

Regressions and plots to determine number of states and state variables

Sanne Evers

4/11/2022

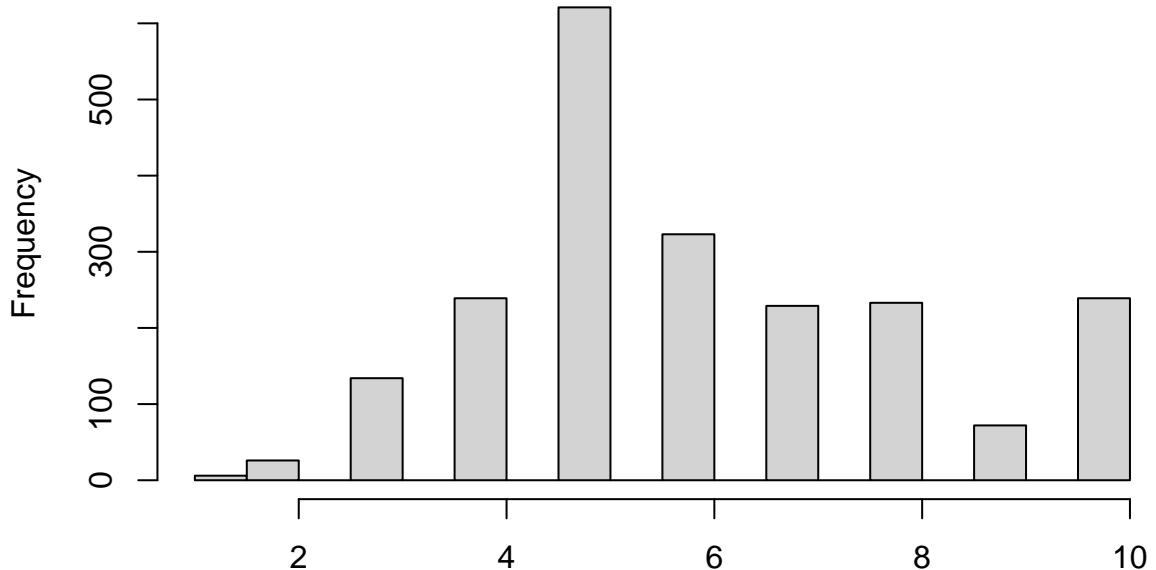
Seedling state; comparing seedlings with small vegetative individuals

Would it make sense to have a separate seedling stage? Seedlings are defined as *newly established plant from last census (or the census before) with longest stem no longer than 10 cm*. So one of the things that I need to know is if there's a difference between seedlings, and vegetative individuals that also only have one stem and are smaller than 10 cm.

First some summary about the seedlings in the data

```
##          [,1]
## n_stems    1.000000
## max_stems  1.000000
## mean_lenght 6.109331
## min_length  1.000000
## max_length 10.000000
```

Size distribution of seedling stem size



```
data$longest_stem_t0[which(data$stage_t0 == "sdl")]
```

Survival

We see that if we look at seedlings as a discrete state (i.e. no dependency on a continuous variable), that there is a significant difference in survival probability between seedlings and same sized vegetative individuals.

```
## # A tibble: 2 x 3
##   stage_t0  mean_length  surv_prob
##   <fct>        <dbl>      <dbl>
## 1 sdl          6.12      0.406
## 2 veg          7.69      0.505

##
## Welch Two Sample t-test
##
## data: survival_t1 by stage_t0
## t = -3.3207, df = 432.77, p-value = 0.000974
## alternative hypothesis: true difference in means between group sdl and group veg is not equal to 0
## 95 percent confidence interval:
## -0.15727347 -0.04032086
## sample estimates:
## mean in group sdl mean in group veg
##             0.4057900            0.5045872
```

However, it appears that survival for seedlings DOES depend on a continuous variable. With `log(longest_stem_t0)` resulting in a lower AIC than just `longest_stem_t0`:

```
##
## Call:
```

```

## glm(formula = survival_t1 ~ log(longest_stem_t0), family = "binomial",
##      data = data %>% filter(stage_t0 == "sdl"))
##
## Deviance Residuals:
##      Min      1Q Median      3Q      Max
## -1.290 -1.033 -0.849  1.249  2.268
##
## Coefficients:
##                               Estimate Std. Error z value Pr(>|z|)
## (Intercept)           -2.4915     0.2487 -10.020 <2e-16 ***
## log(longest_stem_t0)  1.1950     0.1371   8.715 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 2800.4  on 2072  degrees of freedom
## Residual deviance: 2718.5  on 2071  degrees of freedom
## (672 observations deleted due to missingness)
## AIC: 2722.5
##
## Number of Fisher Scoring iterations: 4

```

When we model survival we find a significant differences between seedlings and both vegetative and flowering individuals. However we see **NO** significant difference between vegetative and seedling when we only include vegetative individuals of the same size range as seedlings.

```

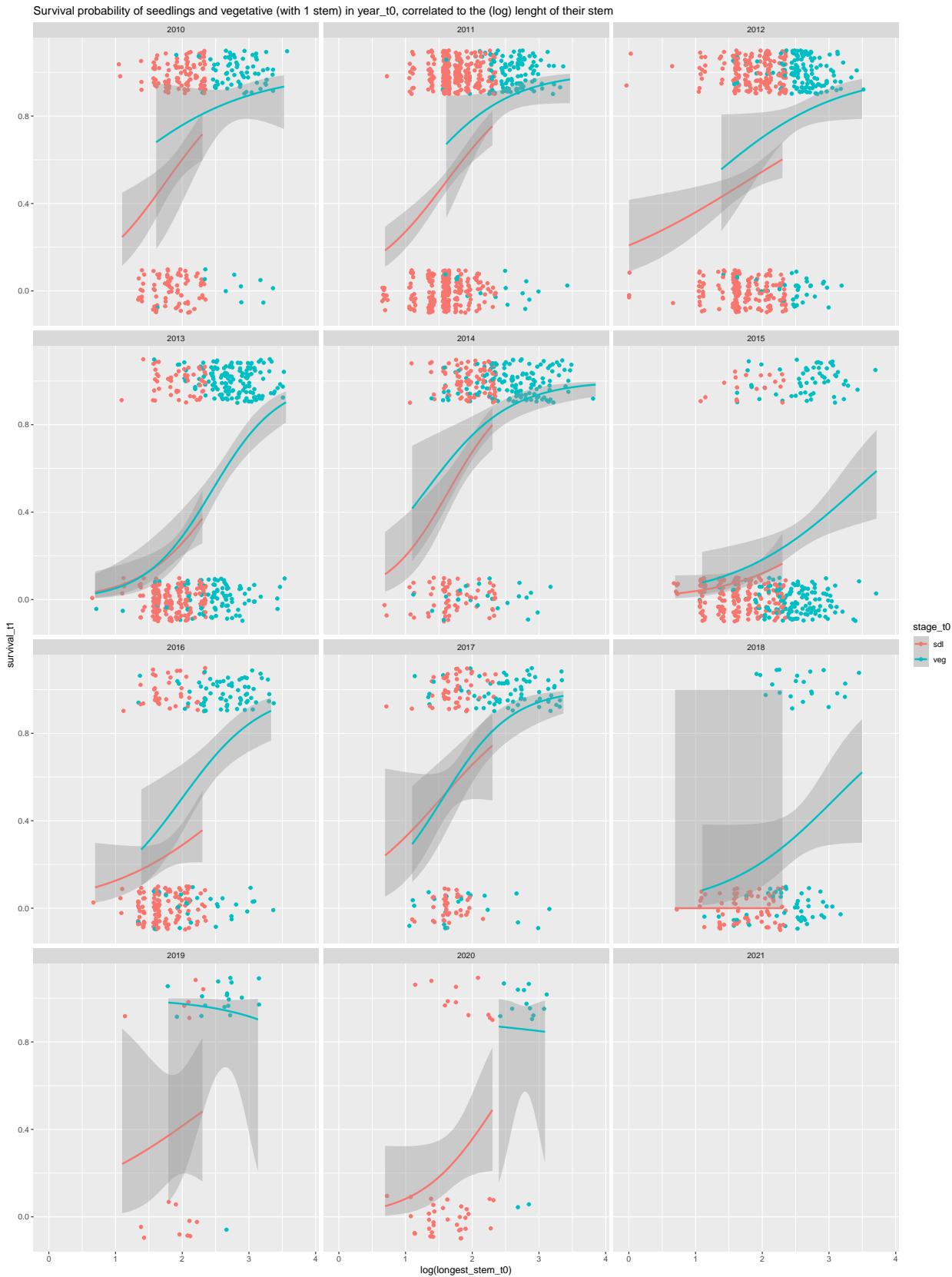
##
## Call:
## glm(formula = survival_t1 ~ log(longest_stem_t0) + stage_t0,
##      family = "binomial", data = data)
##
## Deviance Residuals:
##      Min      1Q Median      3Q      Max
## -2.2762 -1.0330  0.6638  0.8791  2.3220
##
## Coefficients:
##                               Estimate Std. Error z value Pr(>|z|)
## (Intercept)           -2.62594    0.15009 -17.496 < 2e-16 ***
## log(longest_stem_t0)  1.27038    0.08011  15.858 < 2e-16 ***
## stage_t0flow          -0.36643    0.13498  -2.715  0.00663 **
## stage_t0veg           0.22066    0.09197   2.399  0.01643 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 7446.7  on 5541  degrees of freedom
## Residual deviance: 6639.5  on 5538  degrees of freedom
## (48528 observations deleted due to missingness)
## AIC: 6647.5
##
## Number of Fisher Scoring iterations: 4
##

```

```

## Call:
## glm(formula = survival_t1 ~ log(longest_stem_t0) + stage_t0,
##      family = "binomial", data = data %>% filter(n_veg_stems_t0 ==
##          1 & n_fl_stems_t0 == 0))
##
## Deviance Residuals:
##       Min      1Q  Median      3Q     Max 
## -2.1301 -1.0312  0.5823  1.0626  2.2975 
##
## Coefficients:
##                               Estimate Std. Error z value Pr(>|z|)    
## (Intercept)           -2.56523   0.18389 -13.950 <2e-16 ***
## log(longest_stem_t0)  1.23399   0.09969  12.379 <2e-16 ***
## stage_t0veg          0.14216   0.10620   1.339   0.181    
## ---                
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 4576.8 on 3301 degrees of freedom
## Residual deviance: 4186.4 on 3299 degrees of freedom
## (1153 observations deleted due to missingness)
## AIC: 4192.4
##
## Number of Fisher Scoring iterations: 4
## `geom_smooth()` using formula 'y ~ x'
## Warning: Removed 100 rows containing non-finite values (stat_smooth).
## Warning: Removed 100 rows containing missing values (geom_point).

```



So to conclude. If we include longest_stem as state variable, we shouldn't have a seperate state for seedlings. However, if we do not use longest_stem as (part of) a state variable, Seedlings could be included

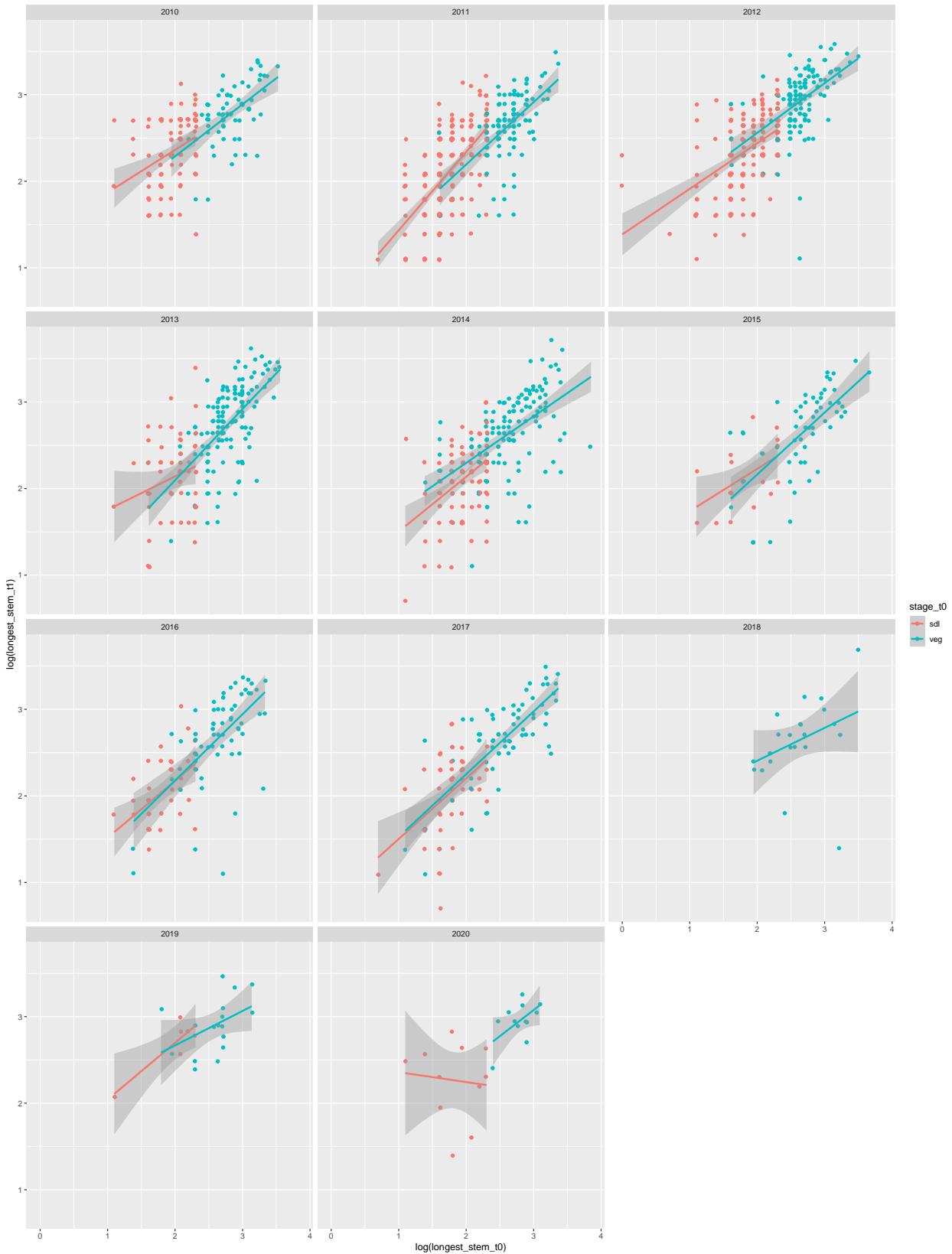
as a separate stage.

Growth

We see something similar when we look at growth - no significant difference if you compare seedlings only with vegetative individuals with n_stems_t0 == 1

```
##  
## Call:  
## lm(formula = log(longest_stem_t1) ~ log(longest_stem_t0) + stage_t0,  
##      data = data %>% filter(longest_stem_t0 > 0 & longest_stem_t1 >  
##          0))  
##  
## Residuals:  
##      Min        1Q    Median        3Q       Max  
## -2.42823 -0.17820  0.04067  0.22607  1.28799  
##  
## Coefficients:  
##                               Estimate Std. Error t value Pr(>|t|)  
## (Intercept)            1.01460   0.03406  29.791  <2e-16 ***  
## log(longest_stem_t0)  0.63942   0.01731  36.947  <2e-16 ***  
## stage_t0flow         -0.02501   0.02940  -0.851   0.395  
## stage_t0veg          0.02349   0.02165   1.085   0.278  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Residual standard error: 0.3562 on 3323 degrees of freedom  
## Multiple R-squared:  0.5388, Adjusted R-squared:  0.5384  
## F-statistic:  1294 on 3 and 3323 DF,  p-value: < 2.2e-16  
##  
## Call:  
## lm(formula = log(longest_stem_t1) ~ log(longest_stem_t0) + stage_t0,  
##      data = data %>% filter(longest_stem_t0 > 0 & longest_stem_t1 >  
##          0 & n_veg_stems_t0 == 1 & n_fl_stems_t0 == 0))  
##  
## Residuals:  
##      Min        1Q    Median        3Q       Max  
## -1.69108 -0.18733  0.03536  0.24269  1.28153  
##  
## Coefficients:  
##                               Estimate Std. Error t value Pr(>|t|)  
## (Intercept)            1.021053   0.046761  21.836  <2e-16 ***  
## log(longest_stem_t0)  0.636426   0.024454  26.025  <2e-16 ***  
## stage_t0veg          0.007749   0.027058   0.286   0.775  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Residual standard error: 0.3708 on 1667 degrees of freedom  
## Multiple R-squared:  0.4787, Adjusted R-squared:  0.4781  
## F-statistic: 765.3 on 2 and 1667 DF,  p-value: < 2.2e-16  
## `geom_smooth()` using formula 'y ~ x'
```

Growth of seedlings and vegetative (with 1 stem) in year_t0, correlated to the (log) lenght of their stem



Vegetative and Flowering stage; are there significant differences between the two?

First thing to check: Do vegetative individuals “turn” flowering once, and stay as flowering?? - Doesn’t look like it.

```
## `summarise()` has grouped output by 'stage_t0'. You can override using the
## `.groups` argument.

## # A tibble: 4 x 3
## # Groups:   stage_t0 [2]
##   stage_t0 stage_t1 individuals
##   <fct>    <chr>     <int>
## 1 flow      flow       1876
## 2 flow      veg        888
## 3 veg       flow       888
## 4 veg       veg        2191
```

Survival

There is a significant difference in survival rate between flowering and vegetative individuals, although the plots don’t look to bad

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: survival_t1 ~ log(n_stems_t0) + stage_t0 + population + (1 |
##   year_t0)
## Data: veg_flow %>% filter(n_stems_t0 > 0)
##
##          AIC      BIC      logLik deviance df.resid
##  4556.2   4604.2   -2271.1    4542.2     7010
##
## Scaled residuals:
##    Min     1Q Median     3Q    Max
## -8.1650  0.1351  0.2245  0.3589  4.2194
##
## Random effects:
##   Groups   Name        Variance Std.Dev.
##   year_t0 (Intercept) 1.473     1.214
##   Number of obs: 7017, groups: year_t0, 18
##
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) 1.78917   0.31148  5.744 9.25e-09 ***
## log(n_stems_t0) 0.62103   0.05724 10.850 < 2e-16 ***
## stage_t0veg -0.40631   0.09900 -4.104 4.06e-05 ***
## populationHK 0.11548   0.10419  1.108   0.268
## populationKS -0.79137   0.10199 -7.760 8.52e-15 ***
## populationRU 0.73505   0.13997  5.251 1.51e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) stg_t0 ppltHK ppltKS
```

```

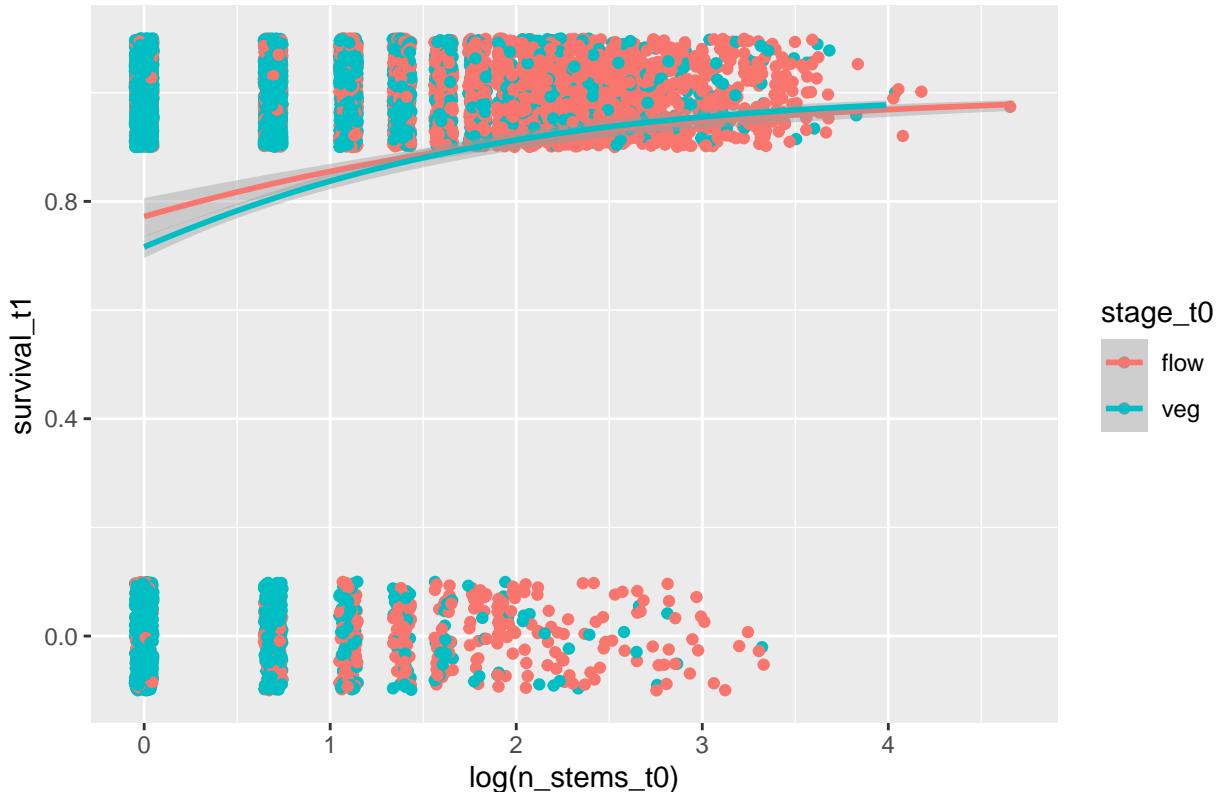
## lg(n_stm_0) -0.226
## stage_t0veg -0.298  0.401
## populatinHK -0.152  0.022  0.137
## populatinKS -0.149 -0.059  0.088  0.374
## populatinRU -0.095  0.093 -0.015  0.232  0.223
## `geom_smooth()` using formula 'y ~ x'

## Warning: Removed 2620 rows containing non-finite values (stat_smooth).

## Warning: Removed 2620 rows containing missing values (geom_point).

```

Survival of vegetative and flowering individuals in year_t0



```

## `geom_smooth()` using formula 'y ~ x'
## Warning: Removed 2620 rows containing non-finite values (stat_smooth).

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: Removed 2620 rows containing missing values (geom_point).

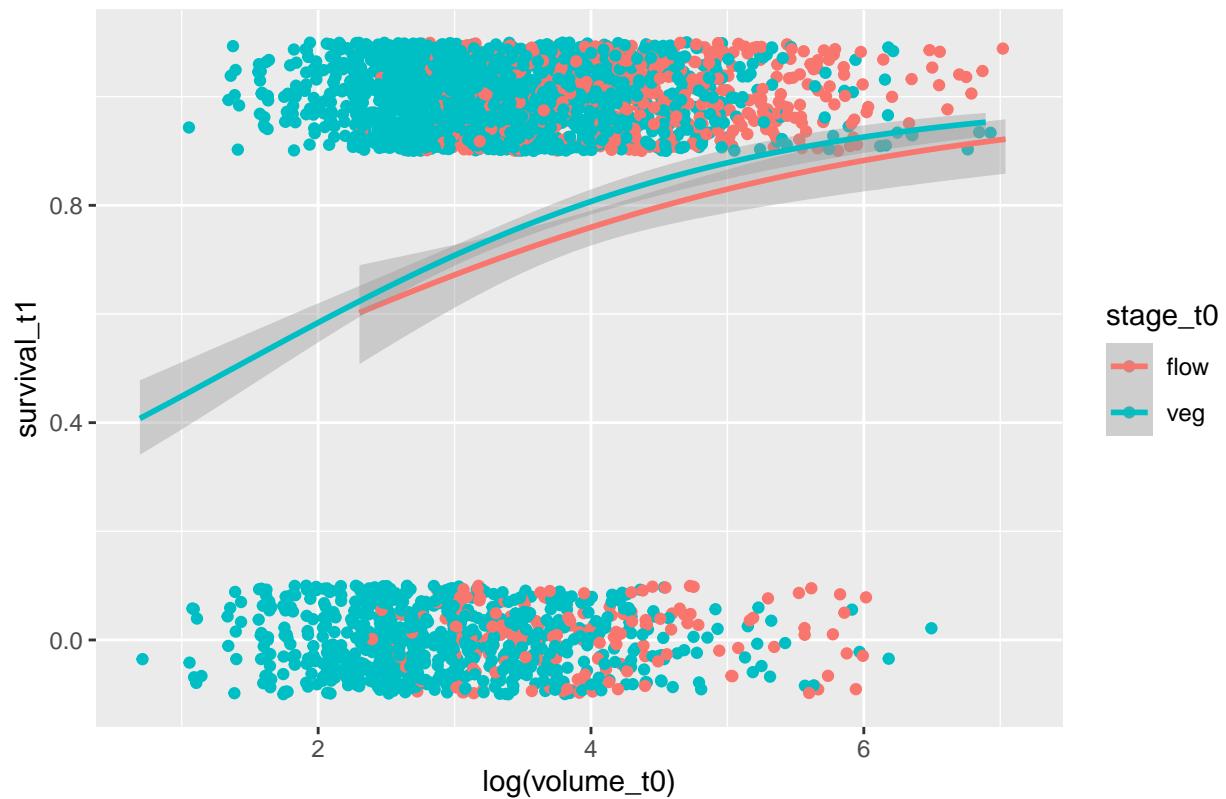
```



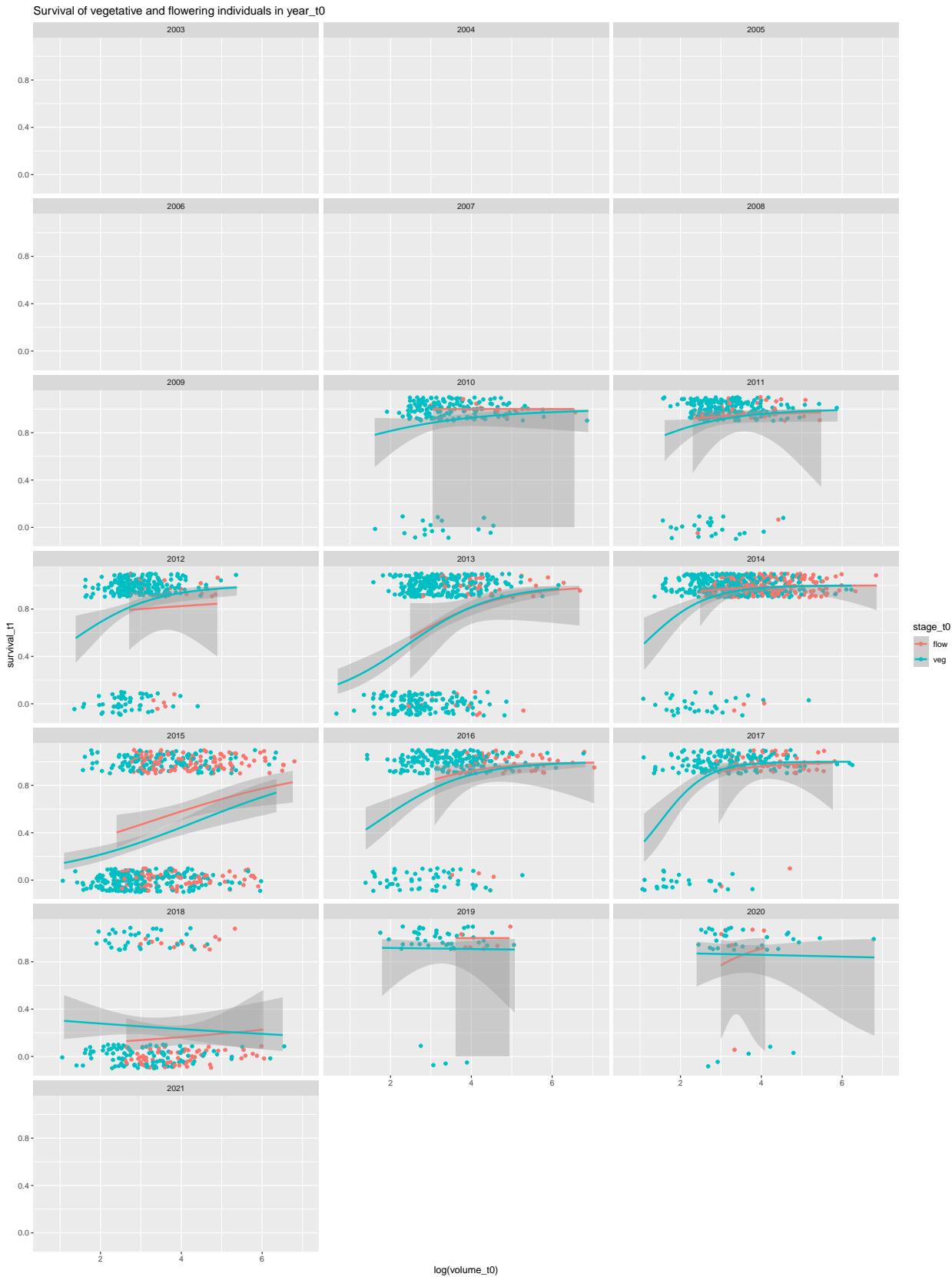
Here are the same graphs, but with volume (`n_stems * longest_stem`) as state variable

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: survival_t1 ~ log(volume_t0) + stage_t0 + population + (1 | year_t0)
## Data: veg_flow %>% filter(volume_t0 > 0)
##
##          AIC      BIC    logLik deviance df.resid
## 2665.0   2707.3  -1325.5   2651.0      3085
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -8.9242 -0.4234  0.2697  0.4712  4.3488
##
## Random effects:
## Groups   Name        Variance Std.Dev.
## year_t0 (Intercept) 1.747     1.322
## Number of obs: 3092, groups: year_t0, 11
##
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.81790   0.49007 -1.669   0.09513 .
## log(volume_t0) 0.69675   0.06184 11.267 < 2e-16 ***
## stage_t0veg -0.25783   0.14382 -1.793   0.07301 .
## populationHK 0.43967   0.13426  3.275   0.00106 **
## populationKS -0.37255   0.14427 -2.582   0.00981 **
## populationRU 0.98013   0.16409  5.973 2.33e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) lg(_0) stg_t0 ppltHK ppltKS
## log(vlm_t0) -0.496
## stage_t0veg -0.400  0.350
## populatinHK -0.168  0.147  0.164
## populatinKS -0.115  0.028  0.148  0.239
## populatinRU -0.091  0.083 -0.038  0.208  0.175
##
## `geom_smooth()` using formula 'y ~ x'
##
## Warning: Removed 6545 rows containing non-finite values (stat_smooth).
## Warning: Removed 6545 rows containing missing values (geom_point).
```

Survival of vegetative and flowering individuals in year_t0



```
## `geom_smooth()` using formula 'y ~ x'  
## Warning: Removed 6545 rows containing non-finite values (stat_smooth).  
## Warning: Removed 6545 rows containing missing values (geom_point).
```

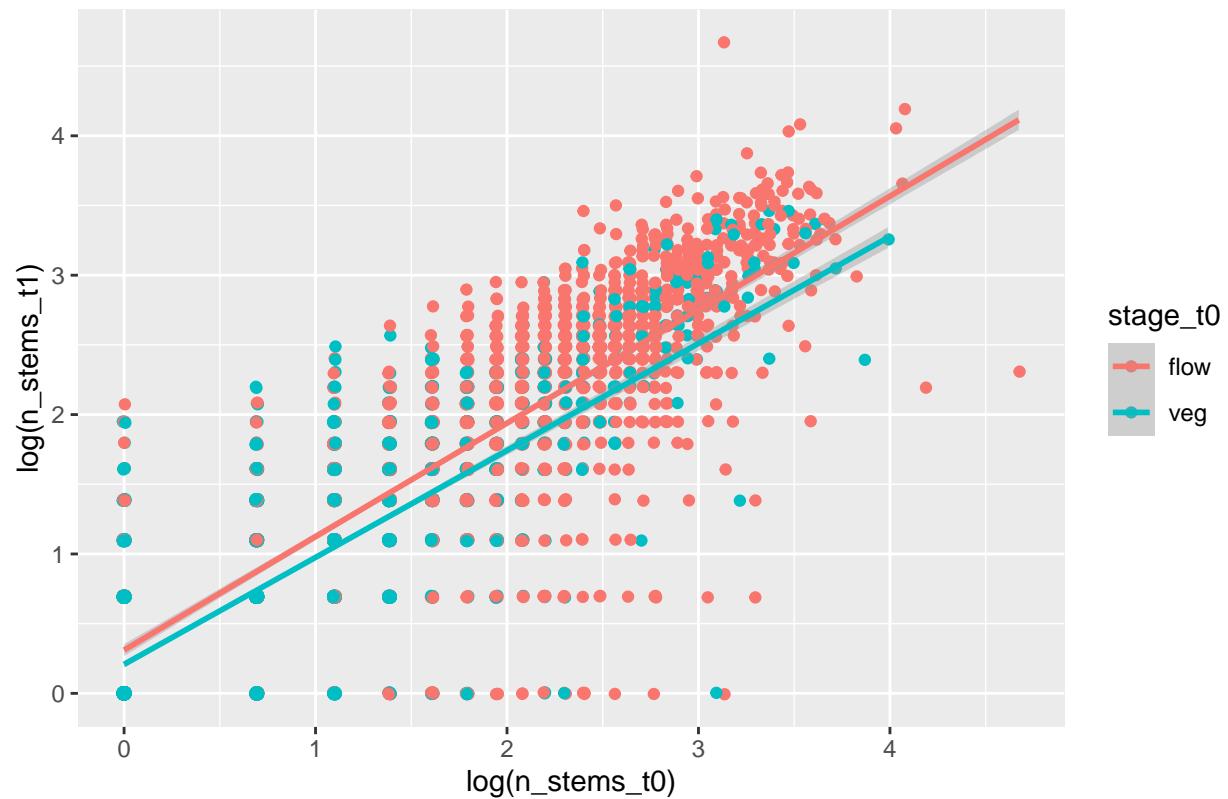


Growth

When using both n_stems and volume (n_stems * longest_stem) give a significant difference between vegetative and flowering individuals. When looking at the graphs, the difference does seem to be minimal in most years.

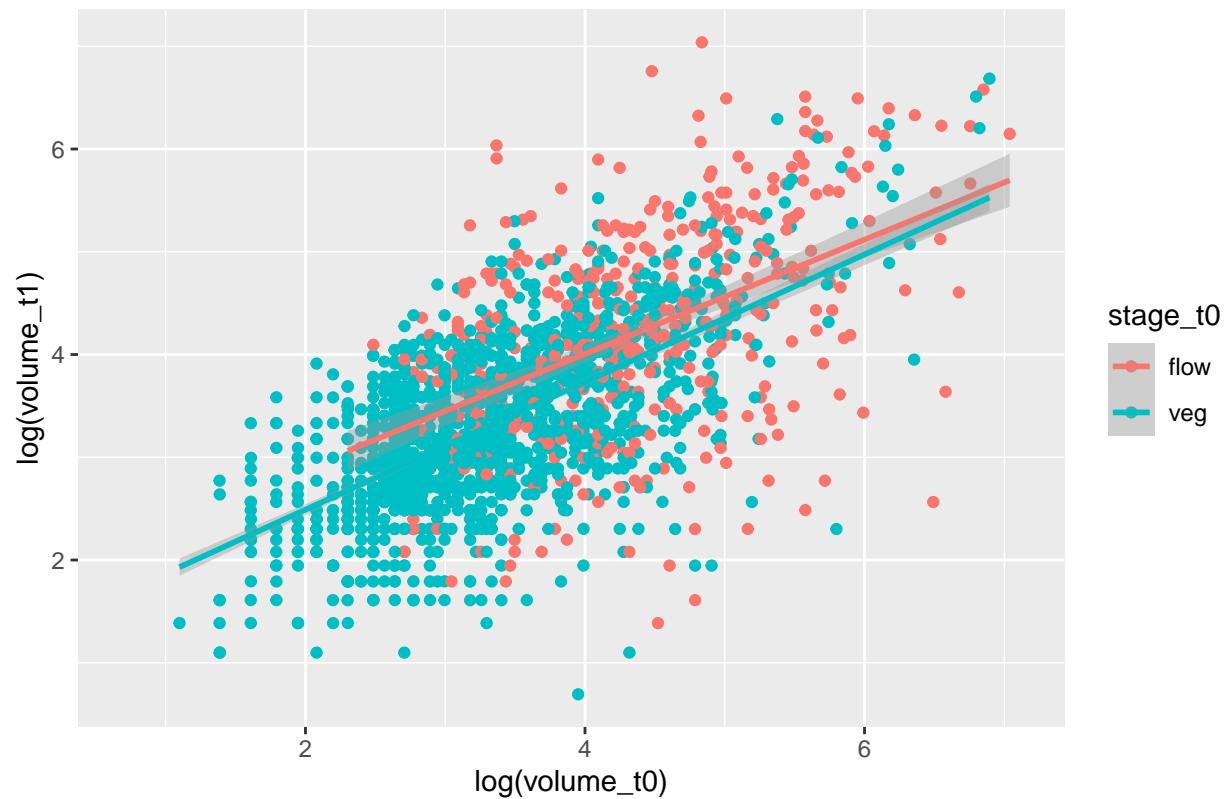
```
## Linear mixed model fit by REML ['lmerMod']
## Formula: log(volume_t1) ~ log(volume_t0) + stage_t0 + population + (1 |
##   year_t0)
##   Data: veg_flow %>% filter(volume_t0 > 0 & volume_t1 > 0)
##
## REML criterion at convergence: 4137.2
##
## Scaled residuals:
##   Min     1Q Median     3Q    Max
## -4.7859 -0.5704 -0.0054  0.6704  3.8902
##
## Random effects:
## Groups   Name        Variance Std.Dev.
## year_t0 (Intercept) 0.01447  0.1203
## Residual           0.44494  0.6670
## Number of obs: 2019, groups: year_t0, 11
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept) 1.67547   0.09531 17.579
## log(volume_t0) 0.61147   0.01792 34.127
## stage_t0veg -0.31791   0.04220 -7.534
## populationHK -0.10878   0.04149 -2.622
## populationKS -0.32313   0.04759 -6.790
## populationRU -0.06616   0.04094 -1.616
##
## Correlation of Fixed Effects:
##          (Intr) lg(_0) stg_t0 pppltHK pppltKS
## log(vlm_t0) -0.825
## stage_t0veg -0.652  0.435
## populatinHK -0.272  0.144  0.172
## populatinKS -0.152  0.018  0.116  0.233
## populatinRU -0.156  0.050 -0.050  0.257  0.224
##
## `geom_smooth()` using formula 'y ~ x'
##
## Warning: Removed 3806 rows containing non-finite values (stat_smooth).
## Warning: Removed 3806 rows containing missing values (geom_point).
```

Growth of vegetative and flowering individuals from year_t0 to year_t1



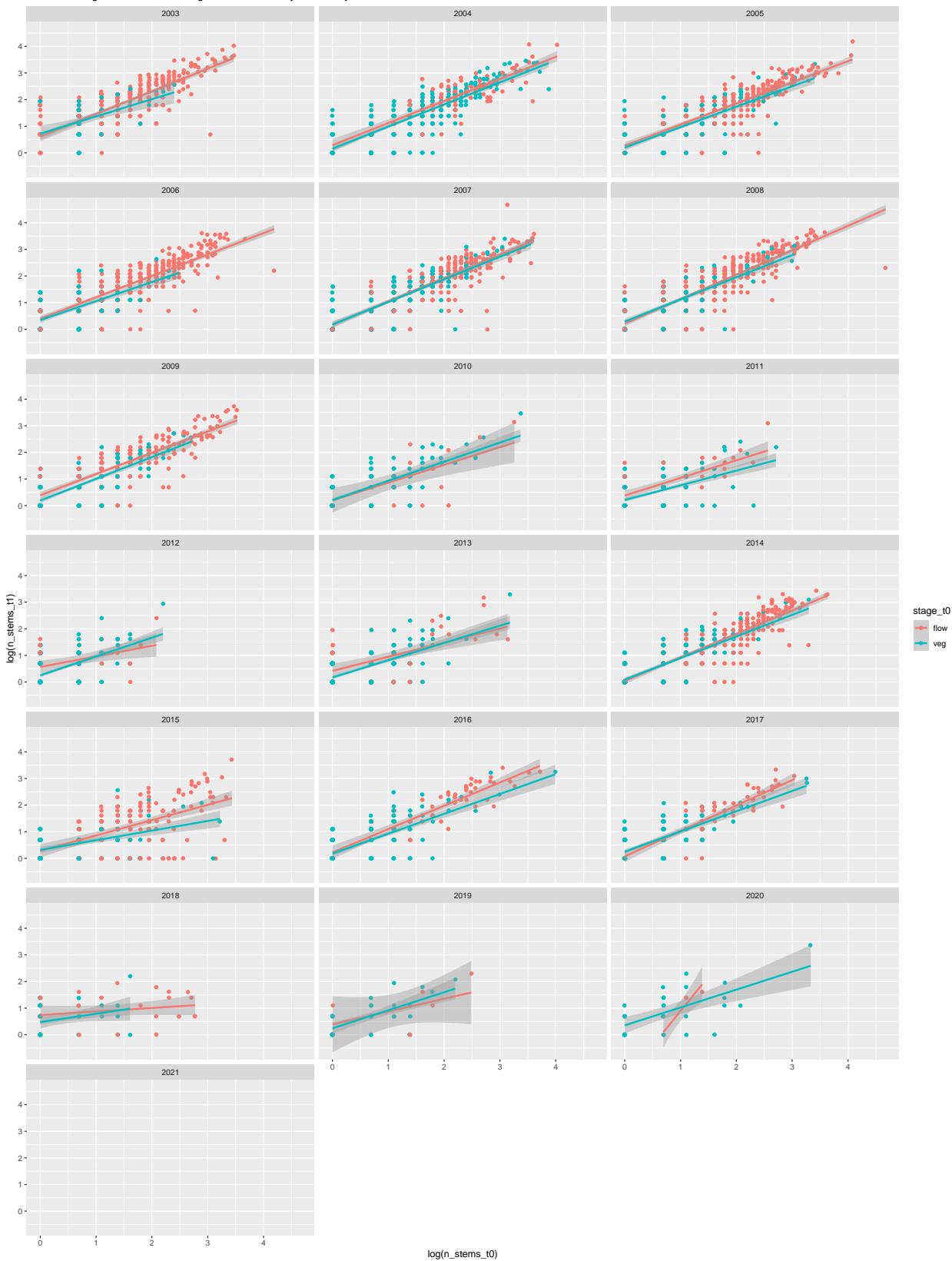
```
## `geom_smooth()` using formula 'y ~ x'
## Warning: Removed 7618 rows containing non-finite values (stat_smooth).
## Warning: Removed 7618 rows containing missing values (geom_point).
```

Growth of vegetative and flowering individuals from year_t0 to year_t1



```
## `geom_smooth()` using formula 'y ~ x'  
## Warning: Removed 3806 rows containing non-finite values (stat_smooth).  
## Warning: Removed 3806 rows containing missing values (geom_point).
```

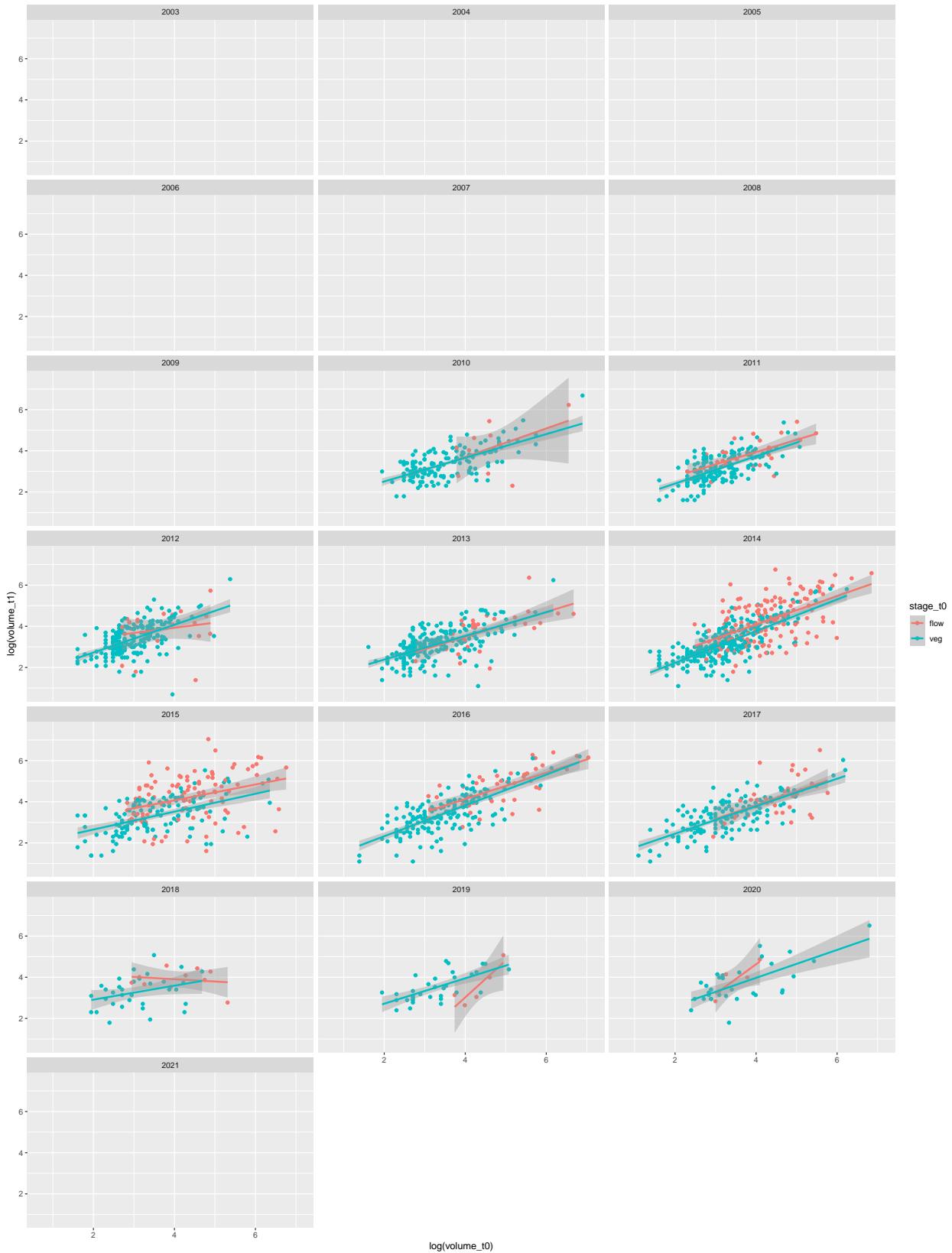
Growth of vegetative and flowering individuals from year_t0 to year_t1



```
## `geom_smooth()` using formula 'y ~ x'
```

```
## Warning: Removed 7618 rows containing non-finite values (stat_smooth).  
## Warning: Removed 7618 rows containing missing values (geom_point).
```

Growth of vegetative and flowering individuals from year_t0 to year_t1



Flower probability

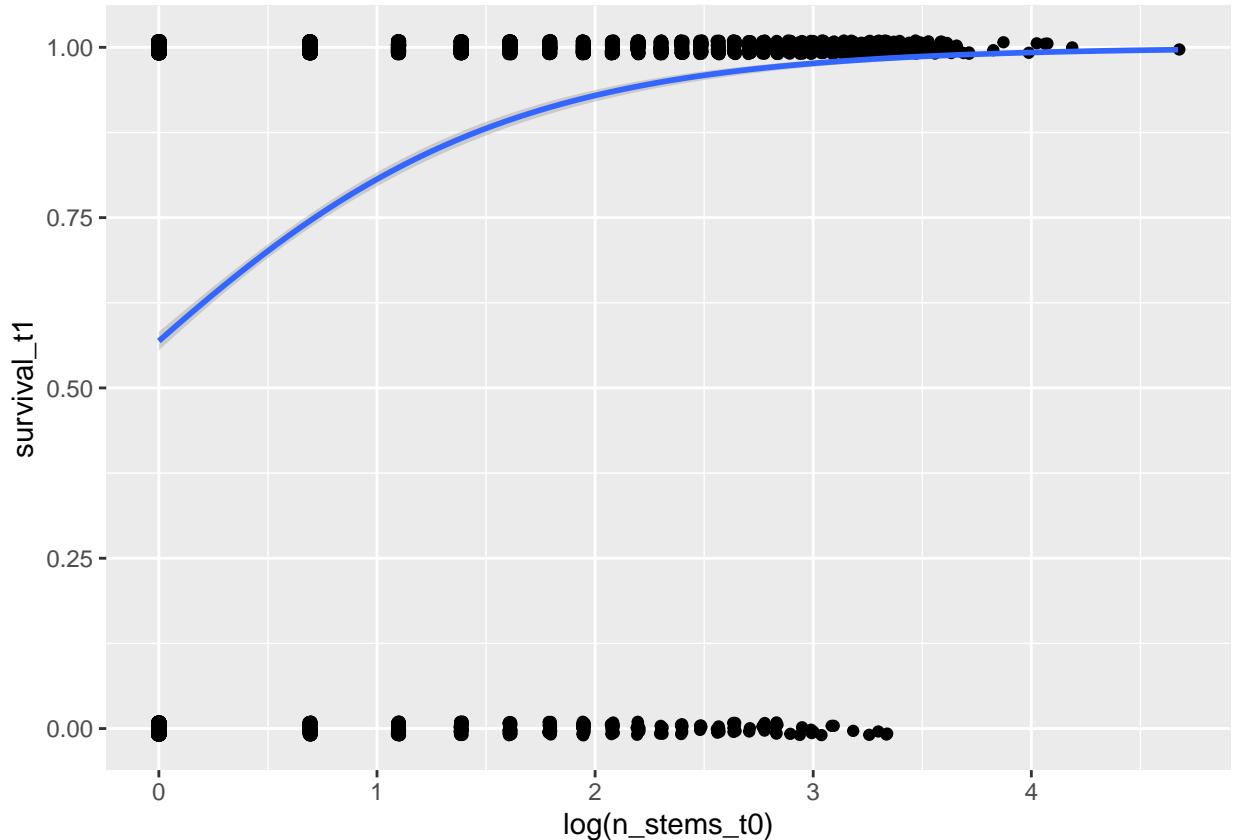
Is the chance larger to flower in t_1 if the individual was flowering in t_0?

State dependencies

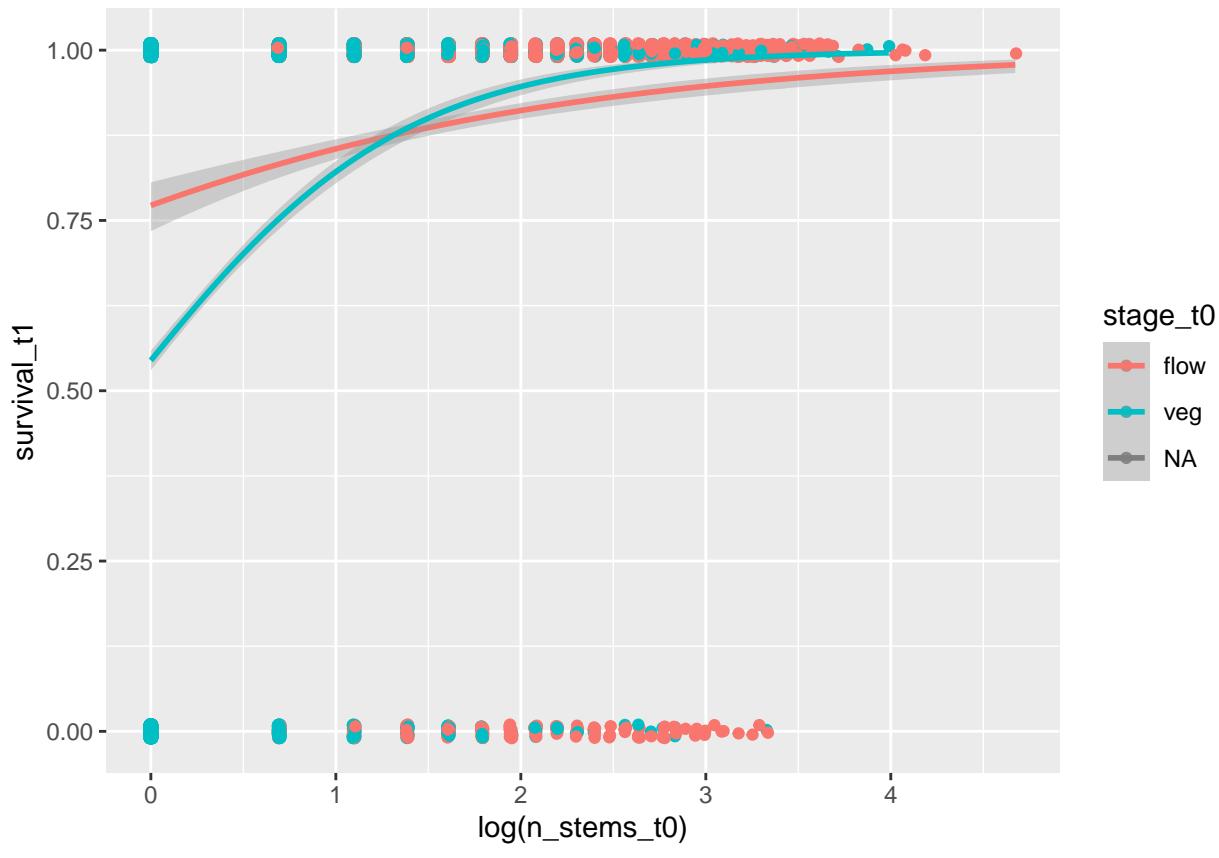
Survival

State variable

```
## `geom_smooth()` using formula 'y ~ x'  
## Warning: Removed 44392 rows containing non-finite values (stat_smooth).  
## Warning: Removed 44392 rows containing missing values (geom_point).
```



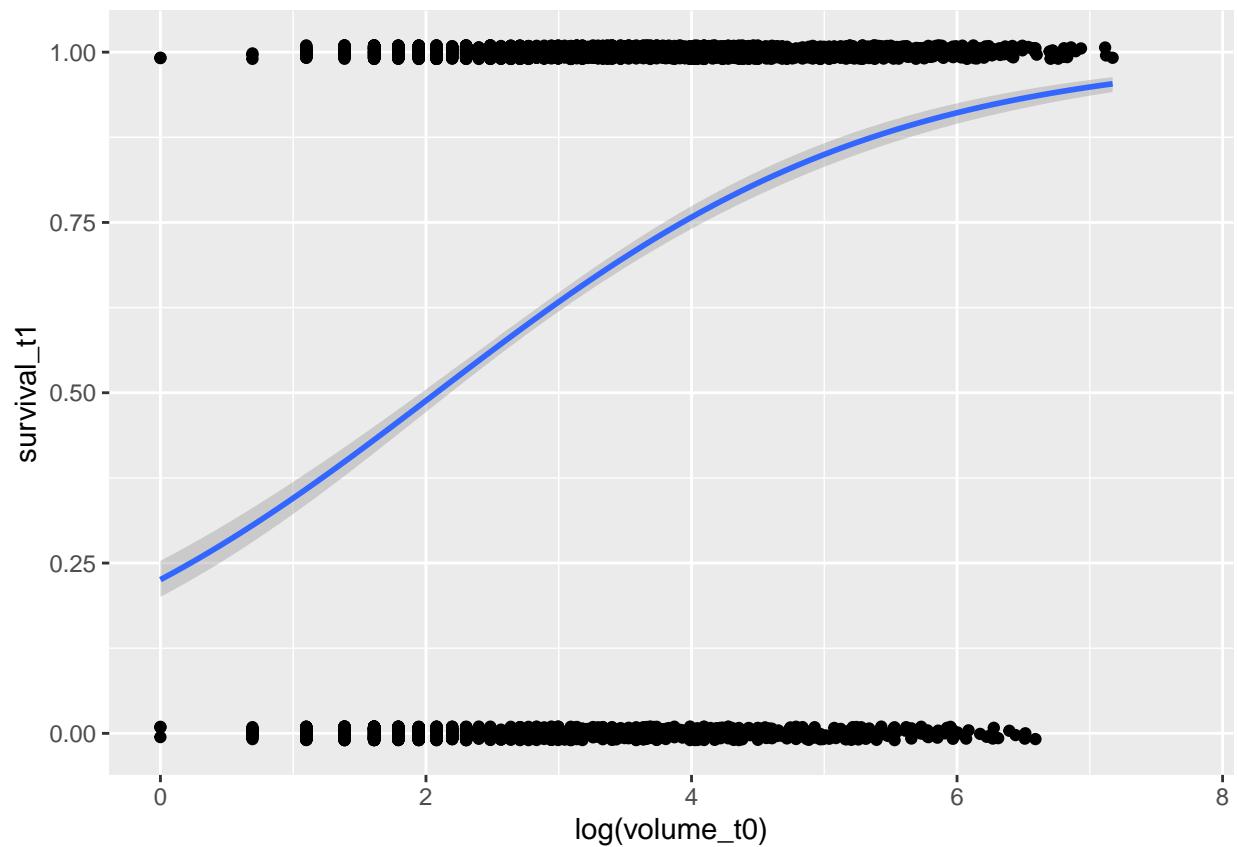
```
## `geom_smooth()` using formula 'y ~ x'  
## Warning: Removed 44392 rows containing non-finite values (stat_smooth).  
## Removed 44392 rows containing missing values (geom_point).
```



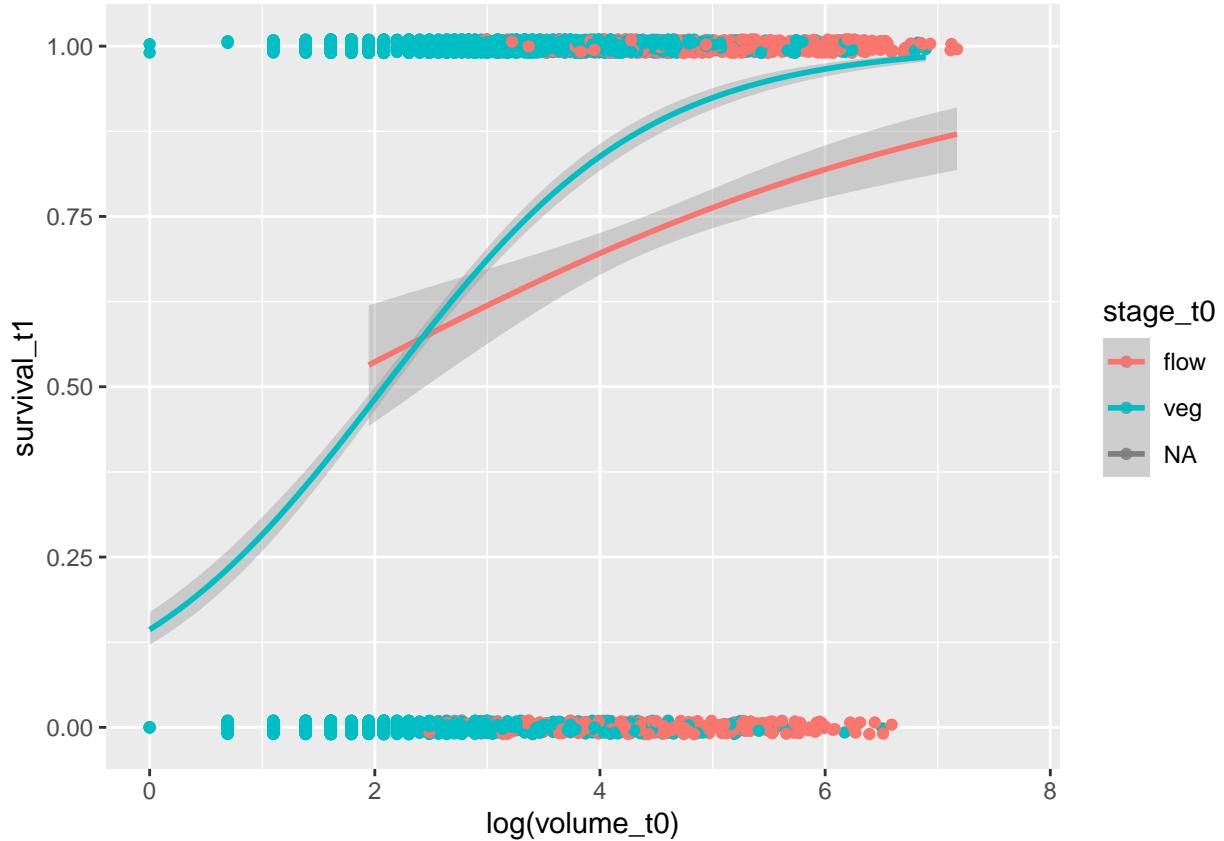
```

## `geom_smooth()` using formula 'y ~ x'
## Warning: Removed 48569 rows containing non-finite values (stat_smooth).
## Warning: Removed 48569 rows containing missing values (geom_point).

```



```
## `geom_smooth()` using formula 'y ~ x'  
## Warning: Removed 48569 rows containing non-finite values (stat_smooth).  
## Removed 48569 rows containing missing values (geom_point).
```



```

##      AIC      dAIC      df weight
## s2  6802.7      0.0 2    1
## s1 10160.7  3358.0 2 <0.001

##
## Call:
## glm(formula = survival_t1 ~ log(volume_t0), family = "binomial",
##      data = data1 %>% filter(stage_t0 != "dead"))
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3424  -1.1060   0.6532   1.0235   1.7256
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.23326   0.07708 -16.00  <2e-16 ***
## log(volume_t0) 0.59330   0.02686  22.09  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 7388.5 on 5500 degrees of freedom
## Residual deviance: 6798.7 on 5499 degrees of freedom
## (6881 observations deleted due to missingness)
## AIC: 6802.7
##

```

```

## Number of Fisher Scoring iterations: 3

Base model selection

Below are the models tested and the summary of the model with the lowest AICc

## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly uniden
## - Rescale variables?

##          2.5 %      97.5 %
## .sig01        NA        NA
## (Intercept) -0.824278597  0.420687407
## log(n_stems_t0) 1.097668686  1.297950516
## populationHK -0.168353238  0.129197147
## populationKS -0.537060639 -0.180931492
## populationRU  0.605674485  1.040037541
## herb_shading_t0  0.087544536  0.203584597
## shrub_shading_t0  0.020929848  0.104422217
## soil_d1       -0.037384429  0.004194910
## soil_d2       -0.016930439  0.025767027
## soil_d3       -0.011953890  0.031312778
## rock          -0.002964150  0.003297016
## slope         0.005075586  0.013467227

## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: survival_t1 ~ log(n_stems_t0) + population + herb_shading_t0 +
##           shrub_shading_t0 + slope + (1 | year_t0)
## Data: data1 %>% filter(stage_t0 != "dead")
##
##      AIC      BIC      logLik deviance df.resid
##    7558.9   7621.8   -3770.4    7540.9     8020
##
## Scaled residuals:
##      Min      1Q      Median      3Q      Max
## -9.2933 -0.5222  0.2673  0.6090  4.8498
##
## Random effects:
## Groups   Name      Variance Std.Dev.
## year_t0 (Intercept) 1.276    1.13
## Number of obs: 8029, groups: year_t0, 15
##
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.196163  0.302282 -0.649  0.51638
## log(n_stems_t0) 1.197583  0.050733 23.606 < 2e-16 ***
## populationHK -0.019739  0.074869 -0.264  0.79205
## populationKS -0.369590  0.090525 -4.083 4.45e-05 ***
## populationRU  0.818495  0.110287  7.421 1.16e-13 ***
## herb_shading_t0  0.141775  0.027151  5.222 1.77e-07 ***
## shrub_shading_t0  0.062086  0.021159  2.934  0.00334 **
## slope         0.009334  0.001993  4.684 2.82e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##

```

```

## Correlation of Fixed Effects:
##          (Intr) l(_0) ppltHK ppltKS ppltRU hrb__0 shr__0
## lg(n_stm_0) -0.079
## populatinHK -0.044 -0.040
## populatinKS -0.046 -0.149  0.255
## populatinRU -0.018  0.047  0.203  0.233
## hrb_shdng_0 -0.160  0.043 -0.057 -0.185 -0.316
## shrb_shdn_0 -0.018 -0.006 -0.216  0.027  0.071 -0.160
## slope       -0.133  0.098 -0.176  0.116  0.023  0.185  0.054

```

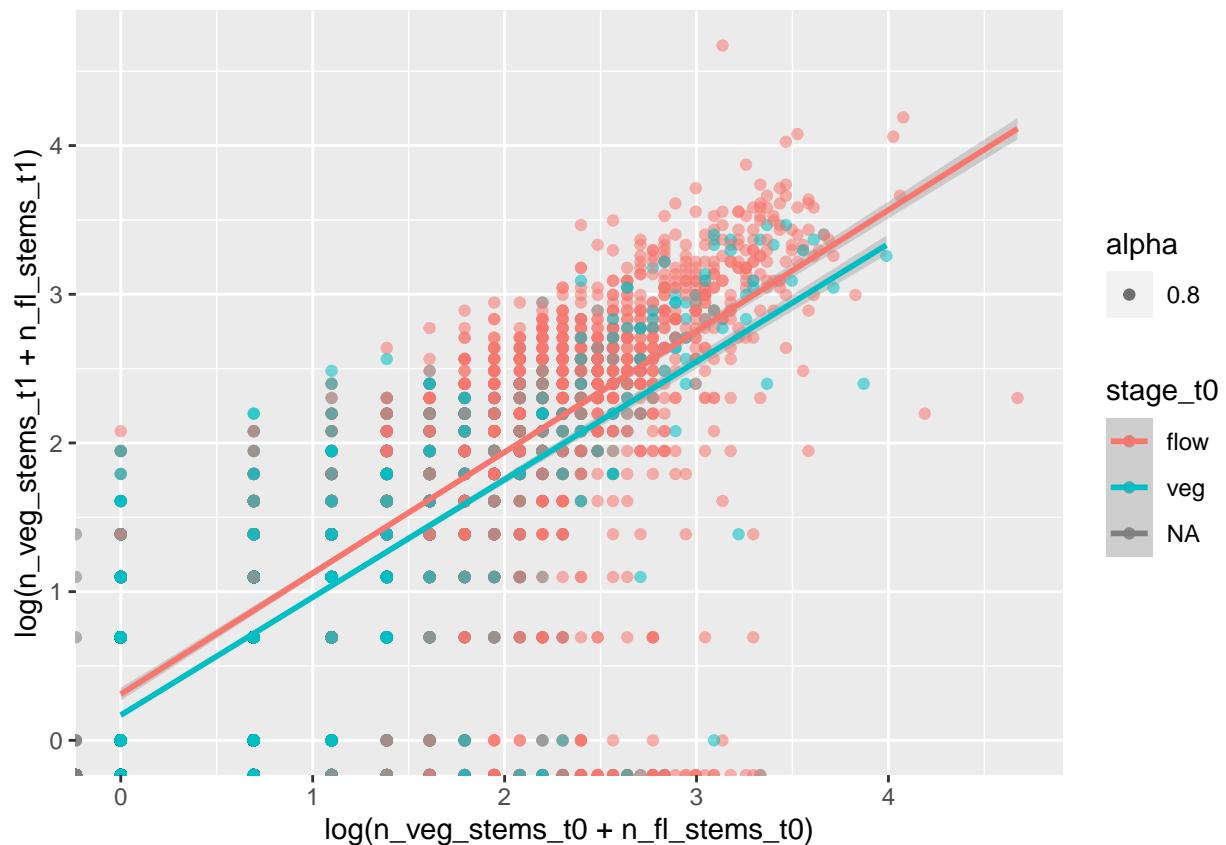
Growth

State variable

```

## `geom_smooth()` using formula 'y ~ x'
## Warning: Removed 47117 rows containing non-finite values (stat_smooth).
## Warning: Removed 41822 rows containing missing values (geom_point).

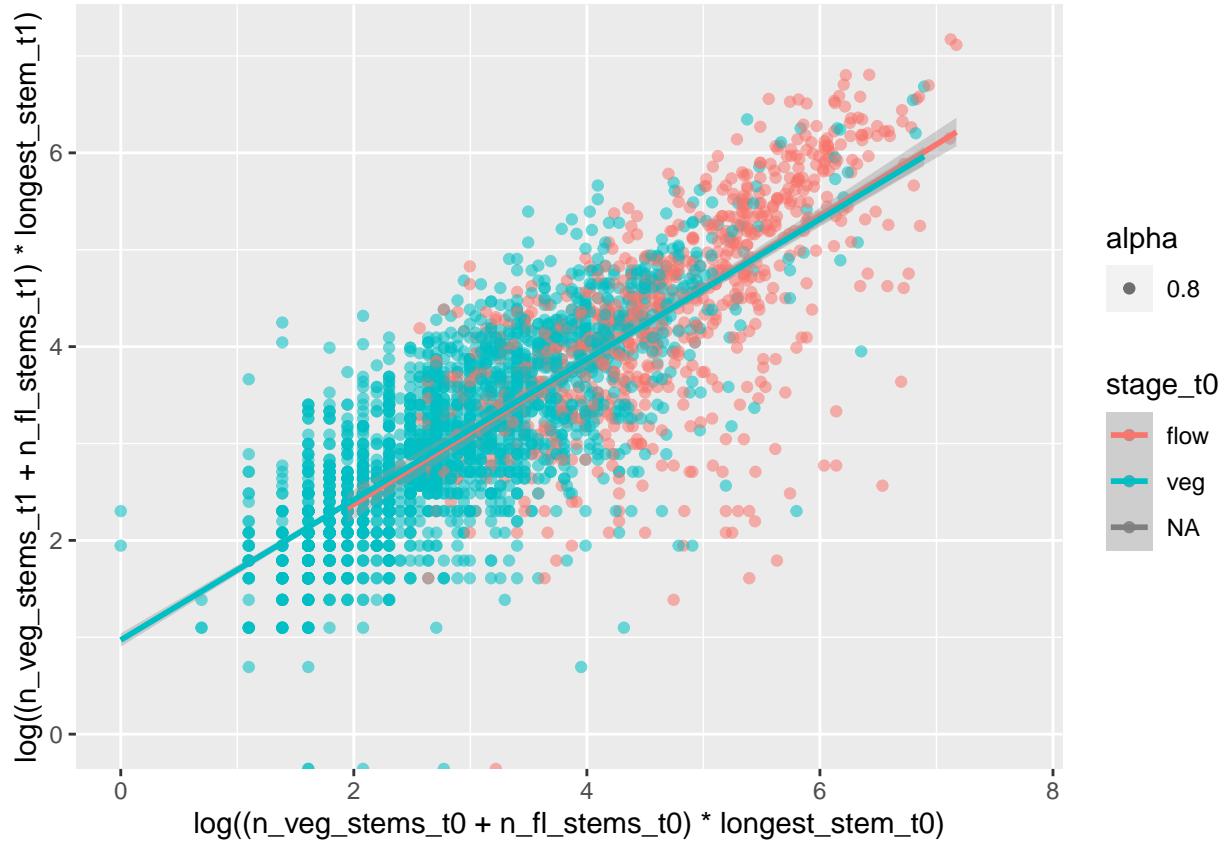
```



```

## `geom_smooth()` using formula 'y ~ x'
## Warning: Removed 50776 rows containing non-finite values (stat_smooth).
## Warning: Removed 50771 rows containing missing values (geom_point).

```



```

##      AIC      dAIC      df weight
## g2  6327.0      0.0 3    1
## g1 11786.2  5459.2 3 <0.001

##
## Call:
## lm(formula = log(volume_t1) ~ log(volume_t0), data = data1 %>%
##     filter(volume_t1 > 0 & volume_t0 > 0))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.2762 -0.3337  0.0110  0.4007  2.2790
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.960609  0.030339  31.66  <2e-16 ***
## log(volume_t0) 0.727722  0.008749  83.18  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6318 on 3292 degrees of freedom
## Multiple R-squared:  0.6776, Adjusted R-squared:  0.6775
## F-statistic: 6918 on 1 and 3292 DF,  p-value: < 2.2e-16

```

Base model selection

Below are the models tested and the summary of the model with the lowest AICc

```

##          2.5 %      97.5 %
## .sig01        NA        NA
## .sigma        NA        NA
## (Intercept) 0.0659765219 0.2453382088
## log(n_stems_t0) 0.7993138453 0.8320354315
## populationHK -0.0145596544 0.0551419294
## populationKS -0.0449152294 0.0348377690
## populationRU -0.0513390579 0.0414064102
## herb_shading_t0 -0.0359233611 -0.0091432651
## shrub_shading_t0 -0.0054637348 0.0126408847
## soil_d1       -0.0023910288 0.0071746705
## soil_d2       -0.0003729416 0.0095089599
## soil_d3       -0.0034972045 0.0064432675
## slope         -0.0009696693 0.0009085533
## rock          -0.0009711372 0.0005296929

## Linear mixed model fit by REML ['lmerMod']
## Formula: log(n_stems_t1) ~ log(n_stems_t0) + herb_shading_t0 + (1 | year_t0)
##   Data: data1 %>% filter(n_stems_t1 > 0 & n_stems_t0 > 0)
##
## REML criterion at convergence: 8532.5
##
## Scaled residuals:
##    Min     1Q Median     3Q    Max
## -5.7278 -0.3709 -0.1042  0.6852  3.7588
##
## Random effects:
##   Groups   Name      Variance Std.Dev.
##   year_t0 (Intercept) 0.01758  0.1326
##   Residual           0.25804  0.5080
## Number of obs: 5709, groups: year_t0, 15
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept) 0.174459  0.038009  4.590
## log(n_stems_t0) 0.821908  0.008064 101.928
## herb_shading_t0 -0.016407  0.005580 -2.941
##
## Correlation of Fixed Effects:
##            (Intr) l(_0)
## lg(n_stm_0) -0.199
## hrb_shdng_0 -0.317  0.058

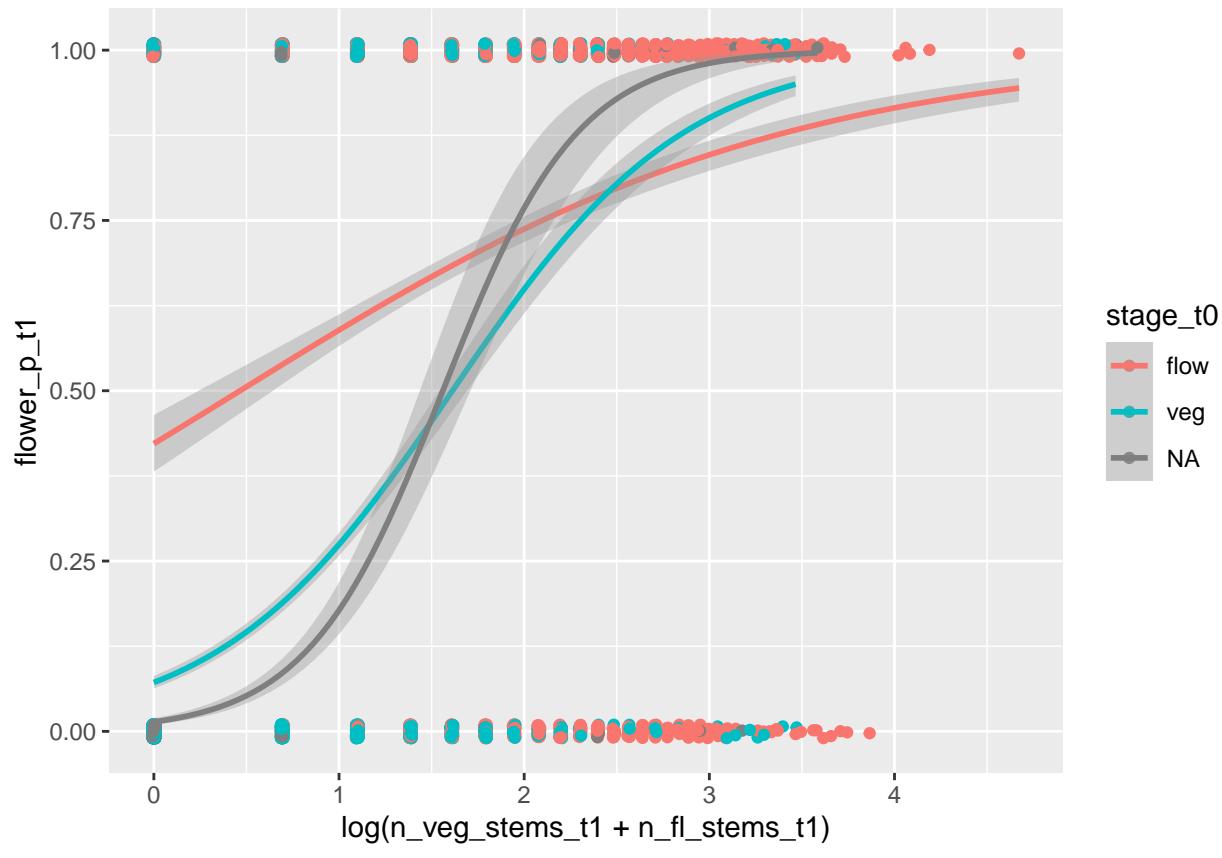
```

Flower probability

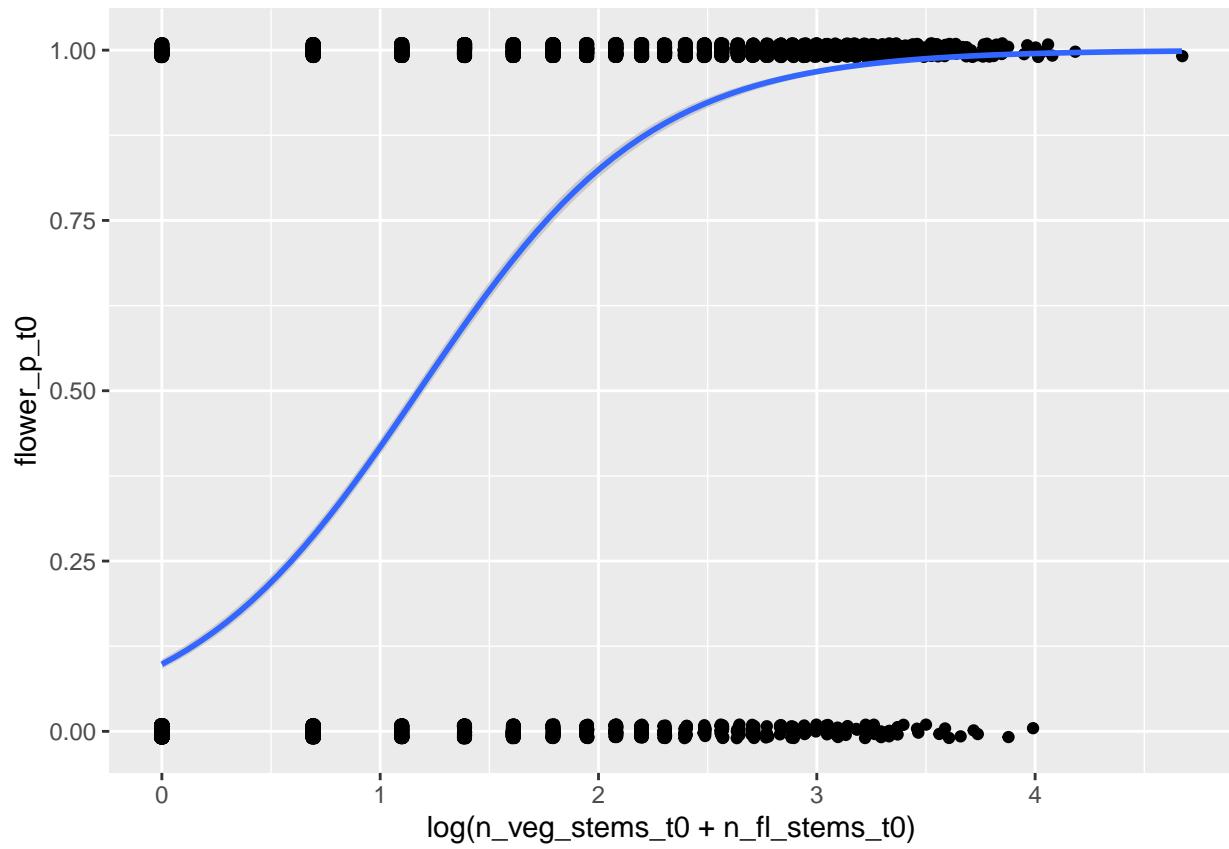
```

## `geom_smooth()` using formula 'y ~ x'
## Warning: Removed 44683 rows containing non-finite values (stat_smooth).
## Warning: Removed 44683 rows containing missing values (geom_point).

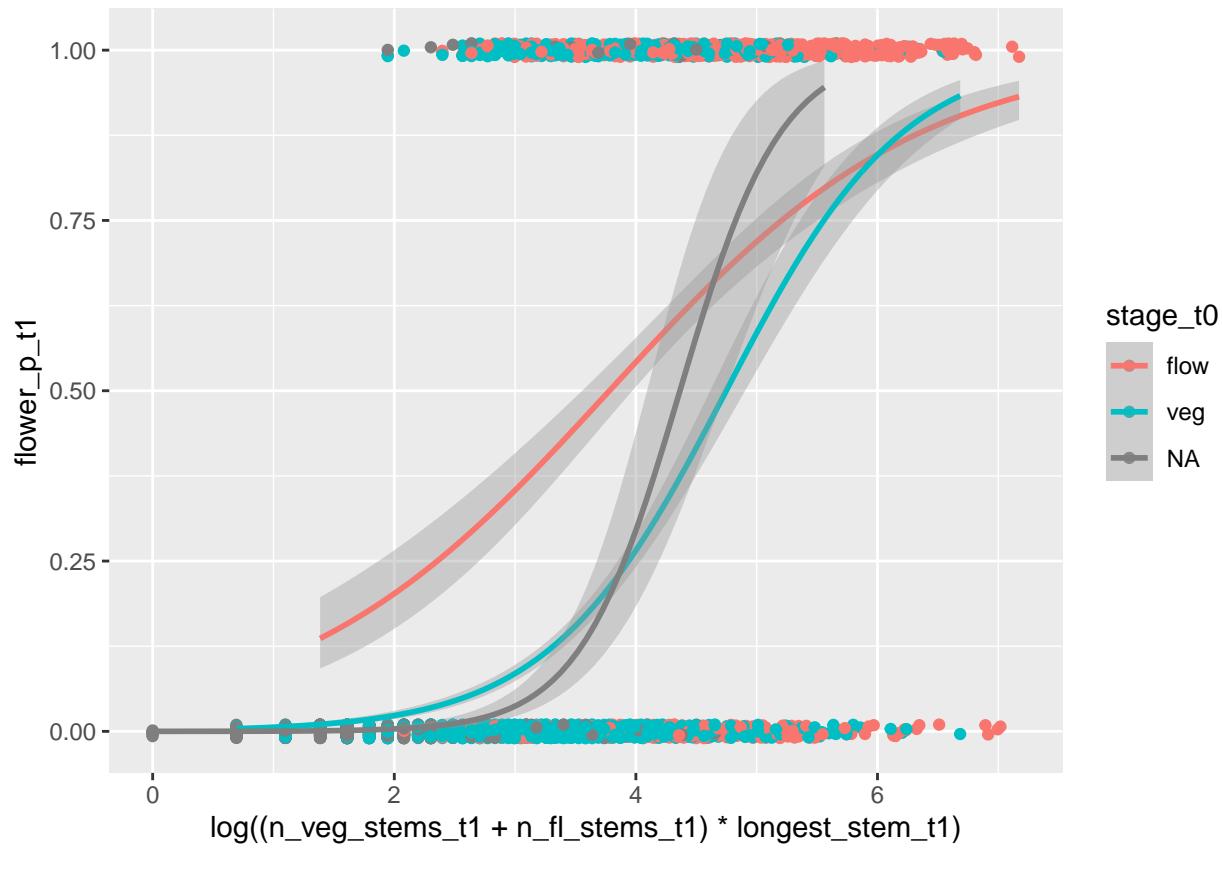
```



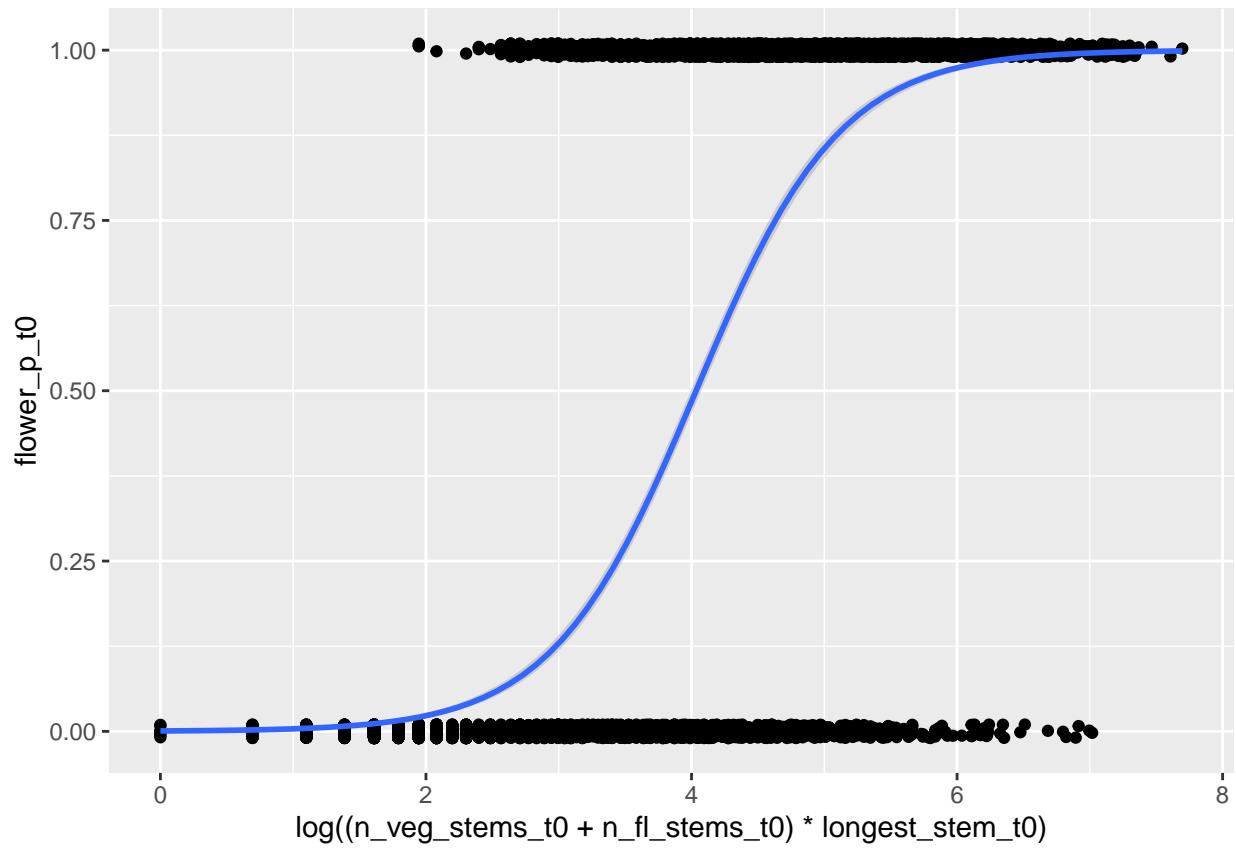
```
## `geom_smooth()` using formula 'y ~ x'
## Warning: Removed 41750 rows containing non-finite values (stat_smooth).
## Warning: Removed 41750 rows containing missing values (geom_point).
```



```
## `geom_smooth()` using formula 'y ~ x'  
## Warning: Removed 48515 rows containing non-finite values (stat_smooth).  
## Warning: Removed 48515 rows containing missing values (geom_point).
```



```
## `geom_smooth()` using formula 'y ~ x'
## Warning: Removed 46248 rows containing non-finite values (stat_smooth).
## Warning: Removed 46248 rows containing missing values (geom_point).
```



```

##      AIC      dAIC   df weight
## fp2  4831.2     0.0  2    1
## fp1 10677.7  5846.4  2 <0.001

##
## Call:
## glm(formula = flower_p_t0 ~ log(volume_t0), family = "binomial",
##      data = data1 %>% filter(volume_t0 > 0))
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -3.3181  -0.3835  -0.1508   0.3920   2.7828
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -7.43932   0.16211 -45.89   <2e-16 ***
## log(volume_t0) 1.84400   0.04011  45.98   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 10480.8 on 7821 degrees of freedom
## Residual deviance: 4827.2 on 7820 degrees of freedom
## AIC: 4831.2
##
## Number of Fisher Scoring iterations: 6

```

Base model selection

Below are the models tested and the summary of the model with the lowest AICc

```

## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.00247872 (tol = 0.002, component 1)

## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly uniden-
## - Rescale variables?

##                2.5 %      97.5 %
## .sig01          NA        NA
## (Intercept) -3.1886945751 -2.4585024235
## log(n_stems_t0) 2.0788904846  2.2381292494
## populationHK    0.5043676182  0.7912521333
## populationKS    0.5646366083  0.8858455365
## populationRU   -0.1799691113  0.2178427151
## herb_shading_t0 -0.0042558658  0.1079635689
## shrub_shading_t0  0.0001677352  0.0714310663
## soil_d1         -0.0252096255  0.0150783302
## soil_d2         -0.0318380435  0.0094922265
## soil_d3         -0.0455677497 -0.0040543401
## rock            -0.0060874975  0.0001771762
## slope           0.0095079903  0.0172951968

## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: flower_p_t0 ~ log(n_stems_t0) + population + shrub_shading_t0 +
##          soil_d3 + slope + (1 | year_t0)
## Data: data1 %>% filter(n_stems_t0 > 0)
##
##      AIC      BIC      logLik deviance df.resid
## 8154.1  8219.6  -4068.1    8136.1     10642
##
## Scaled residuals:
##      Min      1Q      Median      3Q      Max
## -16.7059 -0.3712 -0.2387  0.3682  8.0632
##
## Random effects:
## Groups   Name       Variance Std.Dev.
## year_t0 (Intercept) 0.286    0.5348
## Number of obs: 10651, groups: year_t0, 16
##
## Fixed effects:
##                  Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.914618  0.158430 -18.397 < 2e-16 ***
## log(n_stems_t0) 2.152545  0.040152  53.610 < 2e-16 ***
## populationHK  0.634090  0.072692   8.723 < 2e-16 ***
## populationKS  0.763408  0.078761   9.693 < 2e-16 ***
## populationRU  0.120441  0.094514   1.274  0.2026
## shrub_shading_t0 0.043056  0.018037   2.387  0.0170 *
## soil_d3      -0.017799  0.009513  -1.871  0.0613 .
## slope        0.011758  0.001794   6.552 5.66e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

## 
## Correlation of Fixed Effects:
##          (Intr) l(_0) ppltHK ppltKS ppltRU shr_0 sol_d3
## lg(n_stm_0) -0.245
## populatinHK -0.158  0.061
## populatinKS -0.209 -0.001  0.337
## populatinRU -0.183  0.034  0.293  0.301
## shrb_shdn_0 -0.082 -0.014 -0.142  0.073  0.069
## soil_d3      -0.260 -0.113 -0.006  0.001 -0.022 -0.067
## slope        -0.248  0.093 -0.181  0.172  0.063  0.040  0.161

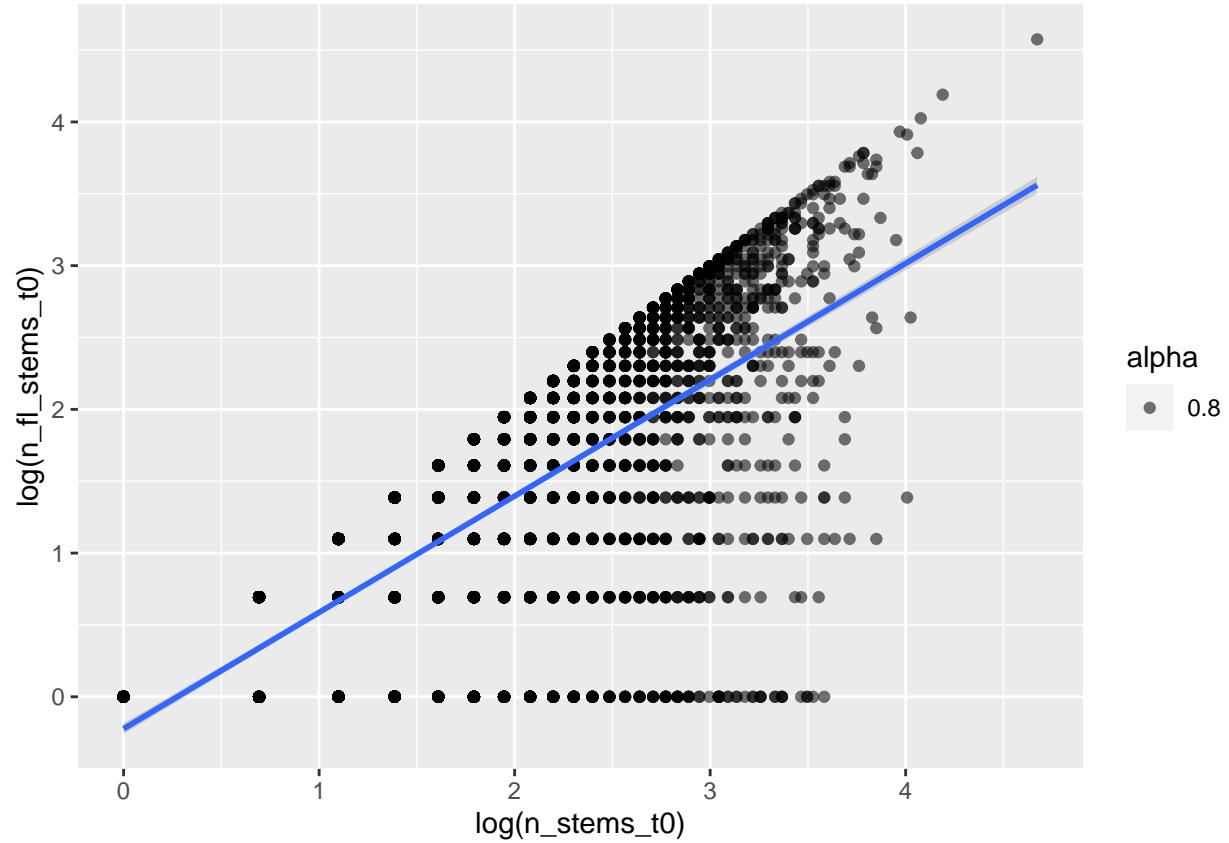
```

Flower/seed numbers

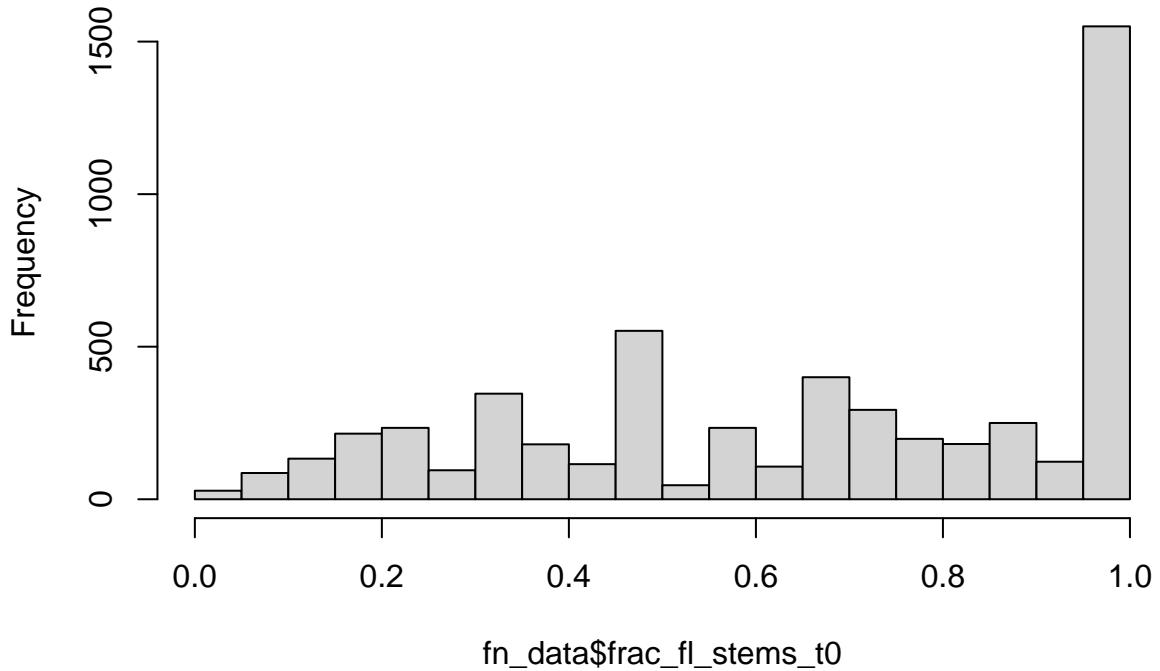
- 1) OK, so taking the obvious; number of flowering stalks, is problematic as n_fl_stems is also used to calculate both n_stems and volume.
- 2) One option would be to calculate the probability of a stalk to “turn” into a flowering stem, and either use a fixed probability or size dependency.
- 3) Another option would be to skip flower stems and calculate the total number of seeds, but that would again be calculated by multiplying the average seeds per stem on an individual by the n_fl_stems. Here another issue would be the large number of individuals that don’t produce any seeds ($n = 677$), which might cause problems when modeling because predictions could be negative (with average numbers (not integers) I can’t use a poisson, and with 0’s I also can’t use a gamma distribution).
- 4) Include an extra vital rate (“abortion probability”) i.e., probability that even if you flower, you don’t produce seeds, and then the n_seeds is modeled, conditional on fp == 1 and ab_p == 0

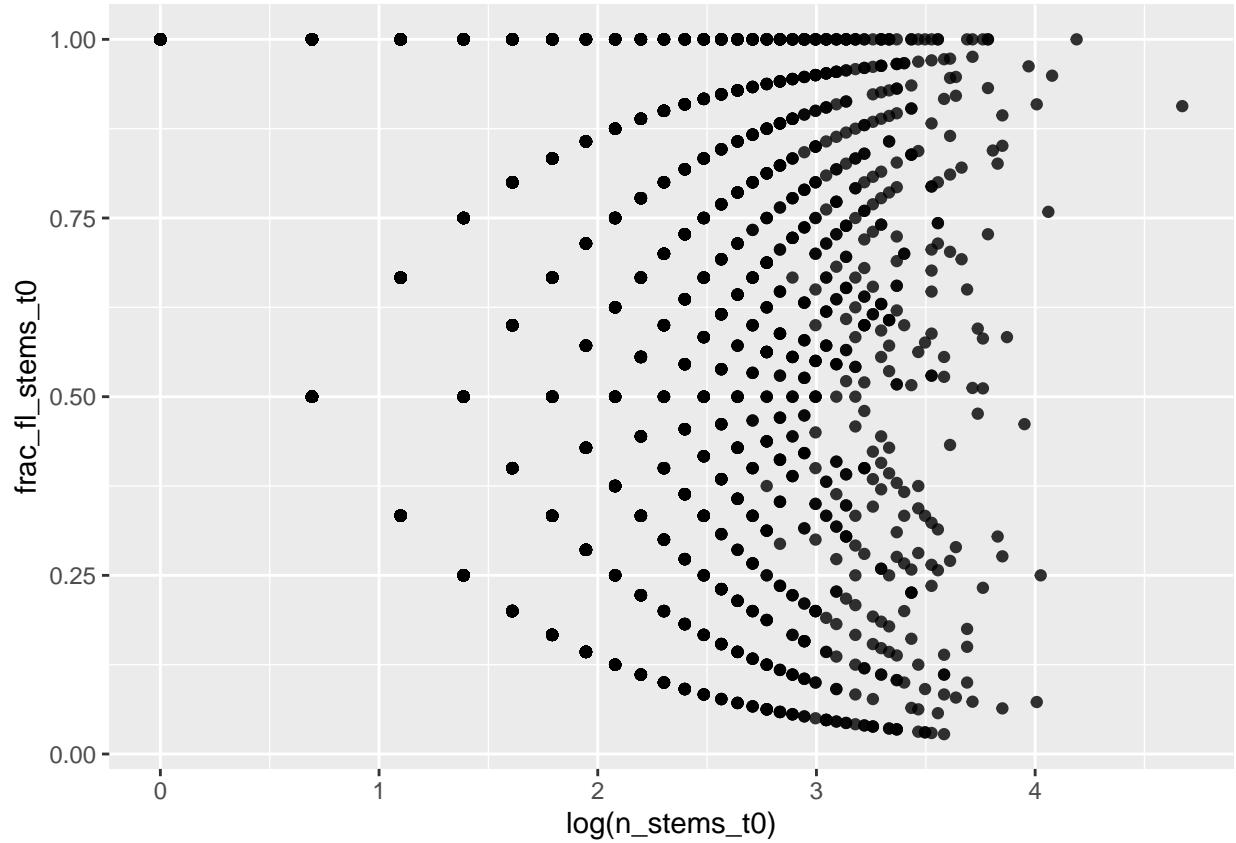
4 is probably the best solution here (see graphs below for the 4 options)

```
## `geom_smooth()` using formula 'y ~ x'
```

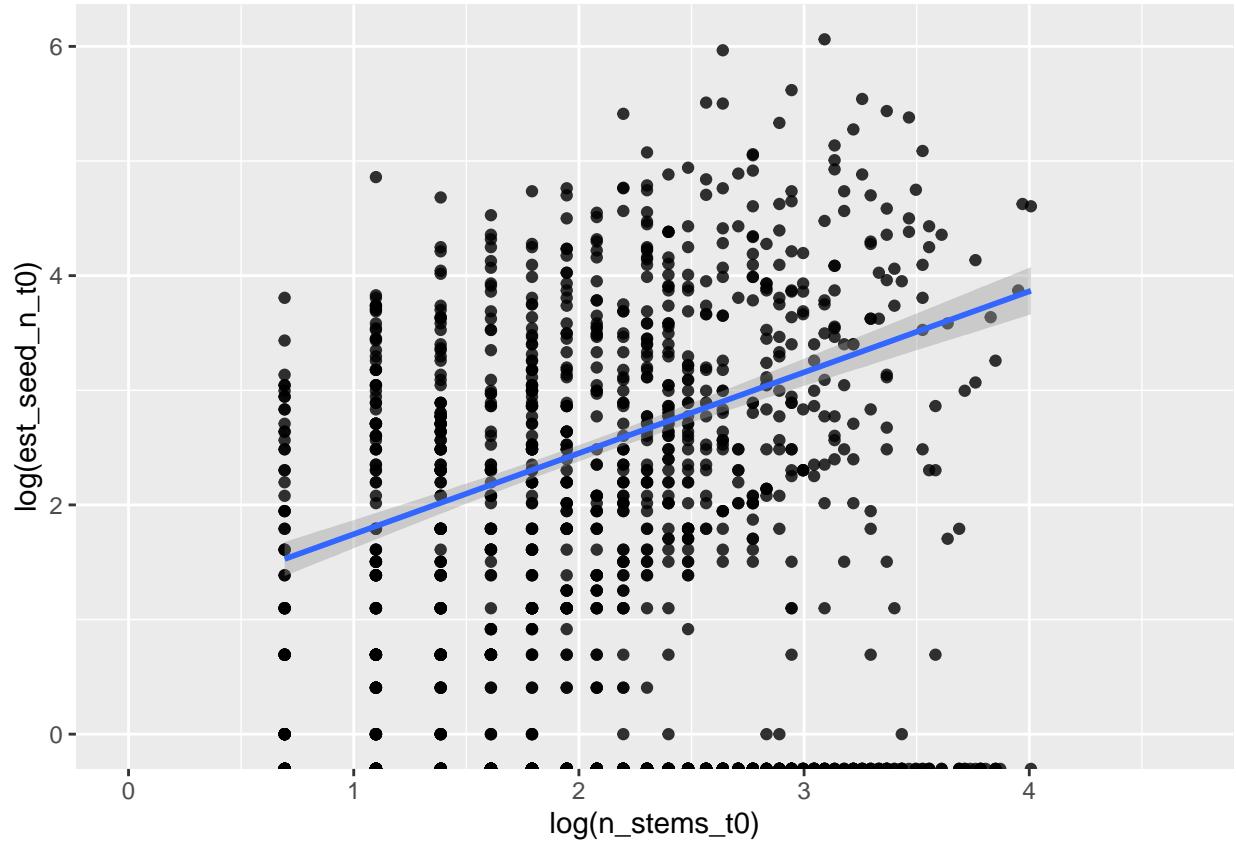


Histogram of fn_data\$frac_fl_stems_t0

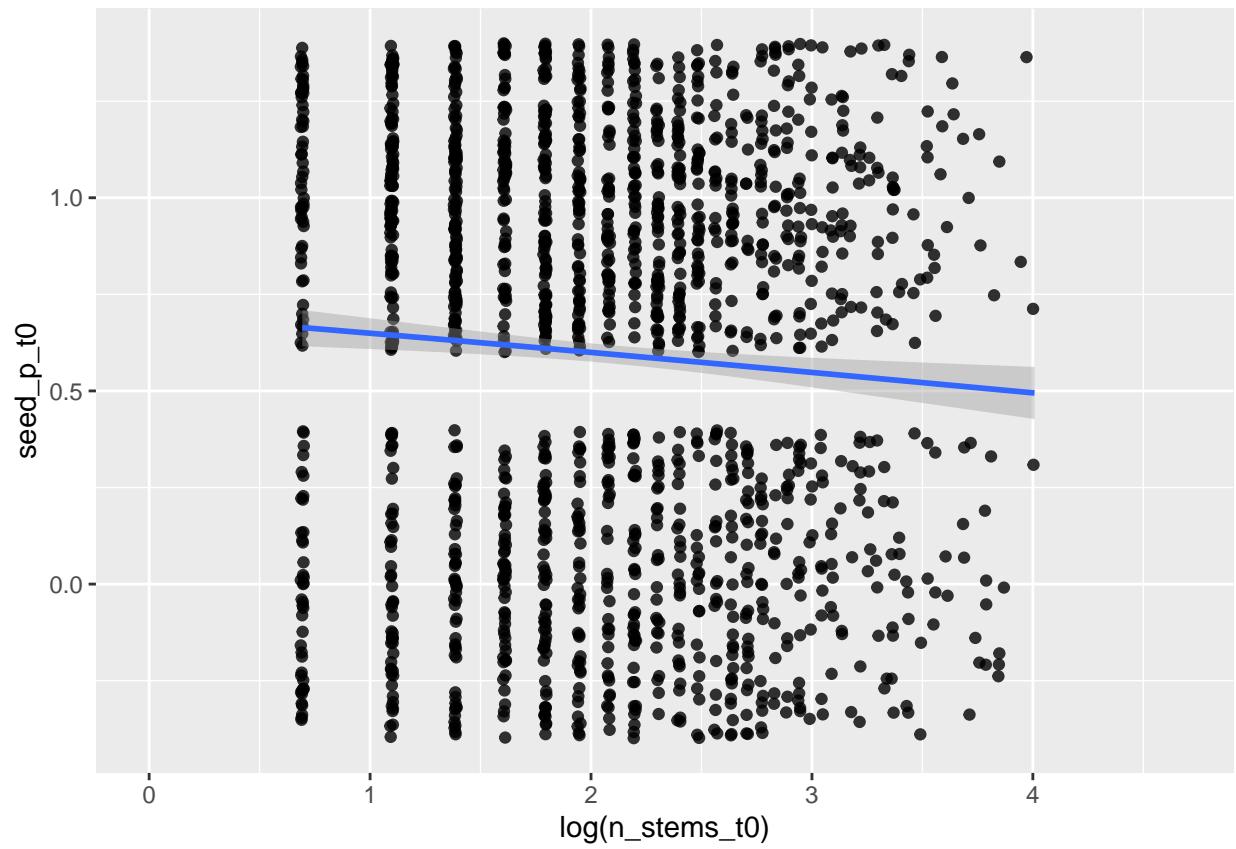




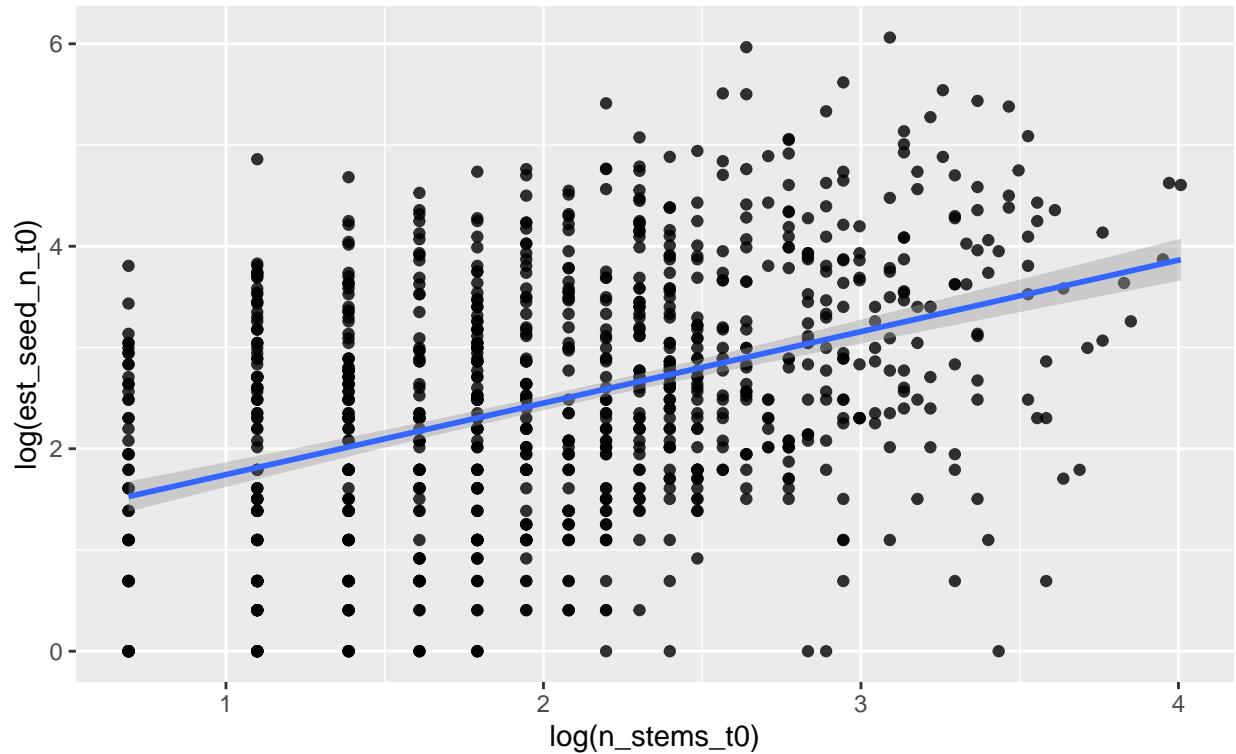
```
## `geom_smooth()` using formula 'y ~ x'  
## Warning: Removed 4370 rows containing non-finite values (stat_smooth).  
## Warning: Removed 3693 rows containing missing values (geom_point).
```



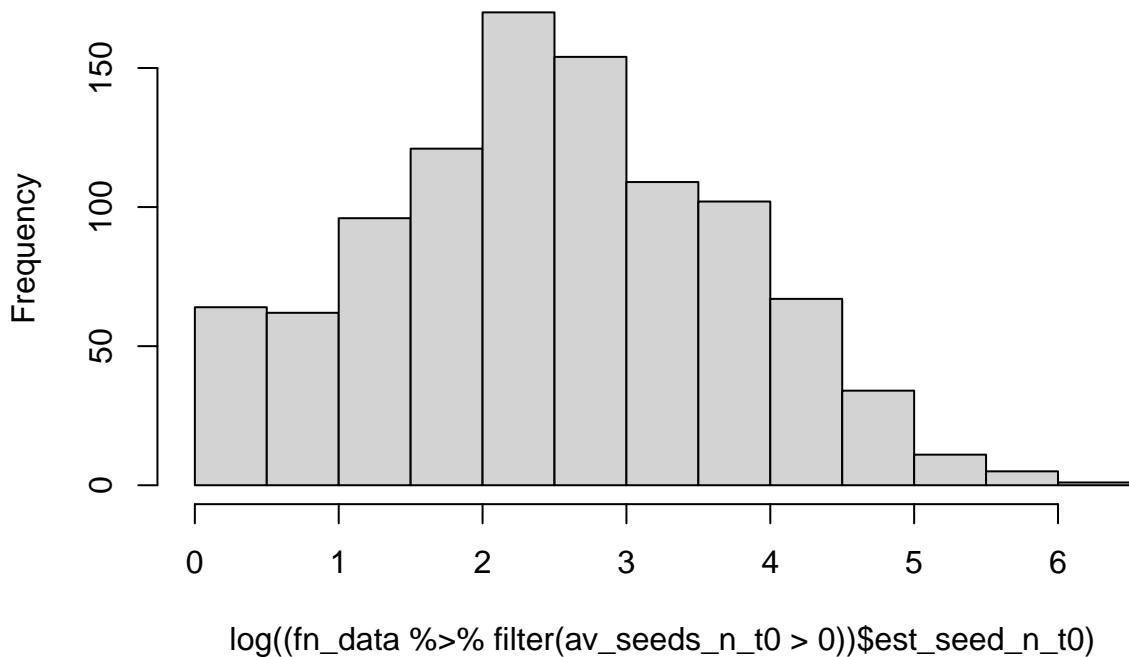
```
## [1] "4th option"
## `geom_smooth()` using formula 'y ~ x'
## Warning: Removed 3693 rows containing non-finite values (stat_smooth).
## Removed 3693 rows containing missing values (geom_point).
```



estimated number of seeds per individual ($\text{av_n_seeds_t0} * \text{n_fl_stems_t0}$)
conditioned on $\text{p_flower} == 1$ and $\text{abortion_p} == 0$



Histogram of $\log((\text{fn_data} \%>\% \text{filter}(\text{av_seeds_n_t0} > 0))\est_seed_n



```
## [1] "gamma shape"
##
## Alpha: 0.93580130
## SE:    0.03670346
##
## Call:
## glm(formula = est_seed_n_t0 ~ log(n_stems_t0), family = Gamma(link = log),
##      data = fn_data %>% filter(seed_p_t0 == 1))
##
## Deviance Residuals:
##       Min      1Q      Median      3Q      Max
## -2.4834 -1.2147 -0.5799   0.2309   4.0356
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.68698   0.12024  14.03   <2e-16 ***
## log(n_stems_t0) 0.69294   0.05584  12.41   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Gamma family taken to be 1.636238)
##
## Null deviance: 1491.7 on 995 degrees of freedom
## Residual deviance: 1238.5 on 994 degrees of freedom
## AIC: 8168.4
```

```

## 
## Number of Fisher Scoring iterations: 6
## [1] "summary model outpute with dispersion corrected using 1/alpha above"
##
## Call:
## glm(formula = est_seed_n_t0 ~ log(n_stems_t0), family = Gamma(link = log),
##      data = fn_data %>% filter(seed_p_t0 == 1))
##
## Deviance Residuals:
##       Min     1Q Median     3Q    Max
## -2.4834 -1.2147 -0.5799  0.2309  4.0356
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) 1.68698   0.09717 17.36   <2e-16 ***
## log(n_stems_t0) 0.69294   0.04512 15.36   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Gamma family taken to be 1.068603)
##
## Null deviance: 1491.7 on 995 degrees of freedom
## Residual deviance: 1238.5 on 994 degrees of freedom
## AIC: 8168.4
##
## Number of Fisher Scoring iterations: 6

```

Base model selection

```

##          2.5 %      97.5 %
## .sig01        NA        NA
## (Intercept) 0.194655687 2.212989494
## log(n_stems_t0) -0.362101778 -0.029112053
## populationHK 1.312483674 1.963946607
## populationKS 0.079155294 0.729320710
## populationRU 0.649748783 1.634546832
## herb_shading_t0 -0.200001638 0.050244587
## shrub_shading_t0 -0.163670630 -0.025700659
## soil_d1        -0.072592323 0.008449277
## soil_d2        -0.069275546 0.014766691
## soil_d3        0.010884636 0.094506163
## rock           -0.007011138 0.006485308
## slope          -0.008027059 0.007096630

## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.00523382 (tol = 0.002, component 1)

## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly uniden-
## - Rescale variables?

##          2.5 %      97.5 %
## .sig01        NA        NA
## .sigma         NA        NA
## (Intercept) 0.371972181 1.278101079
## log(n_stems_t0) 0.704629752 0.874139629

```

```

## populationHK      0.666105897 0.965085473
## populationKS      0.374317198 0.717549429
## populationRU      0.499928249 0.958714292
## herb_shading_t0   -0.047797431 0.079769027
## shrub_shading_t0  0.000715283 0.071341295
## soil_d1           -0.039122596 0.001813340
## soil_d2           -0.027473400 0.014065255
## soil_d3           -0.011700836 0.029244836
## rock              -0.001535844 0.004800721
## slope              -0.002008670 0.005089772

## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: Gamma ( log )
## Formula: est_seed_n_t0 ~ log(n_stems_t0) + population + shrub_shading_t0 +
##          (1 | year_t0)
## Data: fn_data %>% filter(seed_p_t0 == 1)
##
##      AIC      BIC  logLik deviance df.resid
## 7884.2 7923.4 -3934.1    7868.2      988
##
## Scaled residuals:
##      Min      1Q  Median      3Q     Max
## -0.9156 -0.6434 -0.3059  0.2914  7.2600
##
## Random effects:
## Groups   Name        Variance Std.Dev.
## year_t0 (Intercept) 0.349     0.5908
## Residual            1.133     1.0645
## Number of obs: 996, groups: year_t0, 12
##
## Fixed effects:
##             Estimate Std. Error t value Pr(>|z|)
## (Intercept)  0.92513   0.19351   4.781 1.75e-06 ***
## log(n_stems_t0) 0.76891   0.04030  19.082 < 2e-16 ***
## populationHK  0.80298   0.07410  10.836 < 2e-16 ***
## populationKS  0.52802   0.08146   6.482 9.07e-11 ***
## populationRU  0.72138   0.10997   6.560 5.38e-11 ***
## shrub_shading_t0 0.03211   0.01761   1.823   0.0682 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) ppltHK ppltKS ppltRU
## lg(n_stm_0) -0.441
## populatinHK -0.241  0.065
## populatinKS -0.224  0.066  0.497
## populatinRU -0.304  0.177  0.387  0.336
## shrb_shdn_0 -0.122 -0.017 -0.038  0.087  0.182

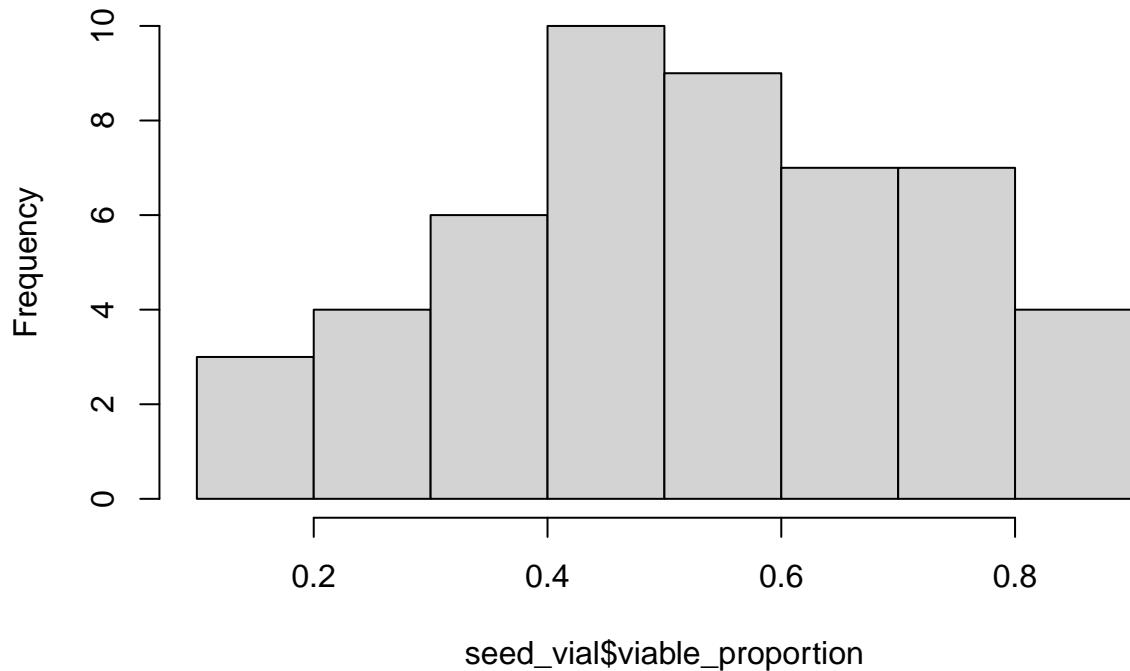
```

State-independent vital rates

Seed viability

```
## New names:  
## * `` -> `...`
```

Histogram of seed_vial\$viable_proportion



```
## # A tibble: 1 x 2  
##   viable_prop viable_sd  
##   <dbl>     <dbl>  
## 1     0.532    0.199  
  
## # A tibble: 8 x 3  
##   locality      viable_prop viable_sd  
##   <chr>          <dbl>     <dbl>  
## 1 Císařská rokle      0.450    0.161  
## 2 Domica            0.51      NA  
## 3 Haknovec          0.592    0.136  
## 4 Koda               0.393    0.237  
## 5 Radotínské údolí    0.717    0.116  
## 6 Vanovice           0.58      NA  
## 7 Zadiel              0.58      NA  
## 8 Železná vrata       0.55      NA  
  
## # A tibble: 12 x 3  
##   year  viable_prop viable_sd  
##   <dbl>     <dbl>     <dbl>  
## 1    2006      0.443    0.162
```

```

## 2 2007      0.378    0.217
## 3 2008      0.538    0.300
## 4 2009      0.454    0.136
## 5 2010      0.433    0.257
## 6 2011      0.514    0.205
## 7 2012      0.569    0.164
## 8 2013      0.666    0.0878
## 9 2014      0.436    0.0705
## 10 2015     0.492    0.192
## 11 2016     0.793    0.0762
## 12 2017     0.714    0.103

```

Lifecycle diagram

