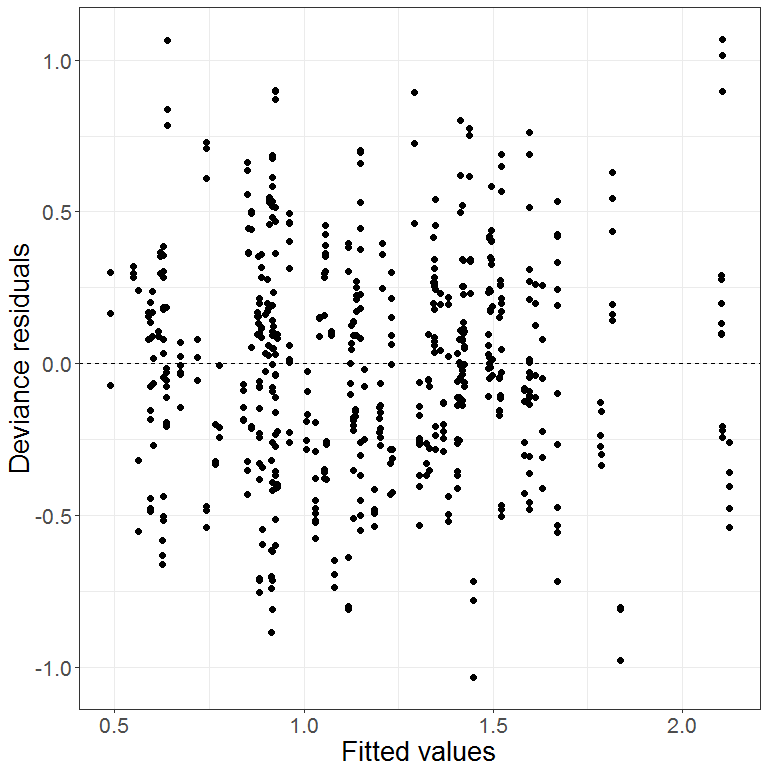
Examining model residuals...



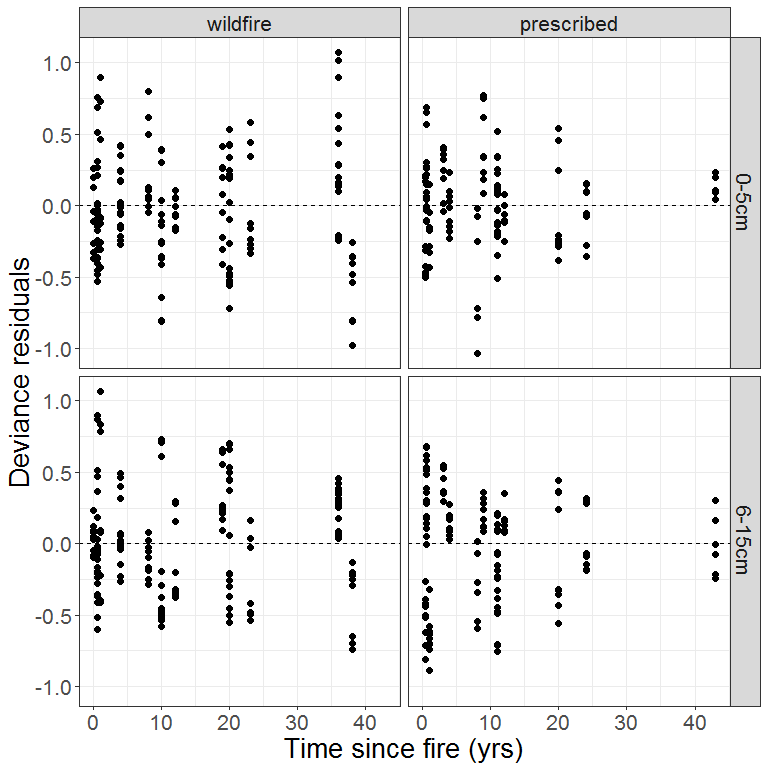
Checking residuals for normality:

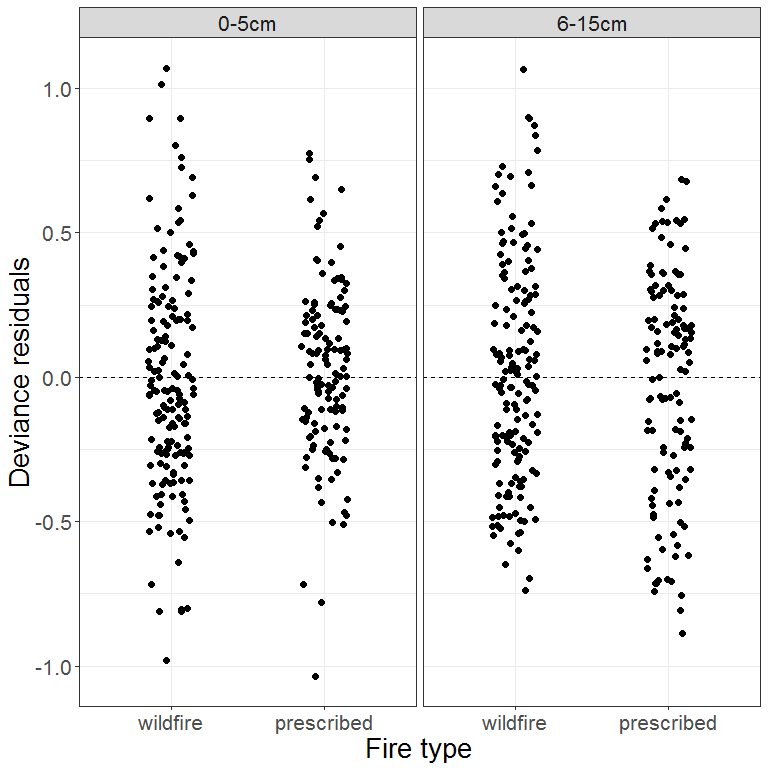


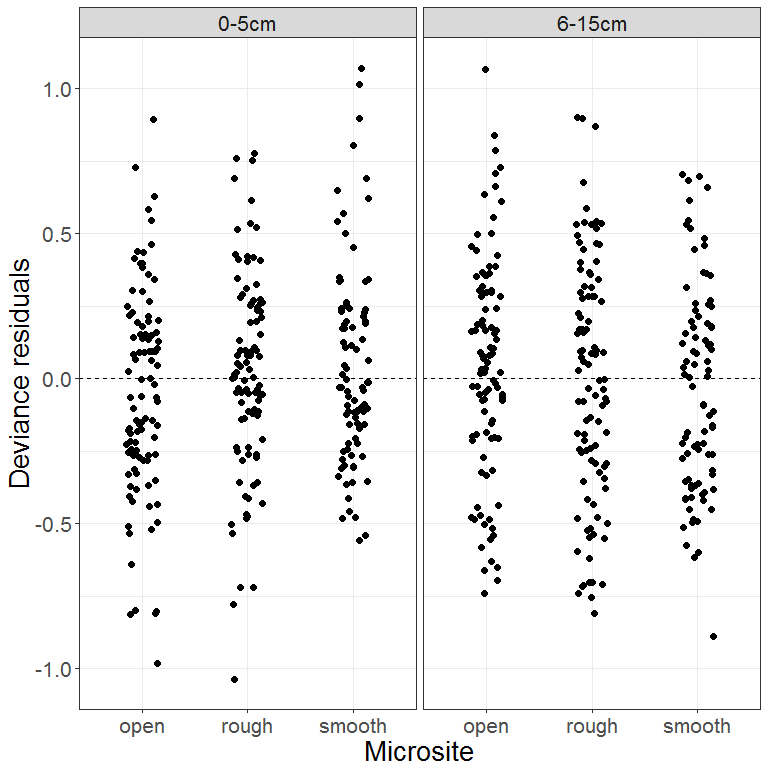
Graph of residuals against fitted values to check for homogeneity.



Residuals against co-variates.

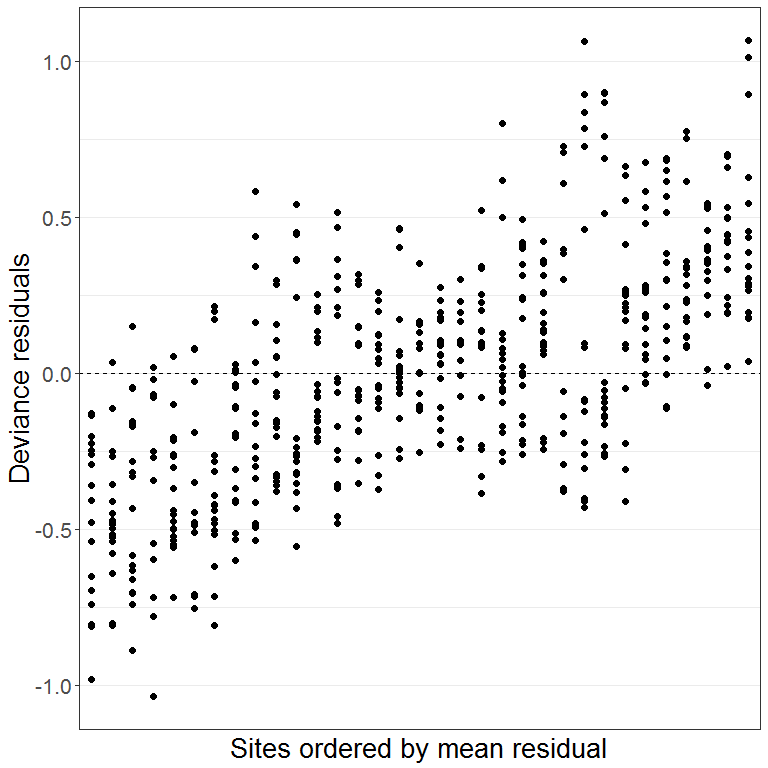






Graph of residuals against sites to check for any systematic patterns and assess our decision to ignore any correlation of responses within sites in the initial model.

To make any pattern easier to see, sites are arranged on the X-axis in ascending order of mean residual value within site.



Residuals are not independent across sites. We confirm this by regressing residuals on site:

## Analysis of Variance Table  
##   
## Response: resid  
## Df Sum Sq Mean Sq F value Pr(>F)   
## fSiteId 32 36.042 1.12632 15.112 < 2.2e-16 \*\*\*  
## Residuals 537 40.023 0.07453   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

### Second model incorporating site random effect

Here we add a random-effect smoother to the model to account for site differences. Because the number of sites is small, we take [this approach described by Simon Wood](http://r.789695.n4.nabble.com/mgcv-gamm-predict-to-reflect-random-s-effects-tp3622738p3622865.html) (author of the mgcv package), using the standard gam function. A binary flag associated with the random effect smoother is set to 1 to include the random effect for model fitting, and set to 0 to exclude it when making predictions.

dat.model <- DAT  
dat.model$flag.site <- 1  
dat.model$flag.ms <- 1  
  
mtotalC.site <- gam(log(percentC) ~ s(tsf.years, by = dft, k=4) +   
 fireType + depth +   
 s(microsite, bs = "re", by = flag.ms) +   
 s(fSiteId, bs = "re", by = flag.site),  
 data = dat.model,  
 method = "REML")

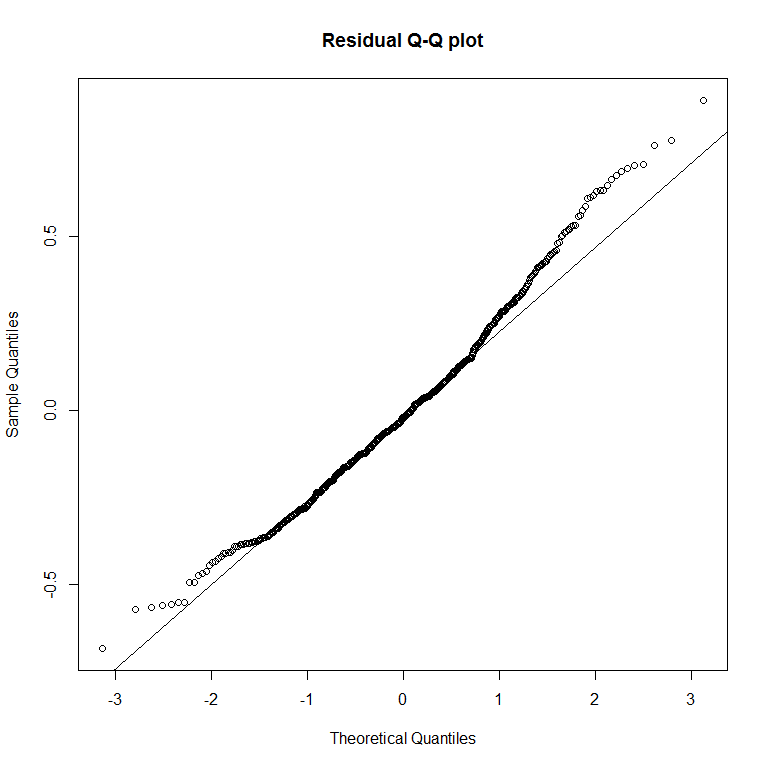
Once again we graph the residuals against sites ordered in the same manner as before.



Residuals now appear independent across sites. We confirm this by repeating the regression:

## Analysis of Variance Table  
##   
## Response: resid  
## Df Sum Sq Mean Sq F value Pr(>F)  
## fSiteId 32 0.127 0.003975 0.0536 1  
## Residuals 537 39.790 0.074097

Checking for normality of residuals:



Not great but we’ll live with it.

Comparing the two models with AIC:

## df AIC  
## mtotalC 13.26487 496.1128  
## mtotalC.site 41.09447 184.2442

The second model, incorporating site random effects, is clearly preferred.

Summary of the fitted model:

##   
## Family: gaussian   
## Link function: identity   
##   
## Formula:  
## log(percentC) ~ s(tsf.years, by = dft, k = 4) + fireType + depth +   
## s(microsite, bs = "re", by = flag.ms) + s(fSiteId, bs = "re",   
## by = flag.site)  
##   
## Parametric coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.52807 0.11605 13.168 <2e-16 \*\*\*  
## fireTypeprescribed -0.19954 0.09694 -2.059 0.04 \*   
## depth6-15cm -0.56397 0.02336 -24.141 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df F p-value   
## s(tsf.years):dftwildfire.0-5cm 2.201 2.524 6.799 0.000735 \*\*\*  
## s(tsf.years):dftprescribed.0-5cm 1.438 1.717 0.335 0.572284   
## s(tsf.years):dftwildfire.6-15cm 1.000 1.000 5.090 0.024454 \*   
## s(tsf.years):dftprescribed.6-15cm 1.000 1.000 0.183 0.668768   
## s(microsite):flag.ms 1.971 2.000 75.080 < 2e-16 \*\*\*  
## s(fSiteId):flag.site 27.229 29.000 16.819 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = 0.722 Deviance explained = 74%  
## -REML = 134.7 Scale est. = 0.075009 n = 570

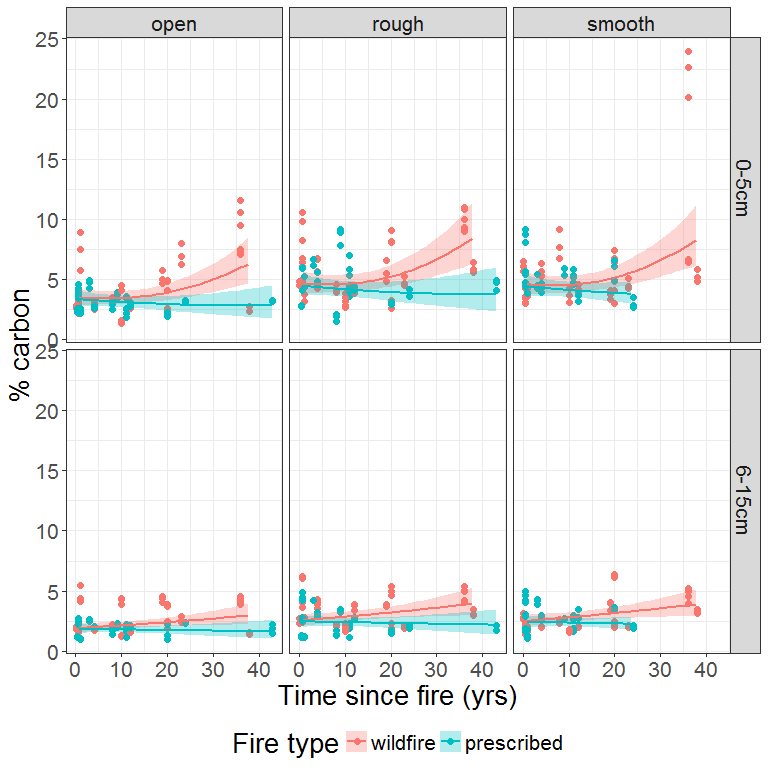
### Model summary with micro-site as fixed term

Changing micro-site from random to fixed effect gives coefficient estimates for the 'rough' and 'smooth' categories in the model summary. The coefficients for other terms (excepting the intercept) will remain unchanged apart from possible minor jiggle.

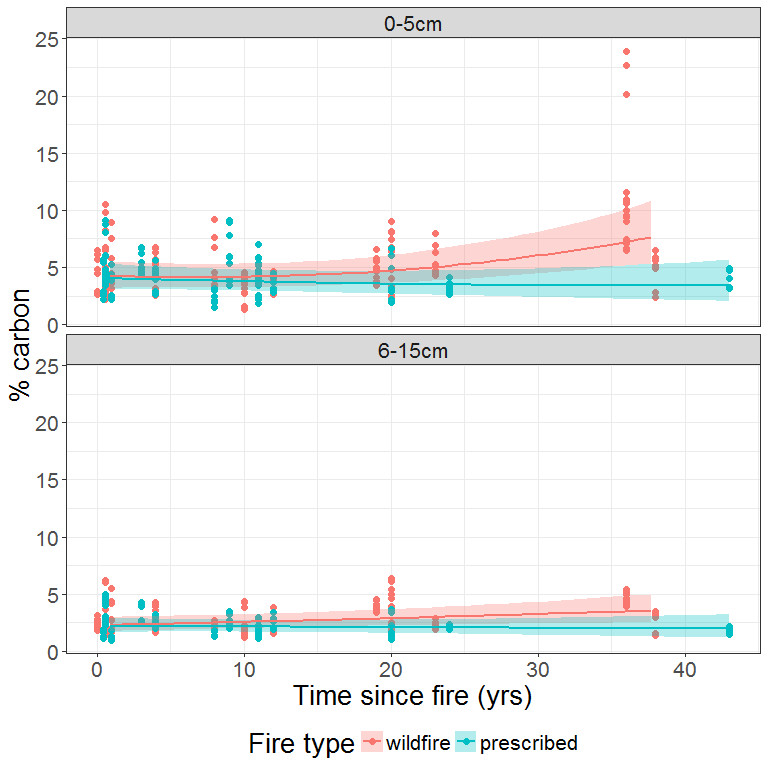
##   
## Family: gaussian   
## Link function: identity   
##   
## Formula:  
## log(percentC) ~ s(tsf.years, by = dft, k = 4) + fireType + depth +   
## microsite + s(fSiteId, bs = "re", by = flag.site)  
##   
## Parametric coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.33660 0.06820 19.600 <2e-16 \*\*\*  
## fireTypeprescribed -0.19943 0.09692 -2.058 0.0401 \*   
## depth6-15cm -0.56397 0.02336 -24.140 <2e-16 \*\*\*  
## micrositerough 0.29559 0.02753 10.739 <2e-16 \*\*\*  
## micrositesmooth 0.27894 0.02889 9.657 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df F p-value   
## s(tsf.years):dftwildfire.0-5cm 2.201 2.524 6.800 0.000734 \*\*\*  
## s(tsf.years):dftprescribed.0-5cm 1.439 1.717 0.335 0.572873   
## s(tsf.years):dftwildfire.6-15cm 1.000 1.000 5.091 0.024439 \*   
## s(tsf.years):dftprescribed.6-15cm 1.000 1.000 0.182 0.669933   
## s(fSiteId):flag.site 27.228 29.000 15.997 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = 0.722 Deviance explained = 74%  
## -REML = 134.9 Scale est. = 0.075011 n = 570

### Model predictions

Predictions taking into account micro-site:

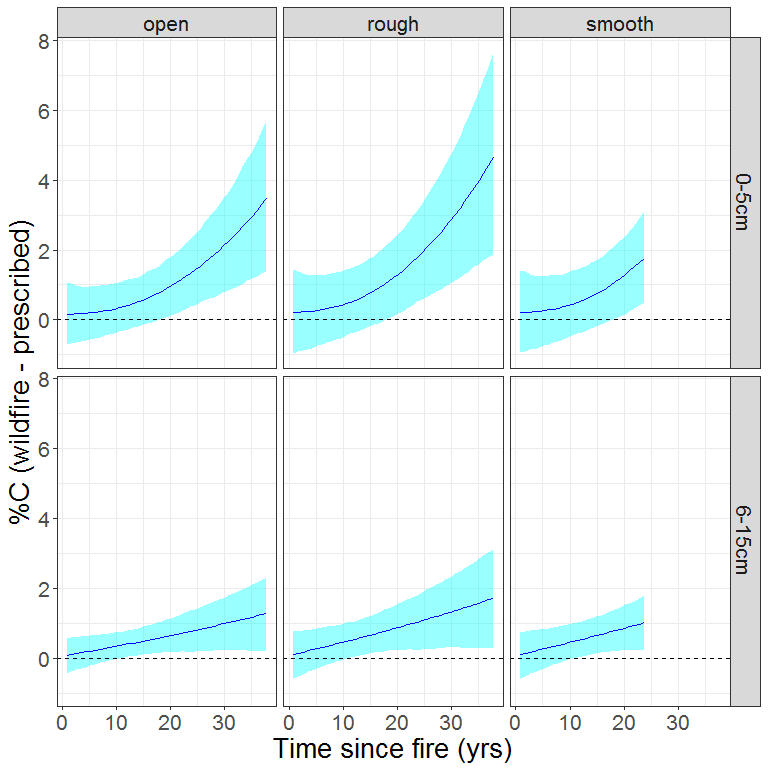


Predictions excluding the effect of micro-site:



### Predicted difference between fire types

Here we use posterior simulation to generate a set of 1000 credible mean trends from the fitted GAM for total carbon and, for each trend, calculate the difference in carbon between the two fire types.

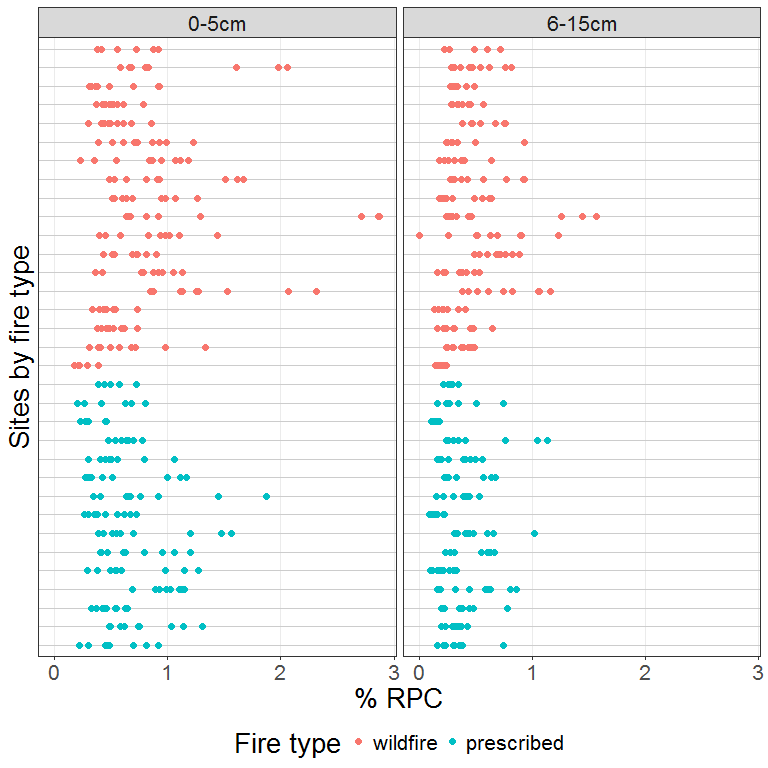


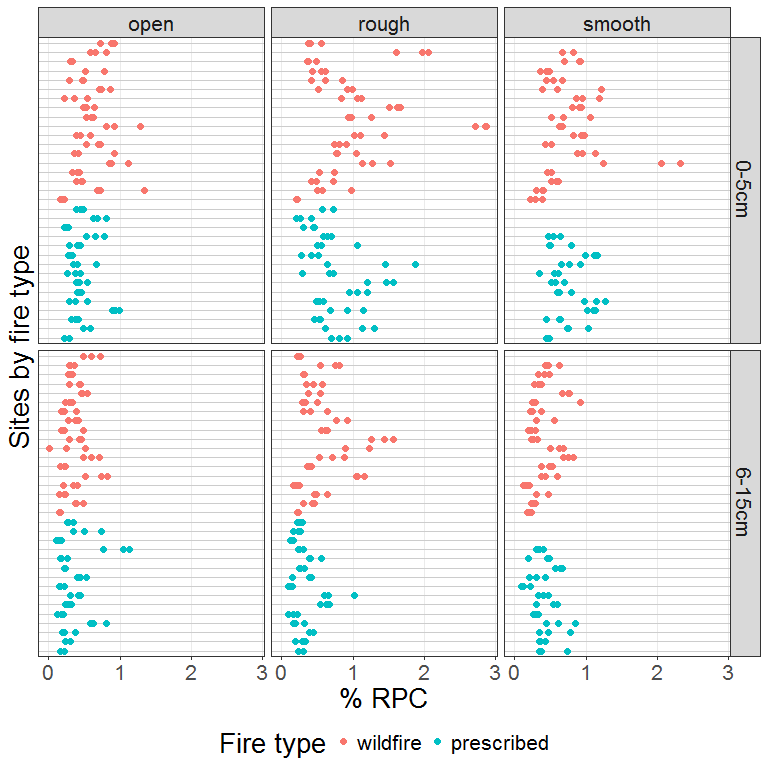
## Recalcitrant pyrogenic carbon

For RPC, we are interested in the absolute value within samples as well as the fraction of total carbon in the form of RPC. Here we begin with a model of the absolute values, controlling for the amount of total carbon in the sample by including that as a co-variate.

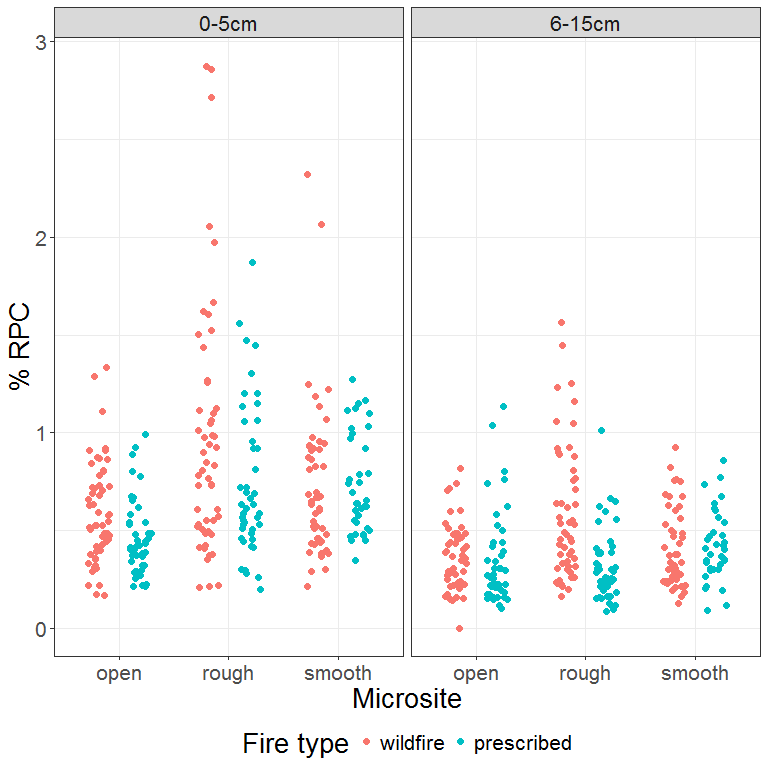
### Data exploration

The following dot plot shows percent carbon for all samples. Sites are ordered by fire type, then site ID value (integer: 1 - 33).

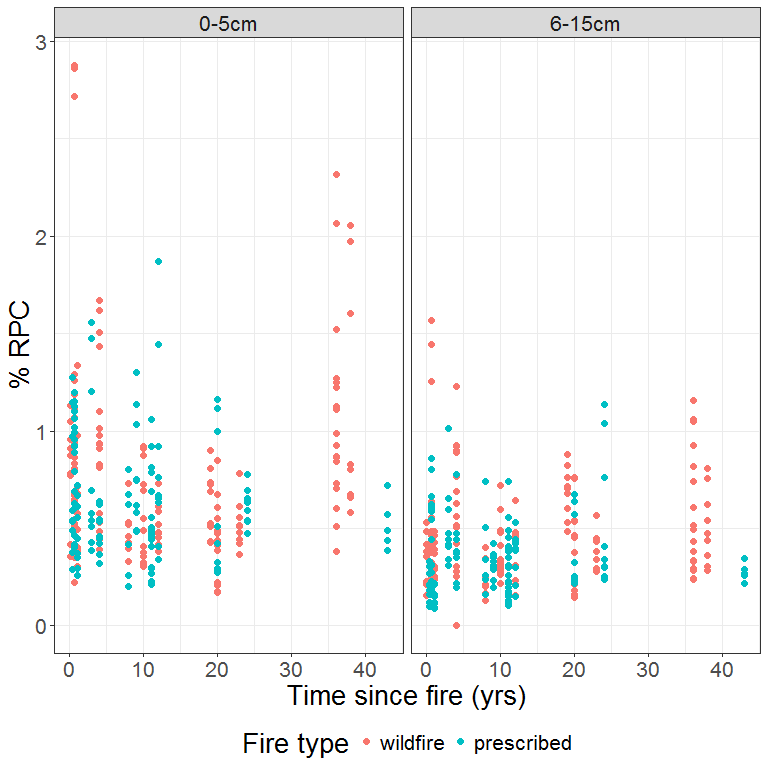


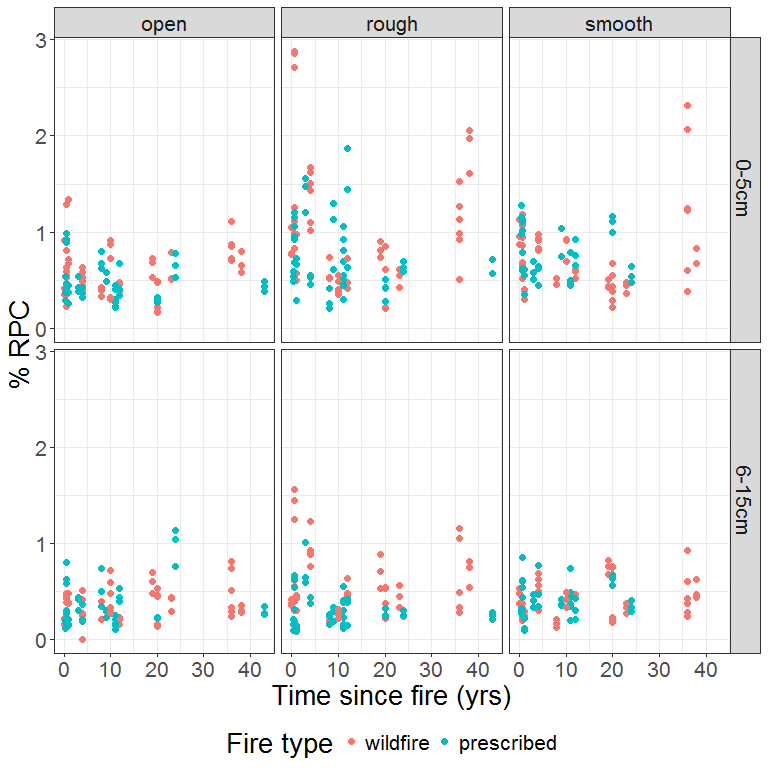


Distribution of values with respect to fire type and microsite.



Relative to time since fire.





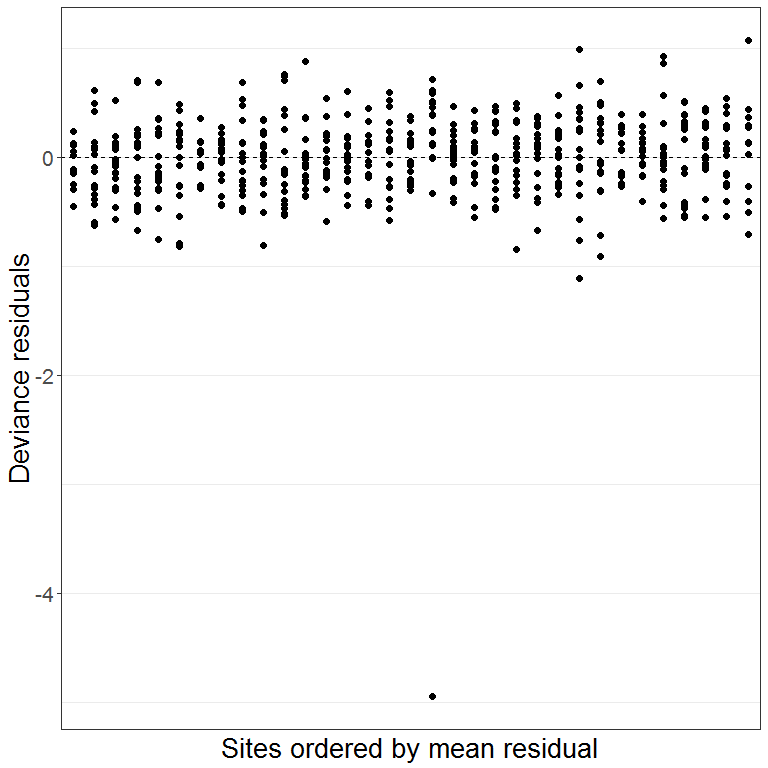
### Model

The form of the model follows that used for total carbon as the response, the only modification being the inclusion of total carbon as a co-variate to control for its effect on the level of recalcitrant carbon.

*Note:* we place the same constraint on the degrees of freedom available for the smooth term for time since fire as was done with the total carbon model.

dat.model <- DAT  
dat.model$flag.site <- 1  
dat.model$flag.ms <- 1  
  
mrpc <- gam(log(percentRPC) ~ s(tsf.years, by = dft, k=4) +   
 fireType + depth +  
 s(log(percentC)) +  
 s(microsite, bs = "re", by = flag.ms) +  
 s(fSiteId, bs = "re", by = flag.site),  
 data = dat.model,  
 method = "REML")

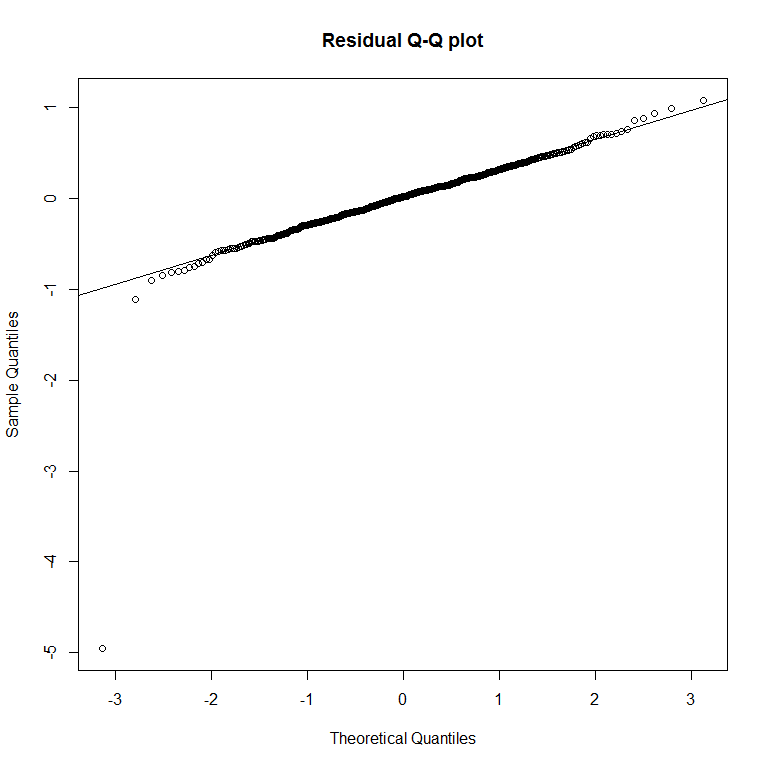
Check residuals against sites ordered in the same manner as for the percent carbon model.



No obvious pattern of residuals across sites. Check with anova:

## Analysis of Variance Table  
##   
## Response: resid  
## Df Sum Sq Mean Sq F value Pr(>F)  
## fSiteId 32 0.809 0.025273 0.1676 1  
## Residuals 537 80.959 0.150762

Checking for normality of residuals:



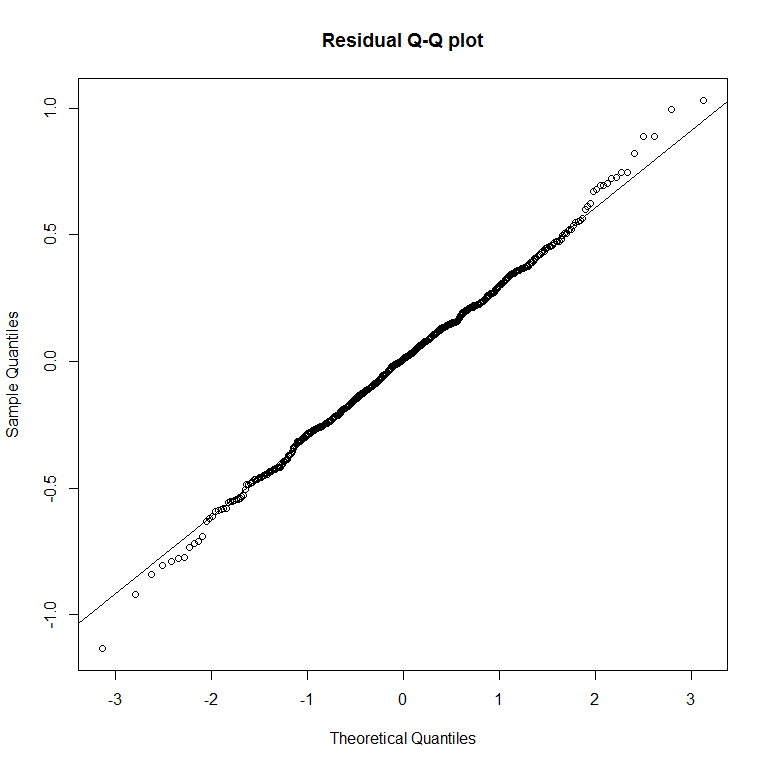
All looks good except for the single sample with a large negative value. Investigating this sample:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| sampleIndex | siteLabel | fireType | depth | microsite | tsf.years | percentC | percentRPC |
| 166 | Red 5W | wildfire | 6-15cm | open | 4 | 1.904756 | 0.0019925 |
| 167 | Red 5W | wildfire | 6-15cm | open | 4 | 2.102373 | 0.5140317 |
| 168 | Red 5W | wildfire | 6-15cm | open | 4 | 1.696940 | 0.2530811 |

The offender is sample number 166. Its value of RPC is very low compared to the other two replicates. Checked with Robert and he suggests that this was most likely a problem with the digestion. The sample will have had little influence on the model, but to be safe we discard it and re-run...

DAT <- filter(DAT, sampleIndex != 166)  
  
dat.model <- DAT  
dat.model$flag.ms <- 1  
dat.model$flag.site <- 1  
  
mrpc <- gam(log(percentRPC) ~ s(tsf.years, by = dft, k=4) +   
 fireType + depth +  
 s(log(percentC)) +  
 s(microsite, bs = "re", by = flag.ms) +  
 s(fSiteId, bs = "re", by = flag.site),  
 data = dat.model,  
 method = "REML")

Re-check the residuals:

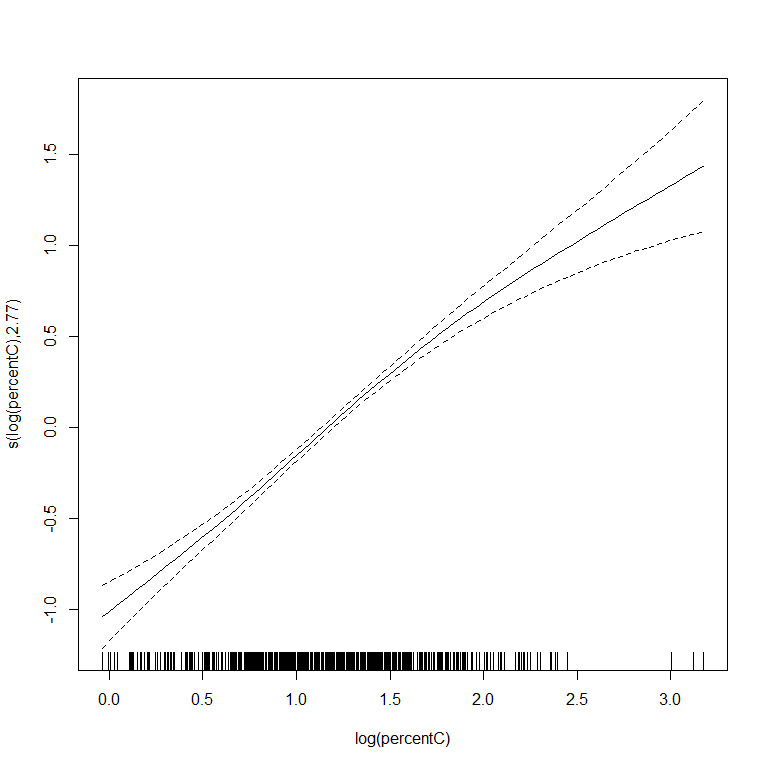


Summary of the fitted model:

##   
## Family: gaussian   
## Link function: identity   
##   
## Formula:  
## log(percentRPC) ~ s(tsf.years, by = dft, k = 4) + fireType +   
## depth + s(log(percentC)) + s(microsite, bs = "re", by = flag.ms) +   
## s(fSiteId, bs = "re", by = flag.site)  
##   
## Parametric coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.70554 0.06003 -11.754 <2e-16 \*\*\*  
## fireTypeprescribed -0.03393 0.08088 -0.420 0.6750   
## depth6-15cm -0.06390 0.03812 -1.676 0.0942 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df F p-value   
## s(tsf.years):dftwildfire.0-5cm 2.707 2.908 10.432 9.36e-07 \*\*\*  
## s(tsf.years):dftprescribed.0-5cm 1.000 1.001 0.102 0.749   
## s(tsf.years):dftwildfire.6-15cm 1.000 1.000 0.823 0.365   
## s(tsf.years):dftprescribed.6-15cm 1.337 1.562 0.521 0.419   
## s(log(percentC)) 2.774 3.562 98.292 < 2e-16 \*\*\*  
## s(microsite):flag.ms 1.039 2.000 1.207 0.121   
## s(fSiteId):flag.site 25.275 29.000 7.279 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = 0.715 Deviance explained = 73.4%  
## -REML = 217.57 Scale est. = 0.10398 n = 569

### Shape of the effect of total carbon

Total carbon (log-transformed) was fitted as a smooth term. Its estimated degrees of freedom (~ 2.8) indicate only a slight departure from linear. The graph below shows the shape of the fitted smooth function:



### Model summary with micro-site as fixed term

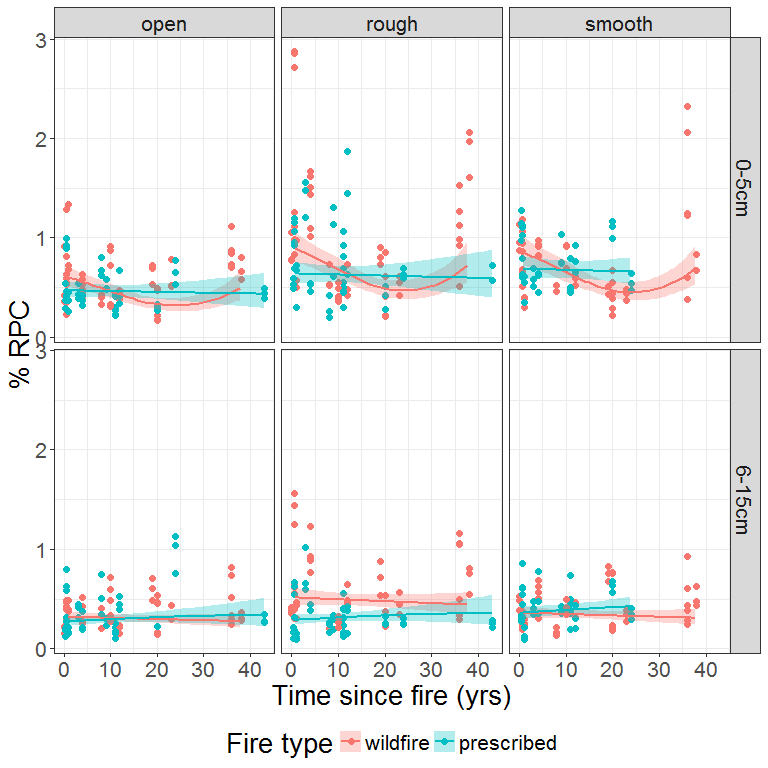
Changing micro-site from random to fixed effect gives coefficient estimates for the 'rough' and 'smooth' categories in the model summary. The coefficients for other terms (excepting the intercept) will remain unchanged apart from possible minor jiggle.

##   
## Family: gaussian   
## Link function: identity   
##   
## Formula:  
## log(percentRPC) ~ s(tsf.years, by = dft, k = 4) + fireType +   
## depth + s(log(percentC)) + microsite + s(fSiteId, bs = "re",   
## by = flag.site)  
##   
## Parametric coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.71155 0.05944 -11.971 <2e-16 \*\*\*  
## fireTypeprescribed -0.03589 0.08094 -0.443 0.6577   
## depth6-15cm -0.06712 0.03908 -1.717 0.0865 .   
## micrositerough 0.04533 0.03561 1.273 0.2035   
## micrositesmooth -0.02292 0.03675 -0.624 0.5330   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df F p-value   
## s(tsf.years):dftwildfire.0-5cm 2.706 2.908 10.409 9.69e-07 \*\*\*  
## s(tsf.years):dftprescribed.0-5cm 1.000 1.001 0.118 0.731   
## s(tsf.years):dftwildfire.6-15cm 1.000 1.000 0.782 0.377   
## s(tsf.years):dftprescribed.6-15cm 1.342 1.570 0.492 0.434   
## s(log(percentC)) 2.693 3.464 89.964 < 2e-16 \*\*\*  
## s(fSiteId):flag.site 25.267 29.000 7.267 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = 0.715 Deviance explained = 73.4%  
## -REML = 220.86 Scale est. = 0.10402 n = 569

### Model predictions

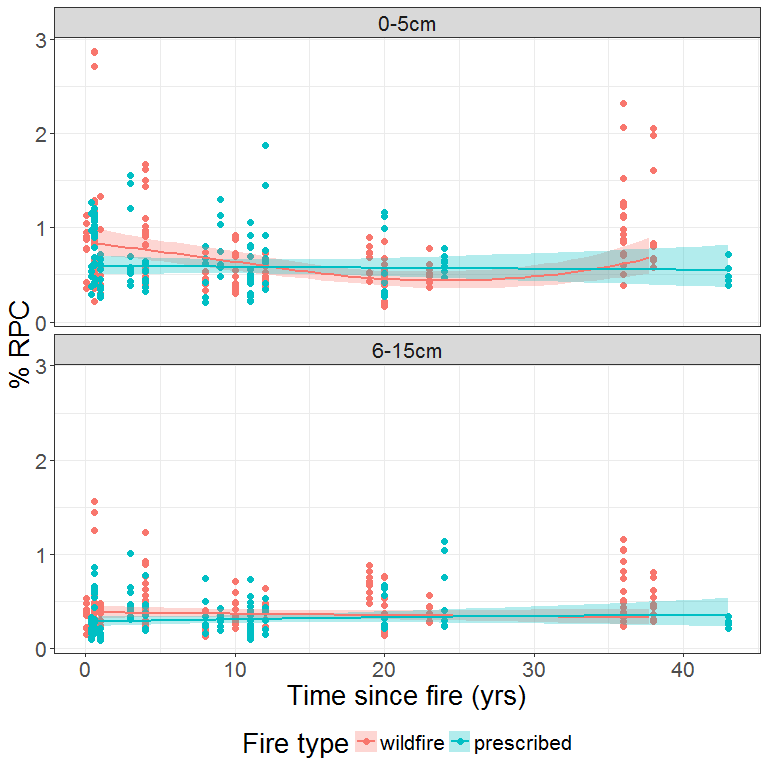
Predictions including micro-site effects.

Within each combination of fire type, depth and micro-site, we derive predictions over times since fire based on the median value of total carbon.



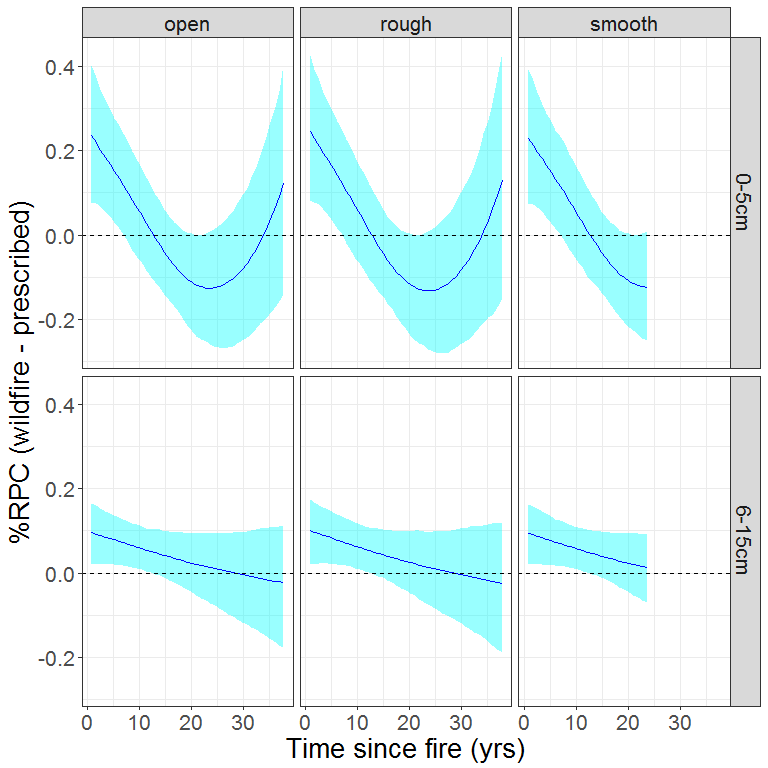
Predictions excluding micro-site effects.

For calculations here, we use the median value of total carbon within combinations of fire type and depth.



### Predicted difference between fire types

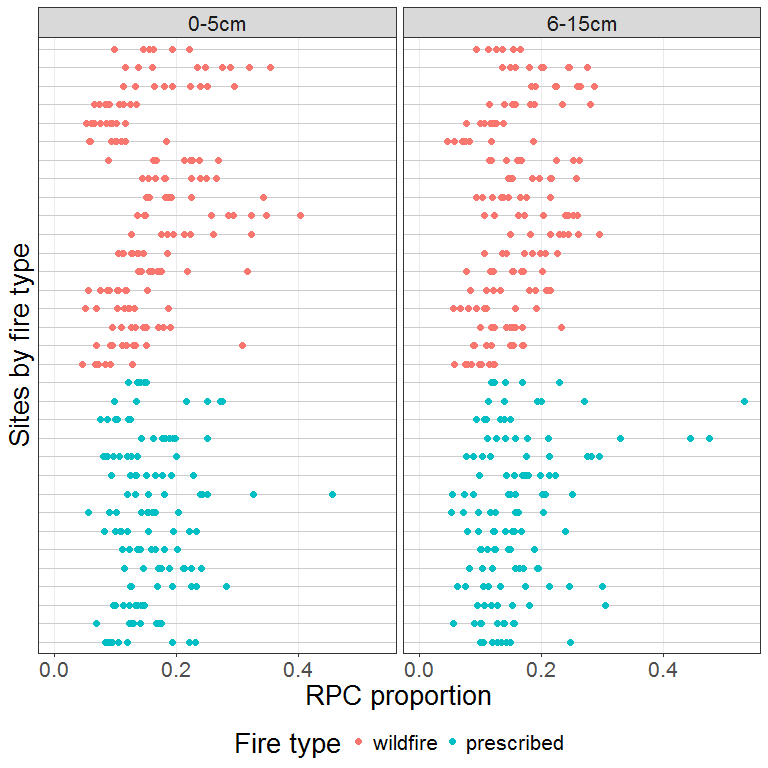
Predicted difference in %RPC between the two fire types based on posterior simulation.

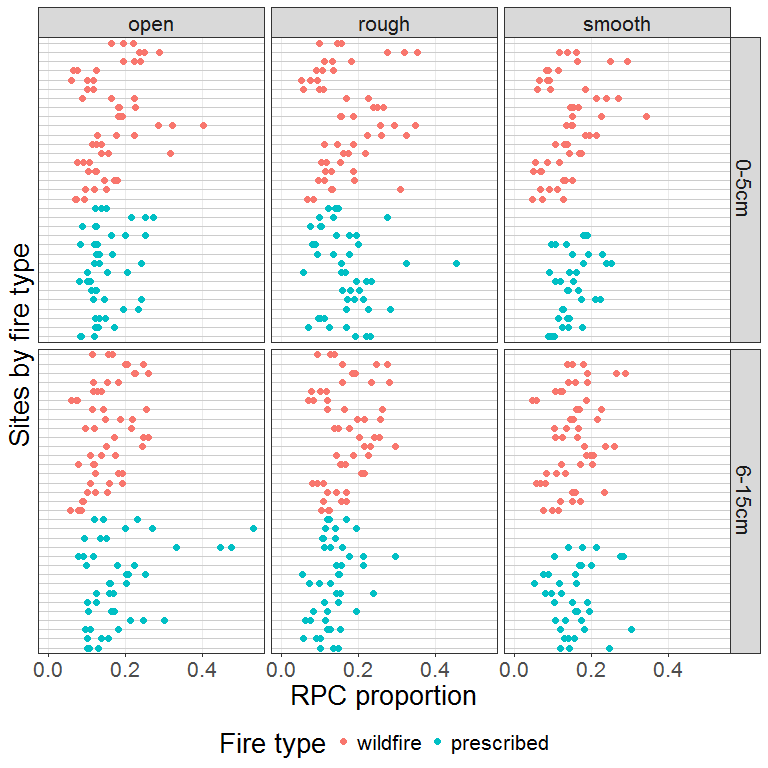


## Proportion of recalcitrant pyrogenic carbon

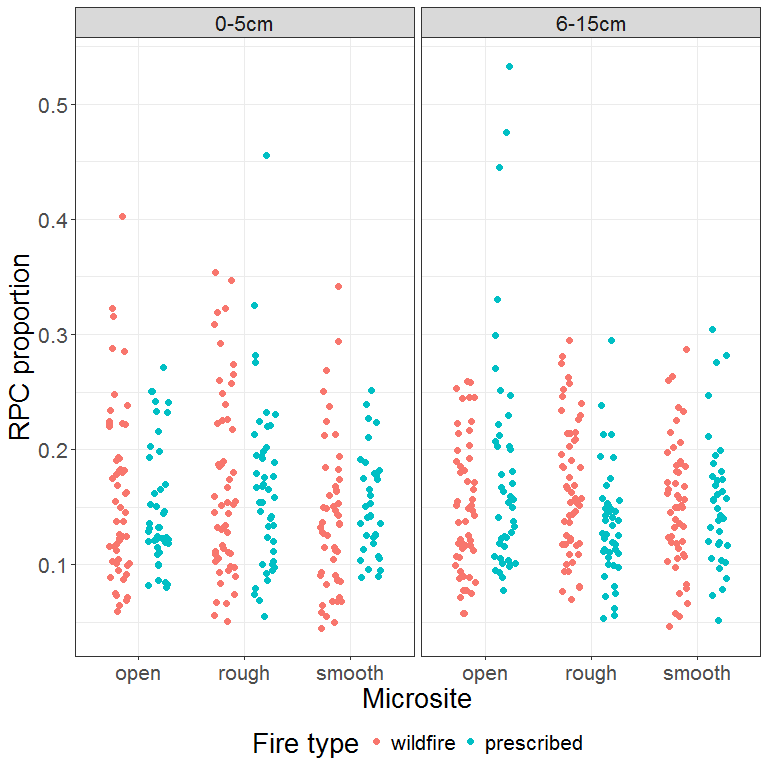
Here we fit a similar model, but with the response being RPC as a proportion of total carbon.

### Data exploration

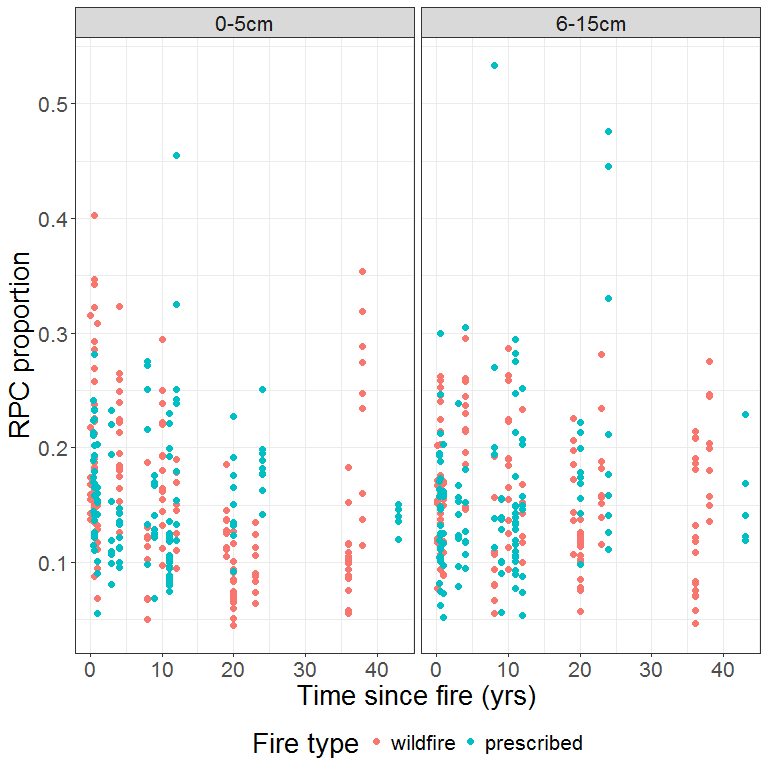


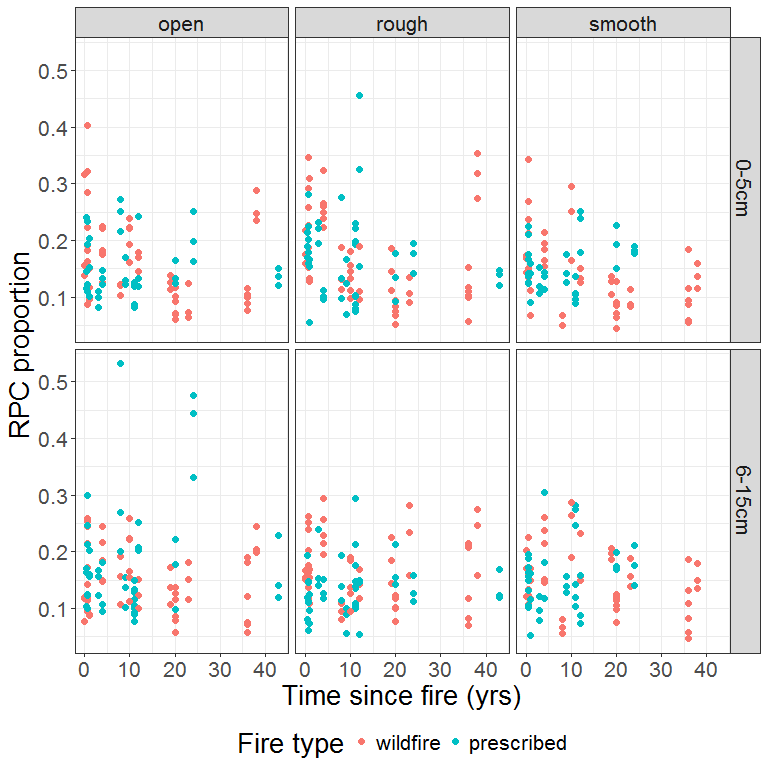


Distribution of values with respect to fire type and microsite.



Relative to time since fire.



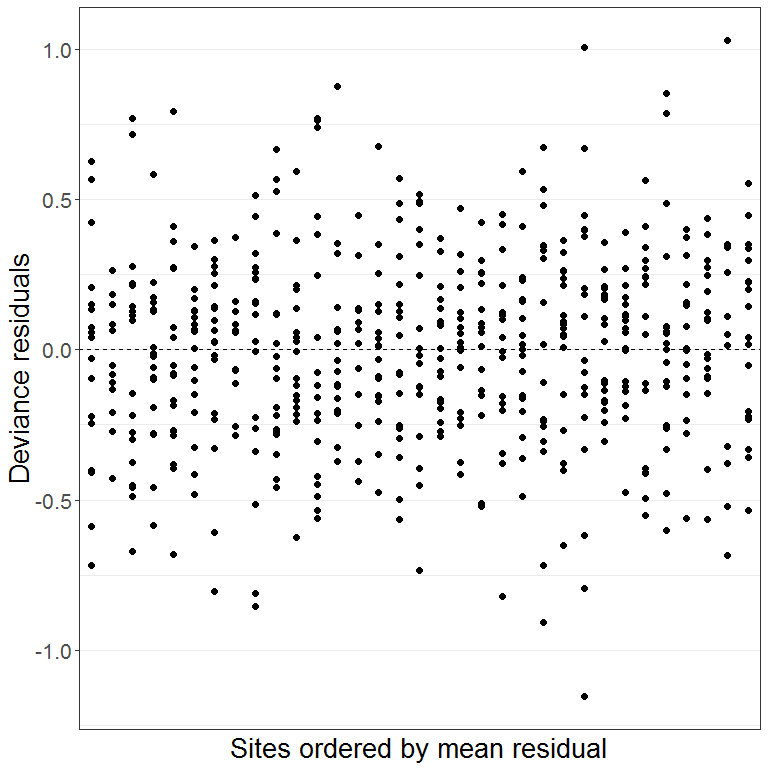


### Model

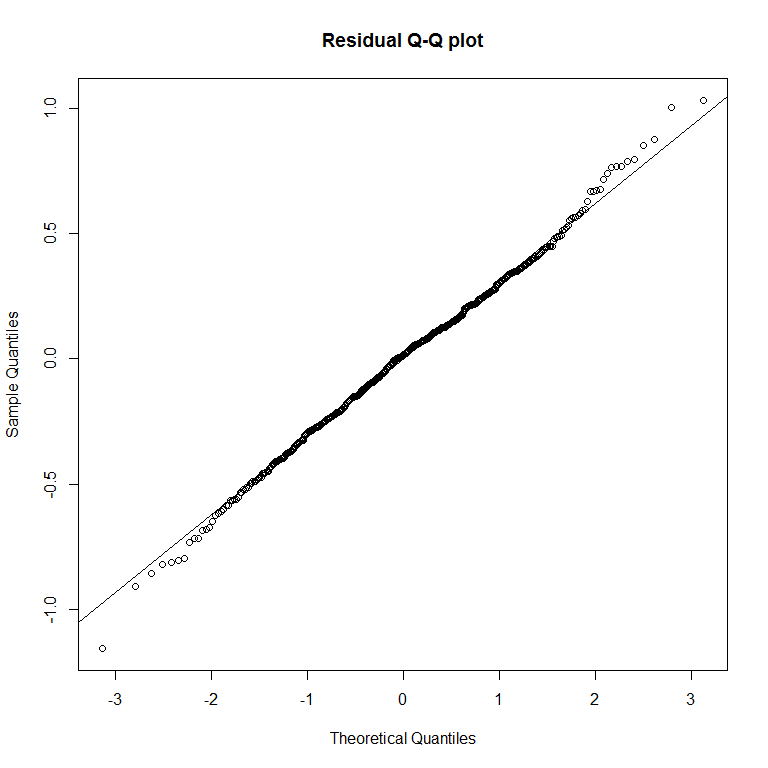
As before, we place a constraint on the degrees of freedom for the smooth term for time since fire.

dat.model <- DAT  
dat.model$flag.ms <- 1  
dat.model$flag.site <- 1  
  
mrpc.prop <- gam(log(propRPC) ~ s(tsf.years, by = dft, k=4) +   
 fireType + depth +  
 s(microsite, bs = "re", by = flag.ms) +  
 s(fSiteId, bs = "re", by = flag.site),  
 data = dat.model,  
 method = "REML")

Check residuals against sites ordered in the same manner as for the percent carbon model.



Checking for normality of residuals:



Summary of the fitted model:

##   
## Family: gaussian   
## Link function: identity   
##   
## Formula:  
## log(propRPC) ~ s(tsf.years, by = dft, k = 4) + fireType + depth +   
## s(microsite, bs = "re", by = flag.ms) + s(fSiteId, bs = "re",   
## by = flag.site)  
##   
## Parametric coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.929871 0.059529 -32.419 <2e-16 \*\*\*  
## fireTypeprescribed -0.003524 0.082720 -0.043 0.966   
## depth6-15cm 0.008754 0.027895 0.314 0.754   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df F p-value   
## s(tsf.years):dftwildfire.0-5cm 2.695 2.903 9.630 2.8e-06 \*\*\*  
## s(tsf.years):dftprescribed.0-5cm 1.000 1.000 0.035 0.8520   
## s(tsf.years):dftwildfire.6-15cm 1.000 1.000 1.288 0.2568   
## s(tsf.years):dftprescribed.6-15cm 1.163 1.292 0.601 0.3915   
## s(microsite):flag.ms 1.128 2.000 1.401 0.0968 .   
## s(fSiteId):flag.site 25.550 29.000 7.706 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = 0.378 Deviance explained = 41.6%  
## -REML = 220.16 Scale est. = 0.10618 n = 569

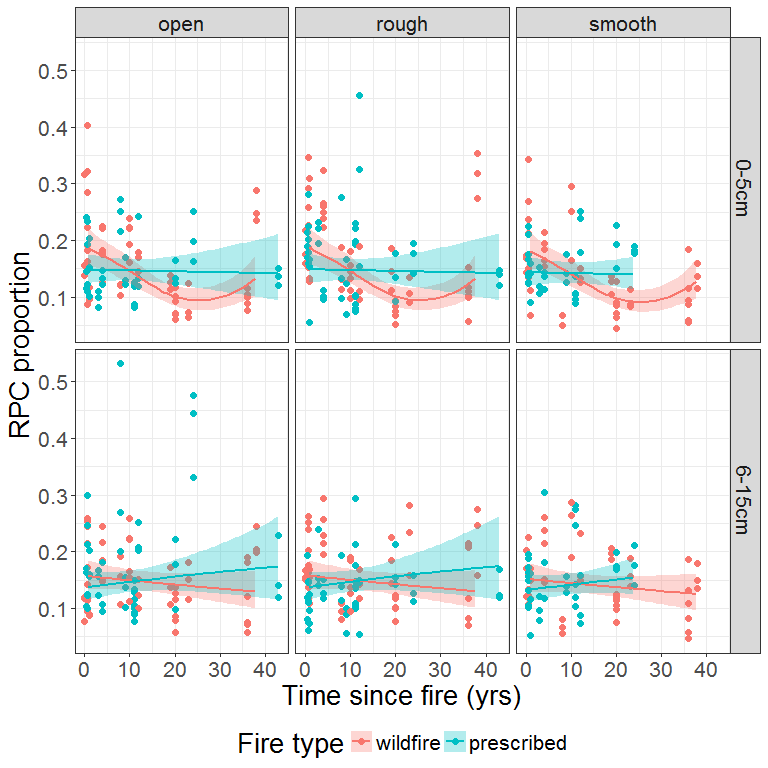
### Model summary with micro-site as fixed term

Changing micro-site from random to fixed effect gives coefficient estimates for the 'rough' and 'smooth' categories in the model summary. The coefficients for other terms (excepting the intercept) will remain unchanged apart from possible minor jiggle.

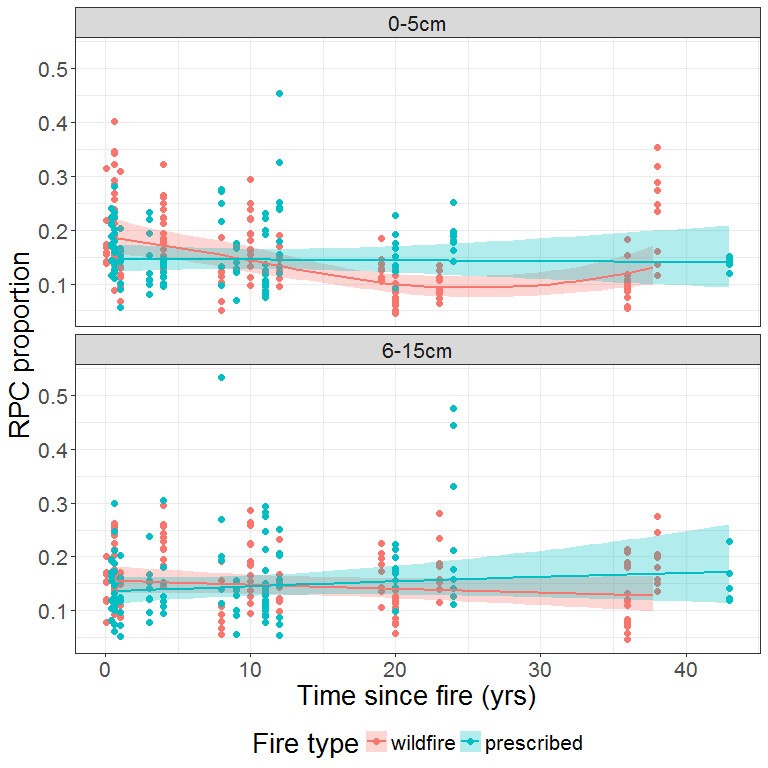
##   
## Family: gaussian   
## Link function: identity   
##   
## Formula:  
## log(propRPC) ~ s(tsf.years, by = dft, k = 4) + fireType + depth +   
## microsite + s(fSiteId, bs = "re", by = flag.site)  
##   
## Parametric coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.912227 0.060521 -31.596 <2e-16 \*\*\*  
## fireTypeprescribed -0.005089 0.082801 -0.061 0.9510   
## depth6-15cm 0.008806 0.027892 0.316 0.7523   
## micrositerough 0.006417 0.032790 0.196 0.8449   
## micrositesmooth -0.060970 0.034379 -1.773 0.0767 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df F p-value   
## s(tsf.years):dftwildfire.0-5cm 2.694 2.902 9.619 2.85e-06 \*\*\*  
## s(tsf.years):dftprescribed.0-5cm 1.000 1.000 0.045 0.831   
## s(tsf.years):dftwildfire.6-15cm 1.000 1.000 1.282 0.258   
## s(tsf.years):dftprescribed.6-15cm 1.169 1.302 0.558 0.409   
## s(fSiteId):flag.site 25.554 29.000 7.719 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = 0.378 Deviance explained = 41.7%  
## -REML = 223.39 Scale est. = 0.10616 n = 569

### Model predictions

Including micro-site random effects:

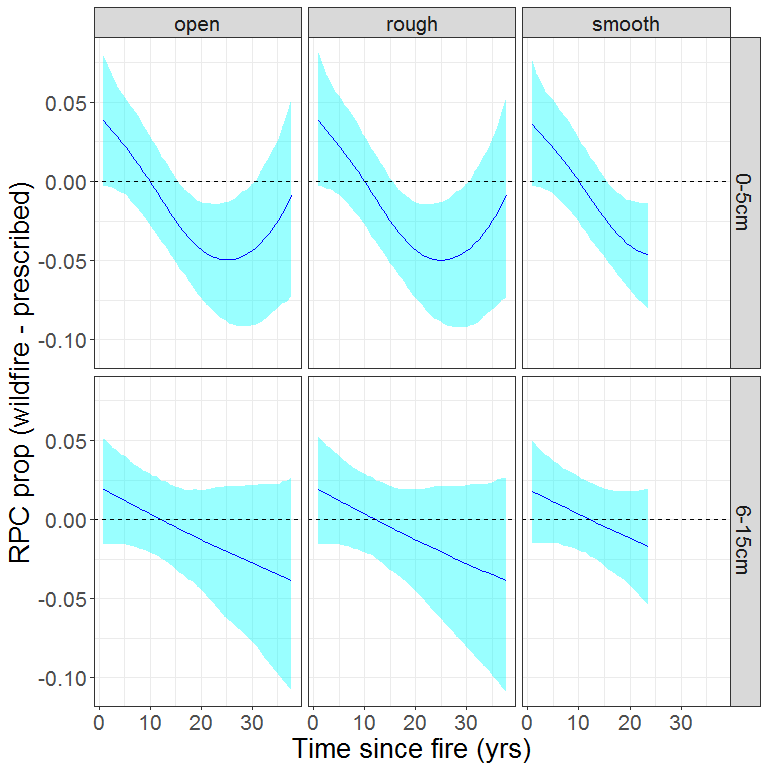


Excluding micro-site random effects:



### Predicted difference between fire types

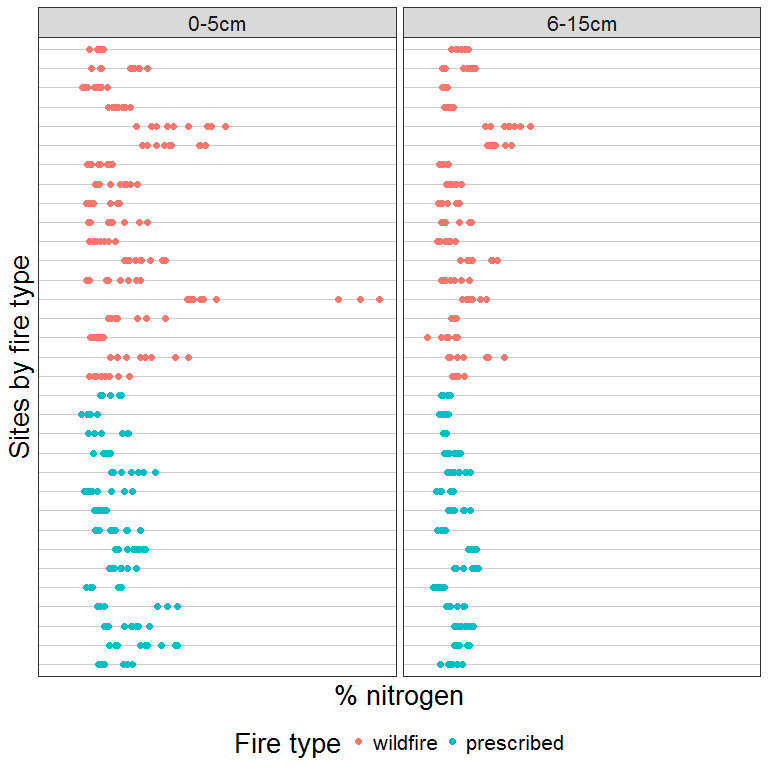
Predicted difference in RPC proportion between the two fire types based on posterior simulation.



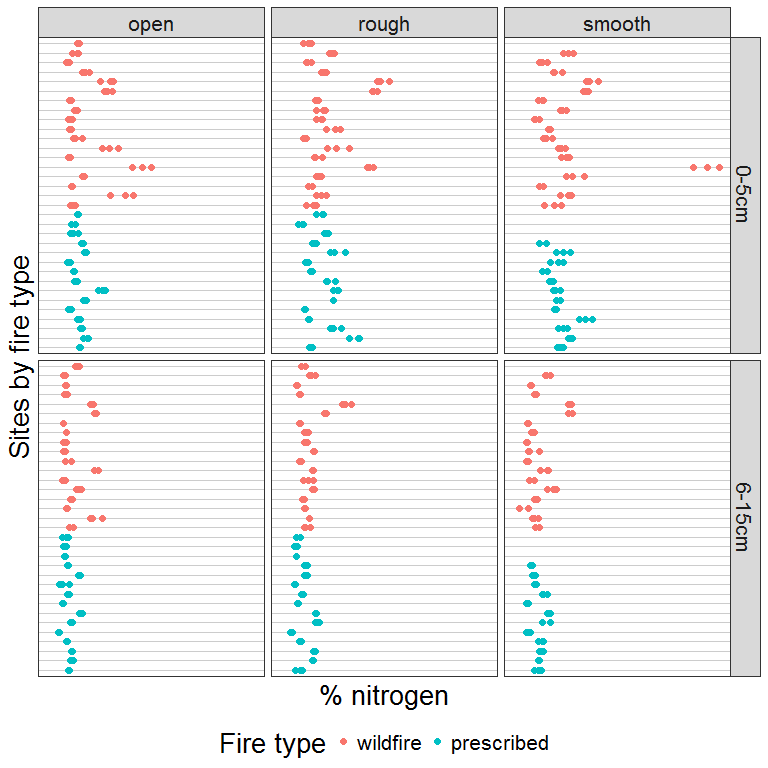
## Total nitrogen

### Data exploration

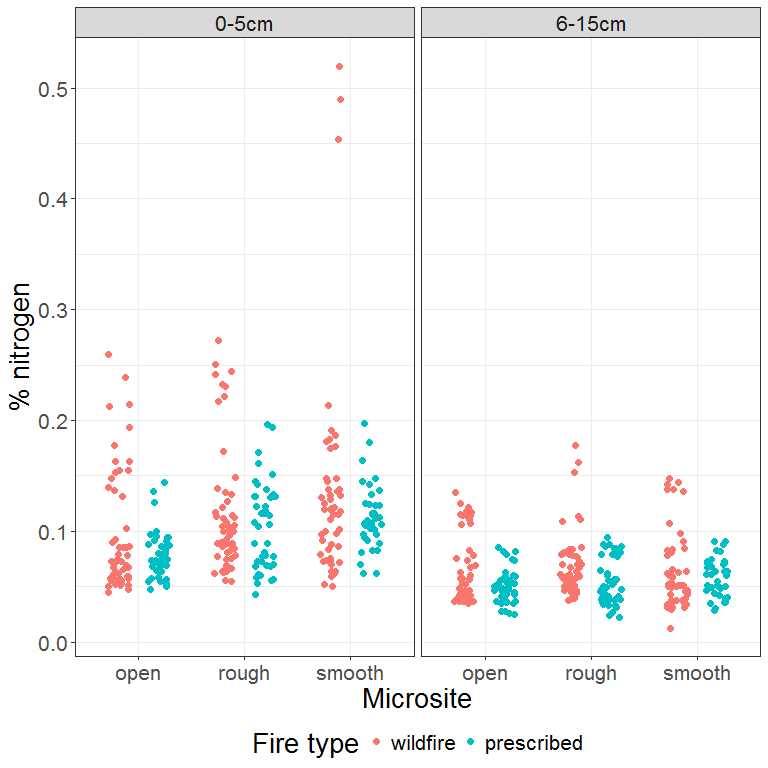
The following dot plot shows percent nitrogen values for all samples. Sites are ordered by fire type, then site ID value (integer: 1 - 33).



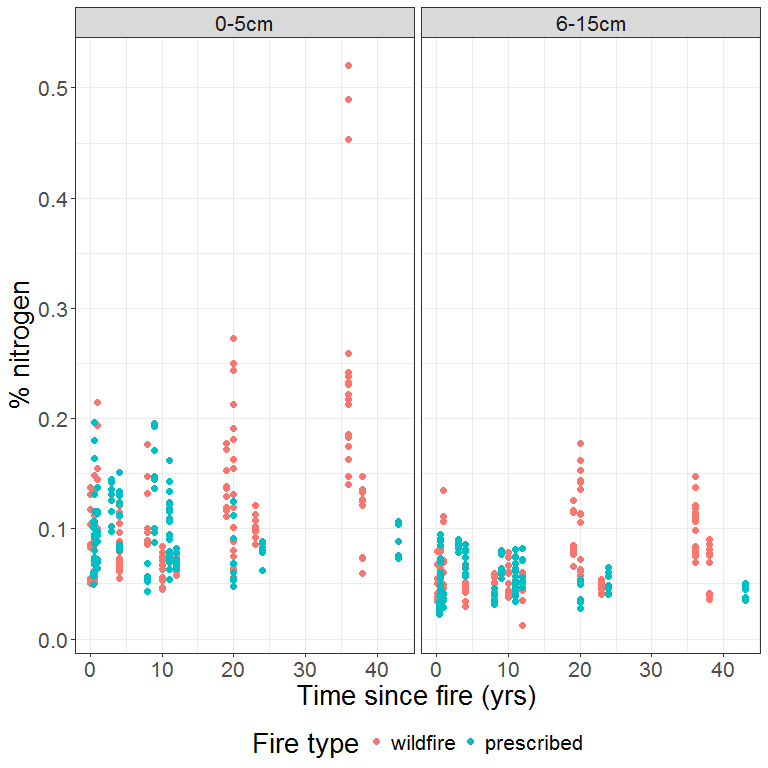
Same data, split by micro-site. Note that 'smooth' micro-site samples are missing for three prescribed fire sites and one wildfire site.

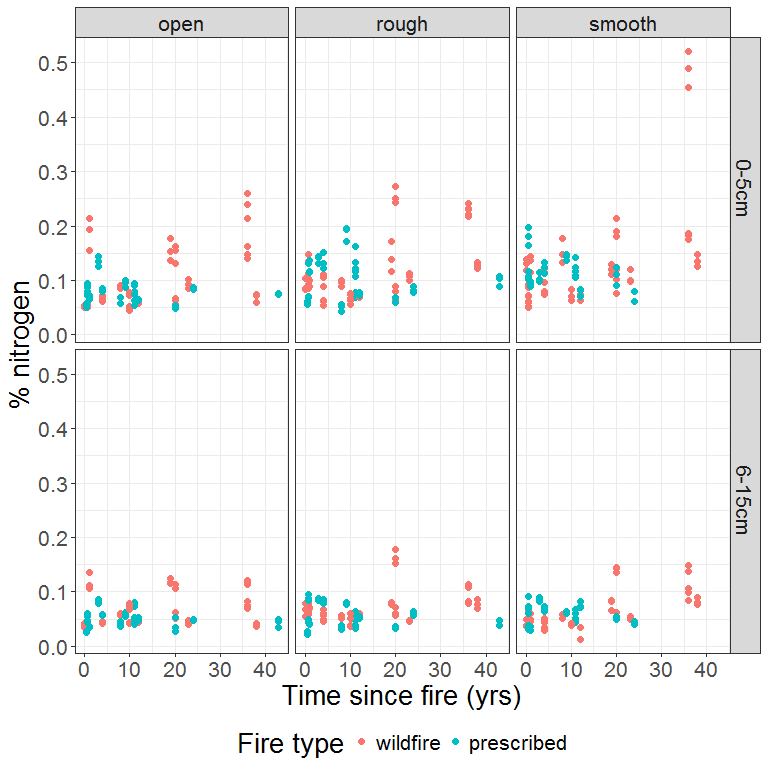


Distribution of values with respect to fire type and microsite.



Pattern of percent nitrogen values relative to time since fire.





### Model

We model total nitrogen using the same approach that proved most informative for carbon. The model includes:

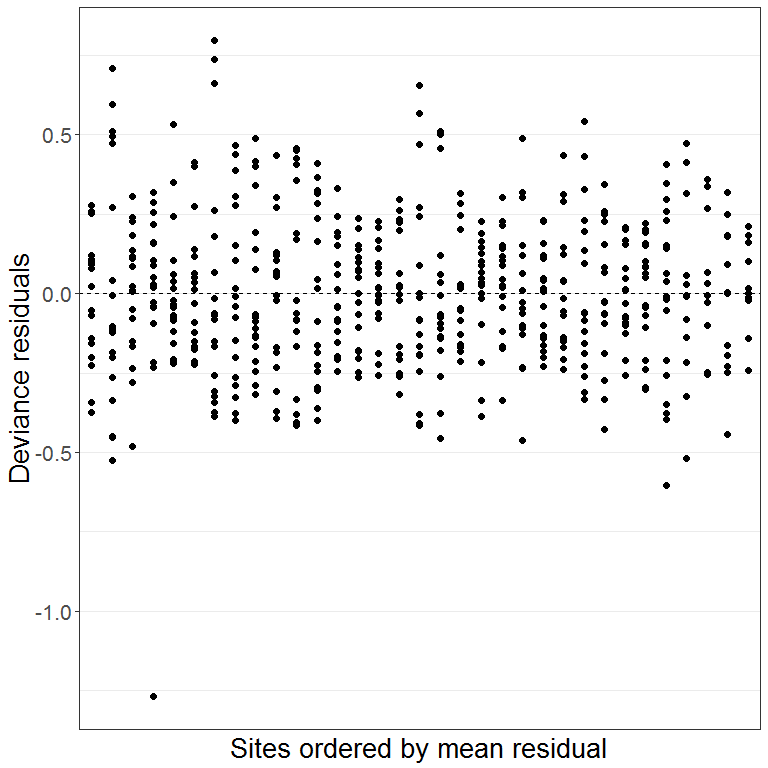
* a smooth term for the interaction between time since fire, depth and fire type;
* fixed categorical terms for fire type, depth and micro-site;
* a random effect for site.

We use a Normal distribution with identity link function. Reponse values are log-transformed.

As for the carbon models, we constrain the degrees of freedom for the smooth term for time since fire.

dat.model <- DAT  
dat.model$flag.site <- 1  
  
mtotalN.site <- gam(log(percentN) ~ s(tsf.years, by = dft, k=4) +   
 fireType + depth +   
 microsite +   
 s(fSiteId, bs = "re", by = flag.site),  
 data = dat.model,  
 method = "REML")

Graph of residuals against sites to check the performance of the random effect term.



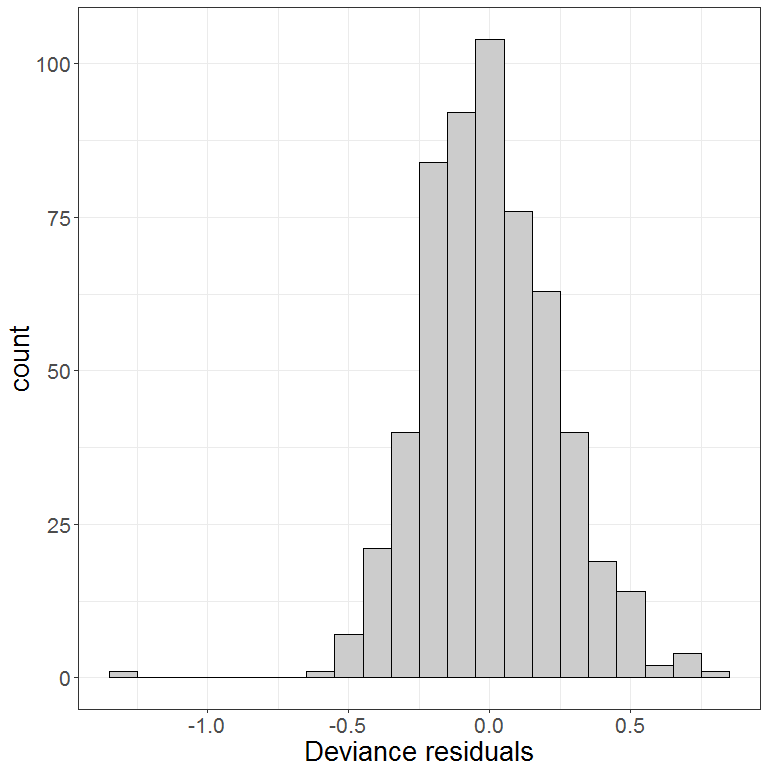
No systematic pattern is evident. However, there is one outlier:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| siteLabel | siteId | fireType | depth | microsite | rep | percentN | fitted |
| Red 15W | 4 | wildfire | 6-15cm | smooth | 1 | 0.0122835 | 0.0437039 |
| Red 15W | 4 | wildfire | 6-15cm | smooth | 2 | 0.0351379 | 0.0437039 |
| Red 15W | 4 | wildfire | 6-15cm | smooth | 3 | 0.0346067 | 0.0437039 |

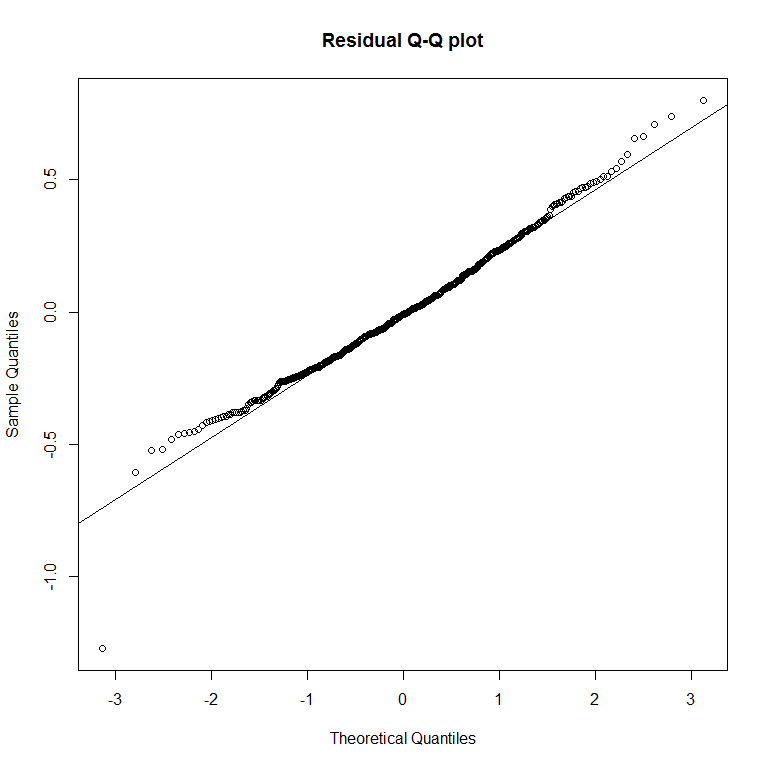
Replicate number 1 has a somewhat lower value for total nitrogen that the other two replicates.

Further model validation...

Distribution of residuals:

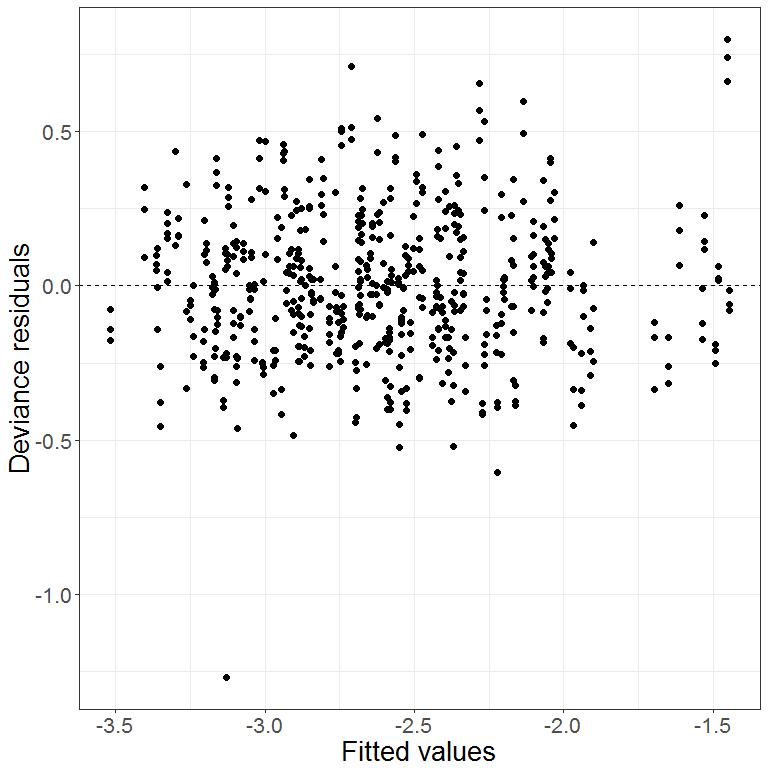


Checking for normality of residuals:

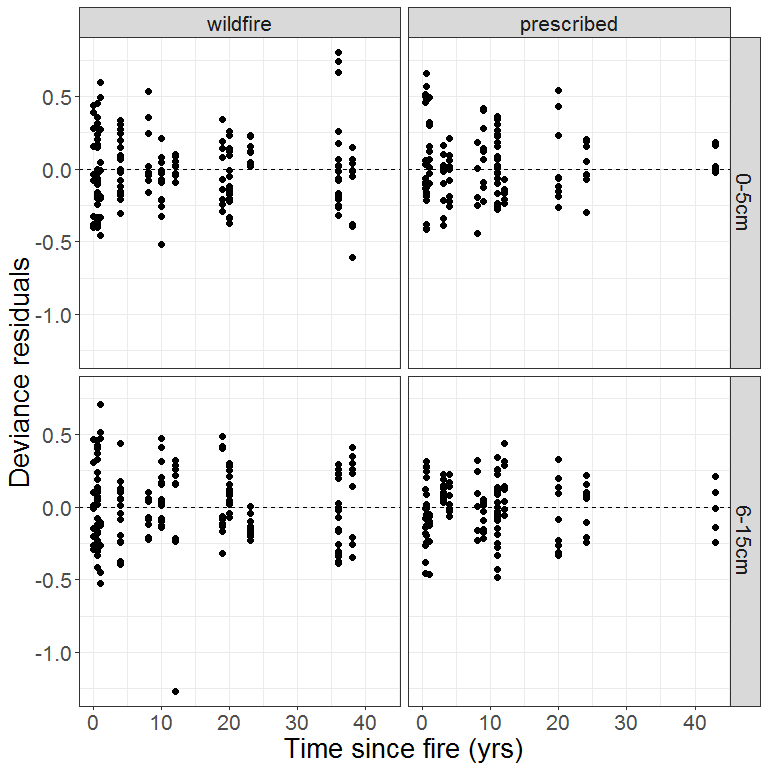


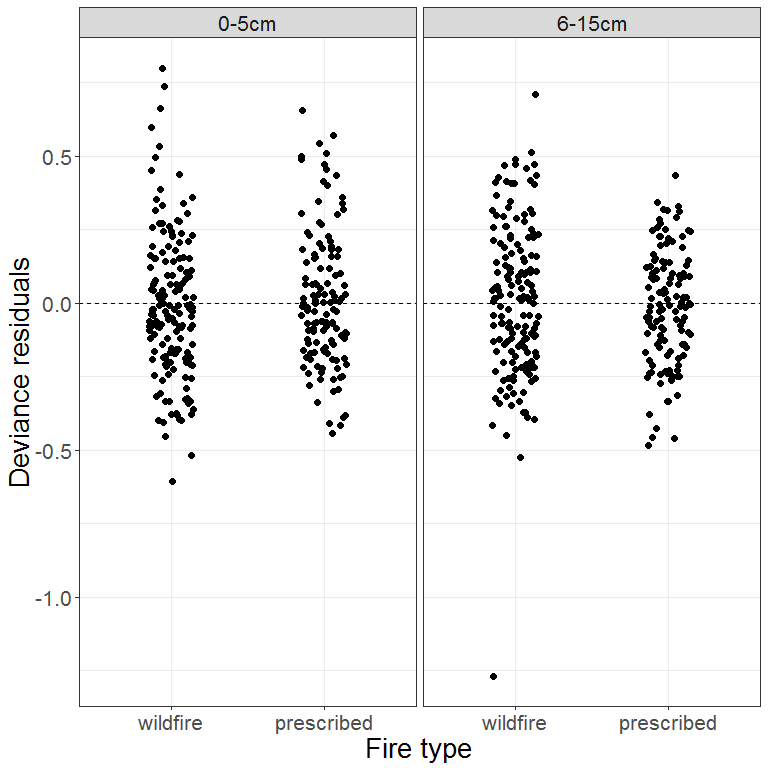
Ignoring the outlier, there is some departure at the tails but nothing dramatic.

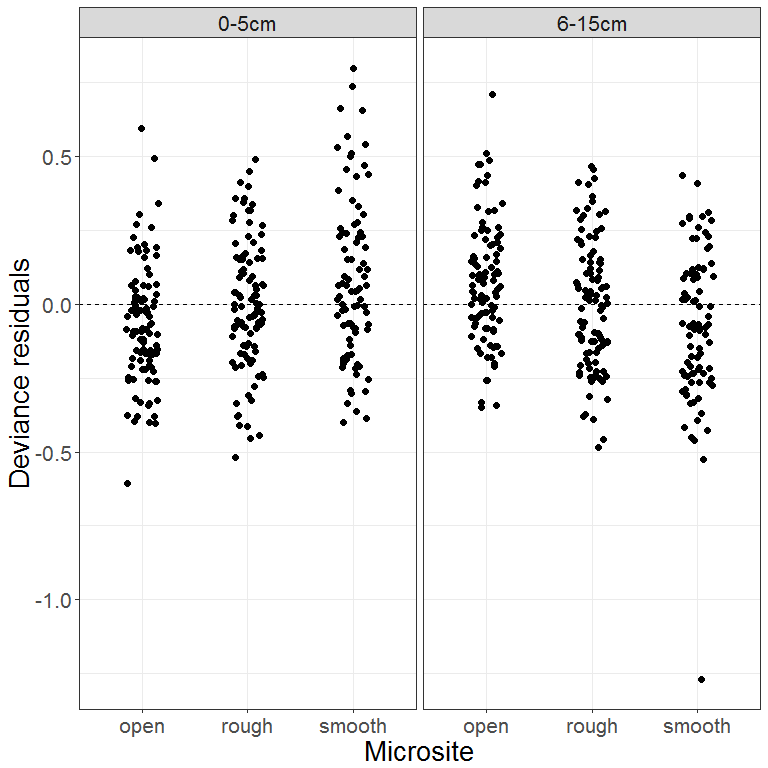
Graph of residuals against fitted values to check for homogeneity.



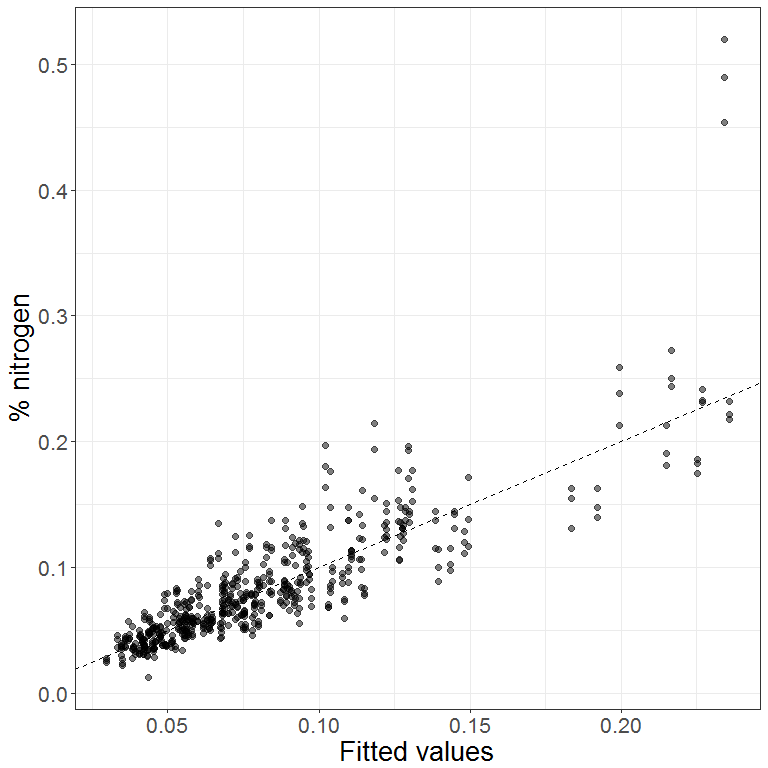
Residuals against co-variates.







Response against fitted values:



The model under-predicts for a group of samples with atypically high nitrogen values (> 0.4%). The samples in question are a group of three replicates:

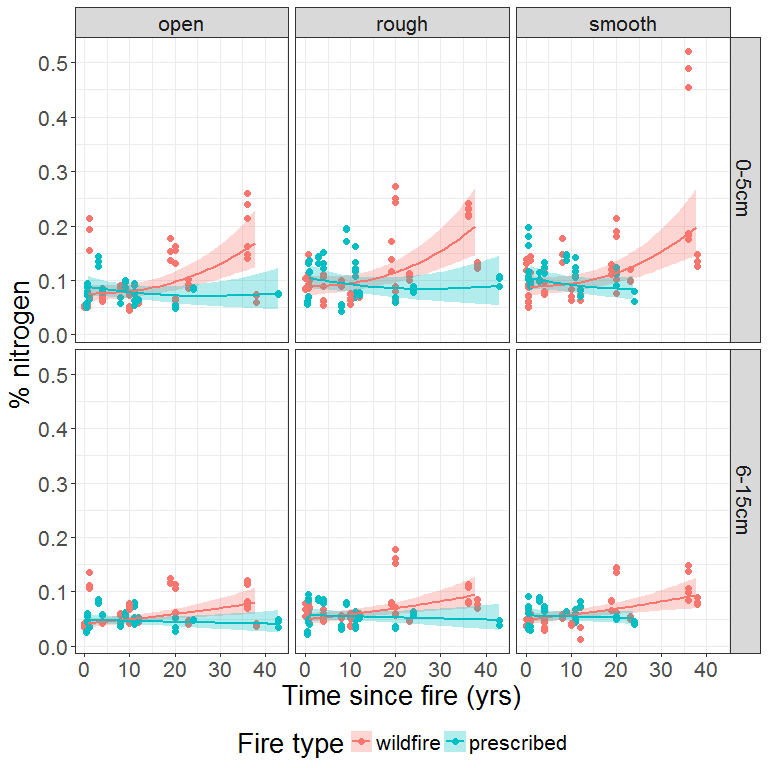
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| siteLabel | siteId | fireType | depth | microsite | rep | percentN | fitted |
| Kt 30+ | 7 | wildfire | 0-5cm | smooth | 1 | 0.4896047 | 0.233984 |
| Kt 30+ | 7 | wildfire | 0-5cm | smooth | 2 | 0.4536229 | 0.233984 |
| Kt 30+ | 7 | wildfire | 0-5cm | smooth | 3 | 0.5199646 | 0.233984 |

### 

### Model summary

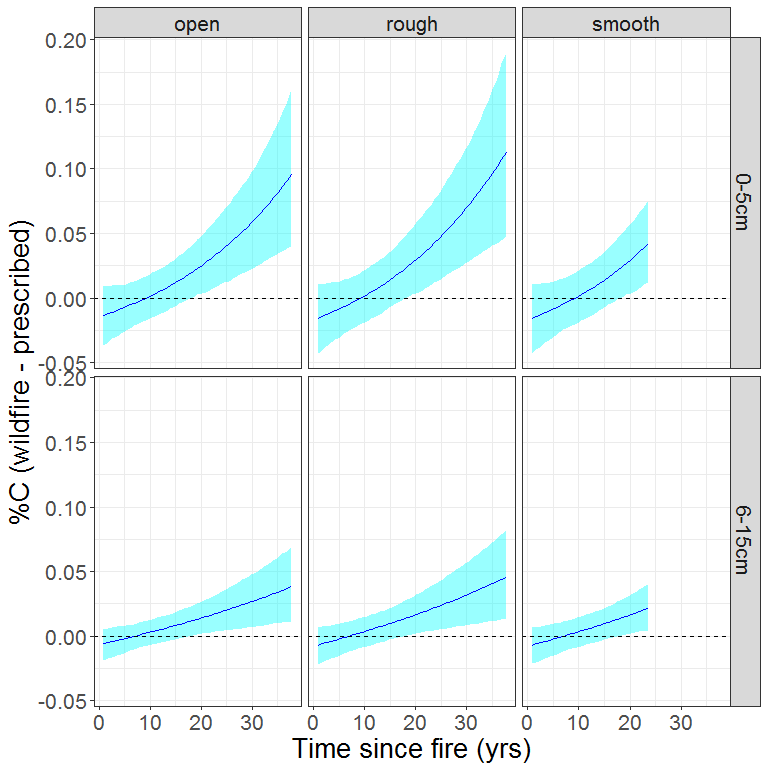
##   
## Family: gaussian   
## Link function: identity   
##   
## Formula:  
## log(percentN) ~ s(tsf.years, by = dft, k = 4) + fireType + depth +   
## microsite + s(fSiteId, bs = "re", by = flag.site)  
##   
## Parametric coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -2.42346 0.07070 -34.276 < 2e-16 \*\*\*  
## fireTypeprescribed -0.10410 0.10168 -1.024 0.306   
## depth6-15cm -0.55316 0.02085 -26.536 < 2e-16 \*\*\*  
## micrositerough 0.16692 0.02450 6.812 2.62e-11 \*\*\*  
## micrositesmooth 0.15915 0.02572 6.187 1.23e-09 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df F p-value   
## s(tsf.years):dftwildfire.0-5cm 2.192 2.520 9.318 4.94e-05 \*\*\*  
## s(tsf.years):dftprescribed.0-5cm 1.923 2.287 1.583 0.1886   
## s(tsf.years):dftwildfire.6-15cm 1.000 1.000 10.181 0.0015 \*\*   
## s(tsf.years):dftprescribed.6-15cm 1.001 1.001 0.308 0.5794   
## s(fSiteId):flag.site 27.704 29.000 23.128 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = 0.768 Deviance explained = 78.3%  
## -REML = 74.109 Scale est. = 0.059283 n = 569

### Model predictions



### Predicted difference between fire types

Here we use posterior simulation to generate a set of 1000 credible mean trends from the fitted GAM for total nitrogen and, for each trend, calculate the difference in nitrogen between the two fire types.

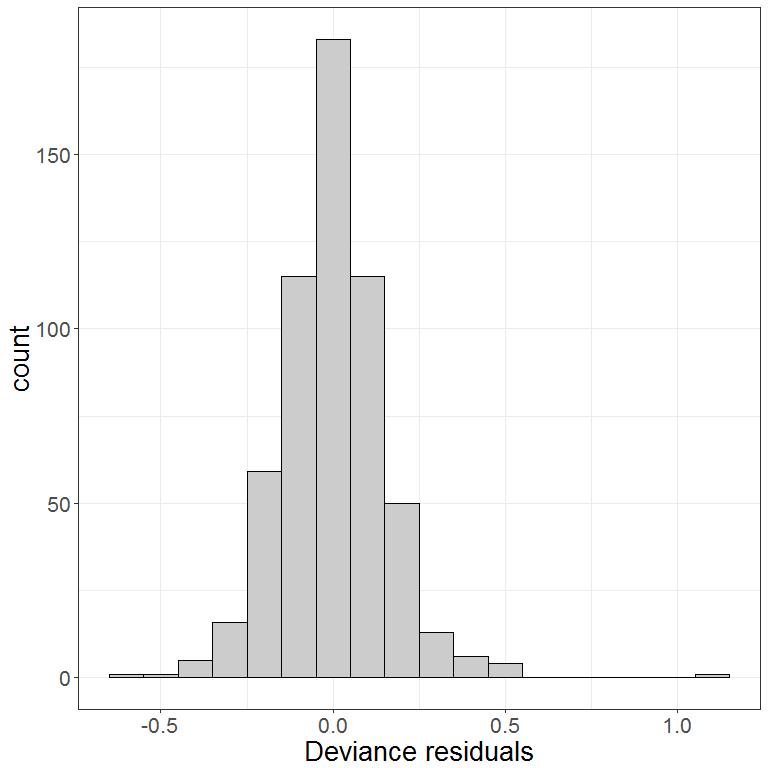


## Carbon to nitrogen ratio

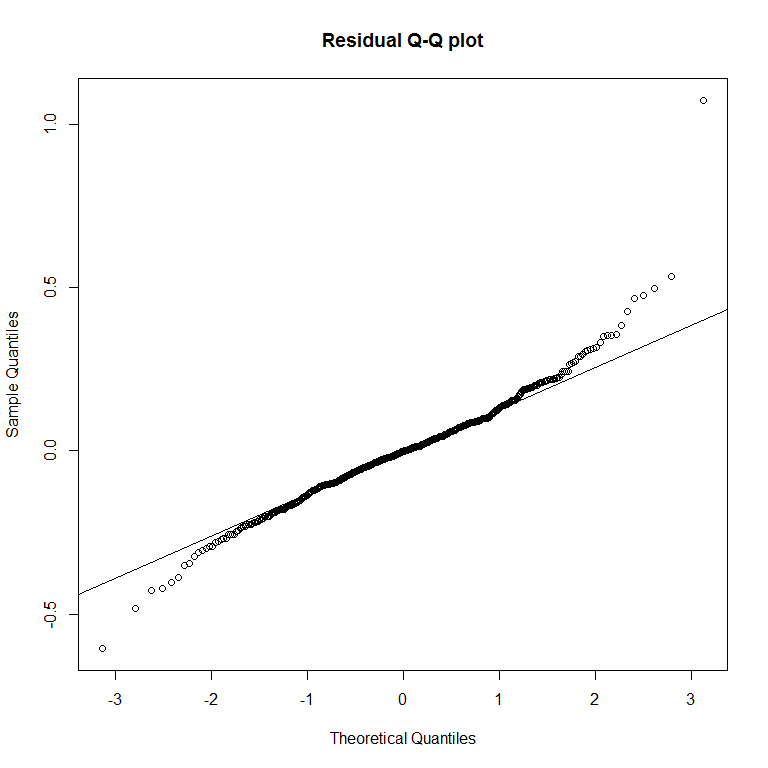
### Model

dat.model <- DAT %>%  
 mutate(cnratio = percentC / percentN)  
  
dat.model$flag.site <- 1  
  
mratio <- gam(log(cnratio) ~ s(tsf.years, by = dft, k=4) +   
 fireType + depth +   
 microsite +   
 s(fSiteId, bs = "re", by = flag.site),  
 data = dat.model,  
 method = "REML")

Distribution of residuals:



Checking for normality of residuals:

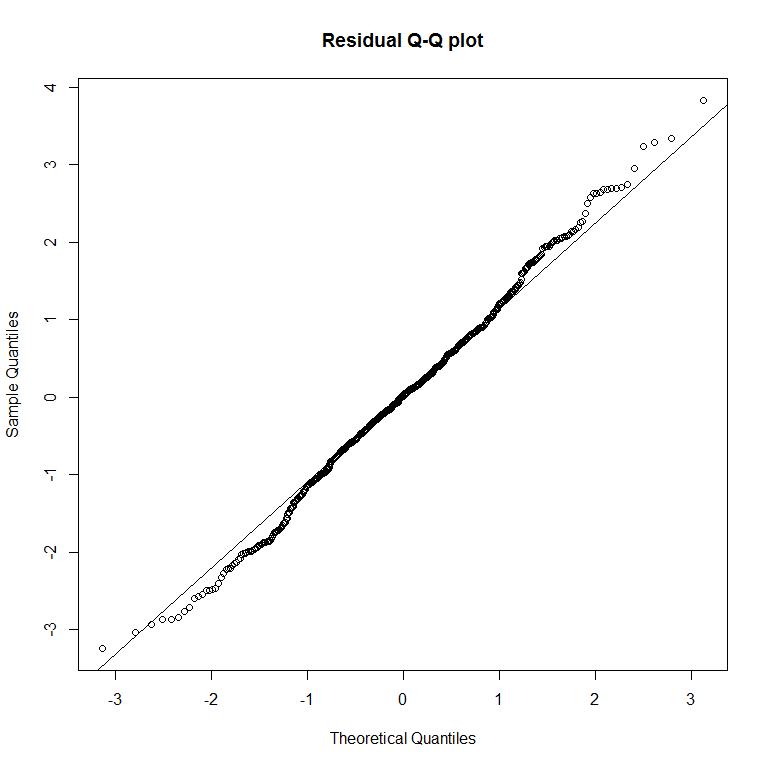


The behaviour at the tails looks more like systematic departure from normality rather than just random wiggle. To address this, we try re-fitting the model with a scaled-t (heavy tailed) distribution.

*Note: this was also done the model of C:N ratio in the Eden soils data.*

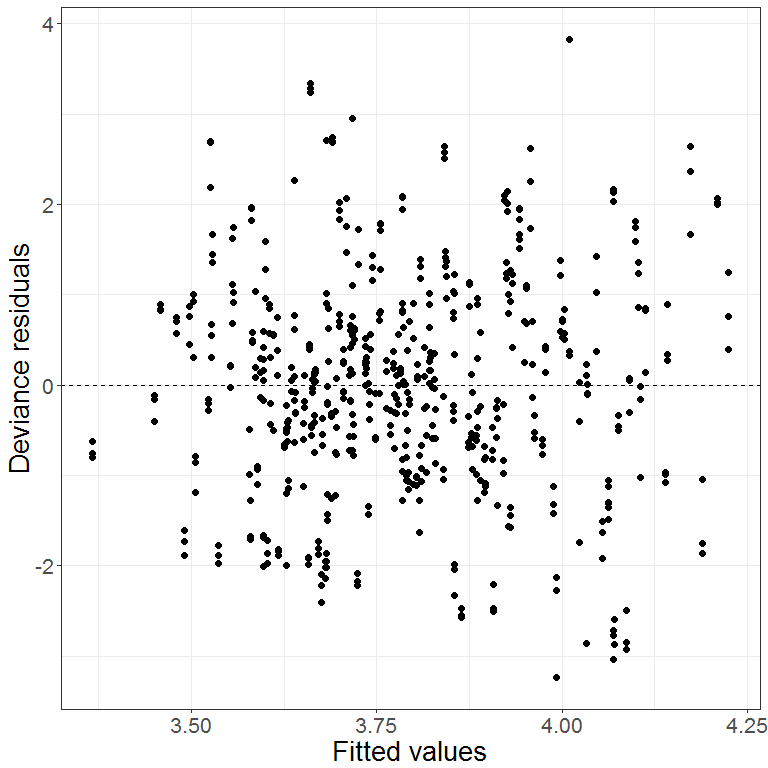
mratio <- gam(log(cnratio) ~ s(tsf.years, by = dft, k=4) +   
 fireType + depth +   
 microsite +   
 s(fSiteId, bs = "re", by = flag.site),  
 data = dat.model,  
 family = scat(),  
 method = "REML")

Re-checking the distribution of residuals:

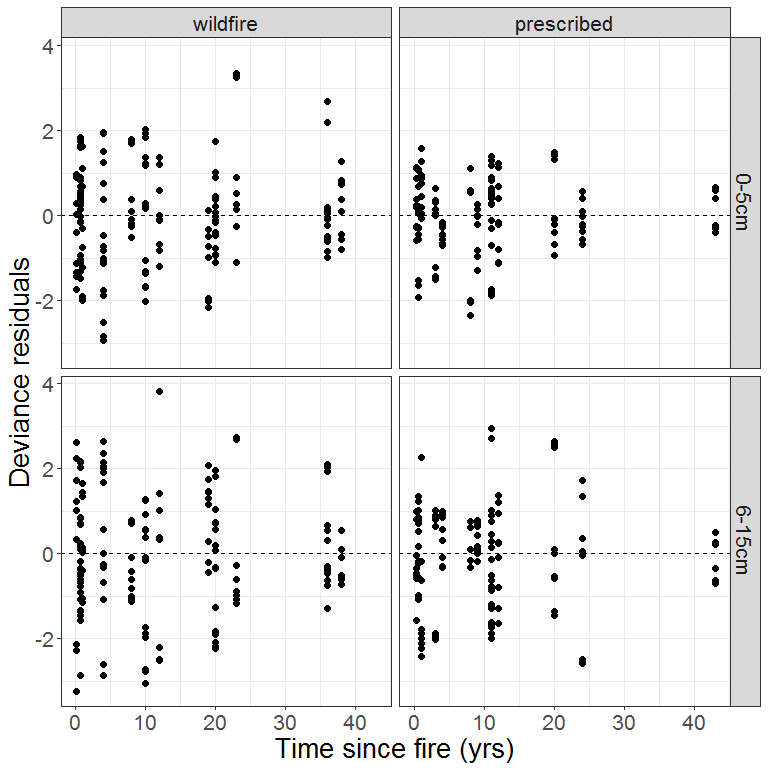


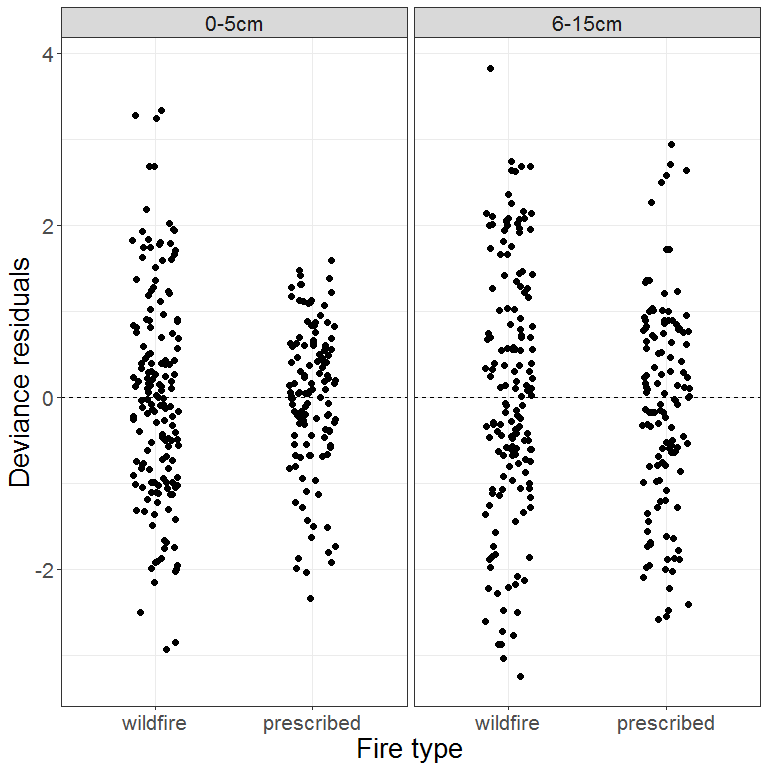
That looks a bit better.

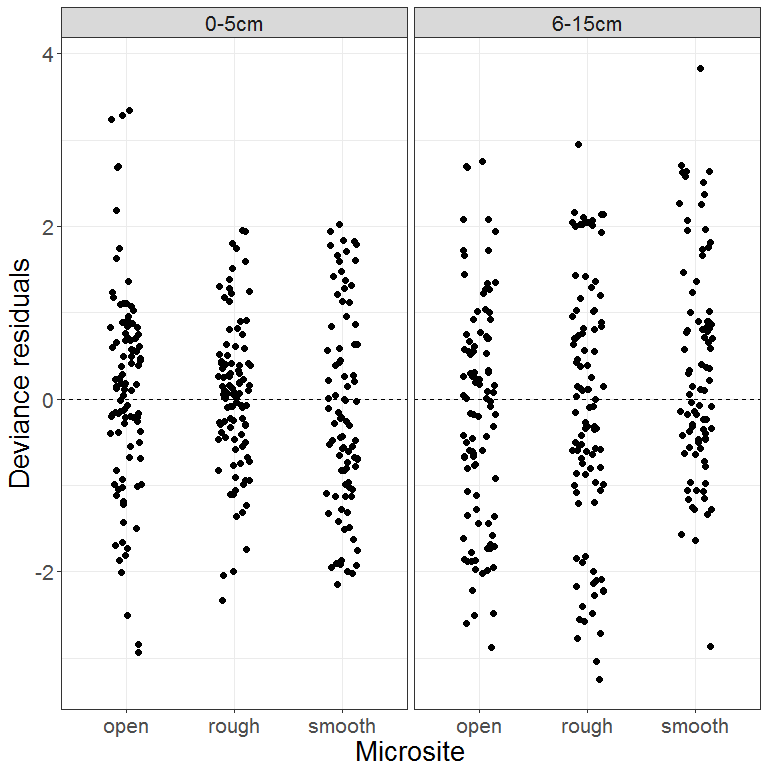
Graph of residuals against fitted values to check for homogeneity.



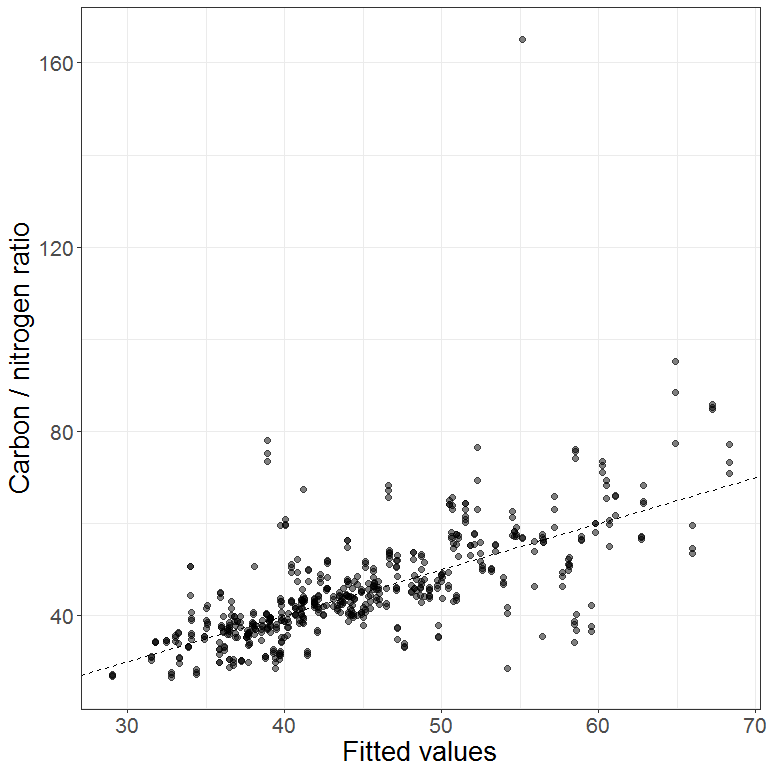
Residuals against co-variates.







Response against fitted values:



Investigating the outlier:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| siteLabel | siteId | fireType | depth | microsite | rep | percentC | percentN | cnratio |
| Red 15W | 4 | wildfire | 6-15cm | smooth | 1 | 2.028206 | 0.0122835 | 165.1168 |

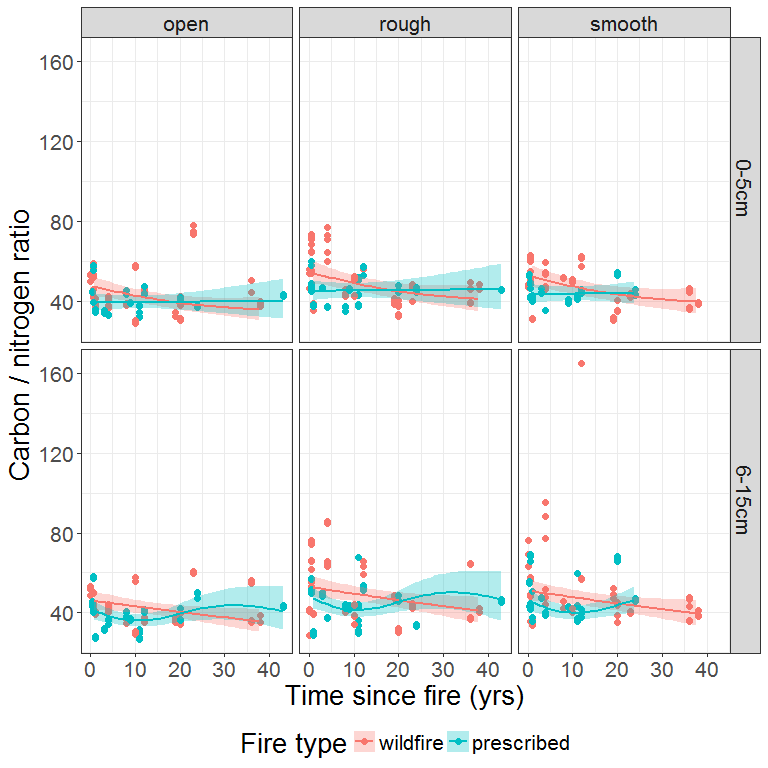
This is the same sample that showed up as an outlier in the total nitrogen model.

**Perhaps discard this sample and re-fit both models? Unlikely that the outlier is having much influence on the overall fit.**

### Model summary

##   
## Family: Scaled t(2.984,0.102)   
## Link function: identity   
##   
## Formula:  
## log(cnratio) ~ s(tsf.years, by = dft, k = 4) + fireType + depth +   
## microsite + s(fSiteId, bs = "re", by = flag.site)  
##   
## Parametric coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 3.755038 0.036789 102.069 < 2e-16 \*\*\*  
## fireTypeprescribed -0.078854 0.052985 -1.488 0.137   
## depth6-15cm -0.003712 0.010797 -0.344 0.731   
## micrositerough 0.138132 0.012601 10.962 < 2e-16 \*\*\*  
## micrositesmooth 0.102309 0.013228 7.734 1.04e-14 \*\*\*  
## ---  
## 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(tsf.years):dftwildfire.0-5cm 1.840 2.169 9.580 0.00921 \*\*   
## s(tsf.years):dftprescribed.0-5cm 1.000 1.000 0.014 0.90427   
## s(tsf.years):dftwildfire.6-15cm 1.000 1.000 5.750 0.01649 \*   
## s(tsf.years):dftprescribed.6-15cm 2.816 2.968 18.862 0.00121 \*\*   
## s(fSiteId):flag.site 27.715 29.000 680.252 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = 0.509 Deviance explained = 48.4%  
## -REML = -209.58 Scale est. = 1 n = 569

### Model predictions



### Predicted difference between fire types

Here we use posterior simulation to generate a set of 1000 credible mean trends from the fitted GAM and, for each trend, calculate the difference in carbon to nitrogen ratio between the two fire types.

