Statistics on Fitness Trial

James Maino

28/02/2019

Load and clean data.

d0 = read\_csv("https://docs.google.com/spreadsheets/d/e/2PACX-1vQmErDprvdoLAIFM9Lf9Fr7JnH3\_2nOYlBmYwTwr814uPDgecxpgWnExdj4WCuCVVDd915GskIWf\_NY/pub?gid=0&single=true&output=csv")

## Parsed with column specification:  
## cols(  
## .default = col\_double(),  
## date = col\_character(),  
## location = col\_character(),  
## month = col\_character(),  
## weather = col\_character(),  
## plot = col\_character(),  
## trtcode = col\_integer(),  
## damage = col\_integer(),  
## abund.q1 = col\_integer(),  
## abund.q2 = col\_integer(),  
## abund.q3 = col\_integer(),  
## abund.q4 = col\_integer(),  
## abund.total = col\_integer(),  
## extrapabund = col\_integer(),  
## kdr\_insects.tested.q1 = col\_integer(),  
## kdr\_insects.tested.q2 = col\_integer(),  
## kdr\_insects.tested.q3 = col\_integer(),  
## kdr\_insects.tested.q4 = col\_integer()  
## )

## See spec(...) for full column specifications.

names(d0)

## [1] "date" "location"   
## [3] "month" "time"   
## [5] "weather" "plot"   
## [7] "trtcode" "damage"   
## [9] "abund.q1" "abund.q2"   
## [11] "abund.q3" "abund.q4"   
## [13] "abund.total" "extrapabund"   
## [15] "kdr\_percent.q1" "kdr\_insects.tested.q1"   
## [17] "kdr\_percent.q2" "kdr\_insects.tested.q2"   
## [19] "kdr\_percent.q3" "kdr\_insects.tested.q3"   
## [21] "kdr\_percent.q4" "kdr\_insects.tested.q4"   
## [23] "kdr\_percent.mean" "OP\_bio\_mortality.8hr.Low.q1"   
## [25] "OP\_bio\_mortality.8hr.Low.q2" "OP\_bio\_mortality.8hr.Low.q3"   
## [27] "OP\_bio\_mortality.8hr.Low.q4" "OP\_bio\_mortality.8hr.High.q1"   
## [29] "OP\_bio\_mortality.8hr.High.q2" "OP\_bio\_mortality.8hr.High.q3"   
## [31] "OP\_bio\_mortality.8hr.High.q4" "OP\_bio\_mortality.24hr.Low.q1"   
## [33] "OP\_bio\_mortality.24hr.Low.q2" "OP\_bio\_mortality.24hr.Low.q3"   
## [35] "OP\_bio\_mortality.24hr.Low.q4" "OP\_bio\_mortality.24hr.High.q1"  
## [37] "OP\_bio\_mortality.24hr.High.q2" "OP\_bio\_mortality.24hr.High.q3"  
## [39] "OP\_bio\_mortality.24hr.High.q4" "lucerne"   
## [41] "capeweed" "grass"   
## [43] "ground" "seradella"   
## [45] "clover" "radish"   
## [47] "other.weed"

d = d0 %>%   
 mutate(trtcode = ifelse(is.na(trtcode), 1, trtcode)) %>%  
 filter(trtcode == 1) %>%  
 mutate(date = as.Date(date, '%d/%m/%Y')) %>%  
 gather(var, val, -date, -location, -month, -time, -weather, -plot, -trtcode, -damage) %>%  
 separate(var, c('var', 'quadrat'), sep = '\\.q', convert = TRUE) %>%  
 filter(!is.na(quadrat)) %>%  
 filter(!is.na(val)) %>%  
 filter(var %in% c('kdr\_percent', 'kdr\_insects.tested')) %>%   
 spread(var, val, convert = TRUE) %>%   
 filter(!is.na(kdr\_percent)) %>%  
 mutate(month = format(date, '%B')) %>%   
 mutate(year = format(date, '%Y')) %>%   
 mutate(year\_month = format(date, '%Y-%m')) %>%   
 mutate(days = as.numeric(date - as.Date(ifelse(location == 'Arthur River', '2017-08-01', '2017-06-26'))))

## Warning: Expected 2 pieces. Missing pieces filled with `NA` in 2420 rows  
## [881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895,  
## 896, 897, 898, 899, 900, ...].

# %>% filter(location == 'Tintinara')  
   
head(d)

## # A tibble: 6 x 14  
## date location month time weather plot trtcode damage quadrat  
## <date> <chr> <chr> <dbl> <chr> <chr> <dbl> <int> <int>  
## 1 2017-06-26 Tintina~ June 11.2 foggy.~ A3 1 4 1  
## 2 2017-06-26 Tintina~ June 13.0 foggy.~ C2 1 4 1  
## 3 2017-06-26 Tintina~ June 13.6 foggy.~ D1 1 1 1  
## 4 2017-06-26 Tintina~ June 14 foggy.~ D2 1 3 1  
## 5 2017-06-26 Tintina~ June 15.4 foggy.~ E3 1 3 1  
## 6 2017-06-27 Tintina~ June 9.5 cloudy G3 1 2 1  
## # ... with 5 more variables: kdr\_insects.tested <int>, kdr\_percent <dbl>,  
## # year <chr>, year\_month <chr>, days <dbl>

Fit full model with fixed effect of time and location and random effect of plot.

d = mutate(d, days\_s = (days - mean(days))/sd(days))  
m1 = lmer(kdr\_percent ~ location/days + (1|plot), weights = kdr\_insects.tested,   
 data=d)  
d$pred\_m1 = predict(m1) # predictions with random effects  
d$pred\_fixed\_m1 = predict(m1, re.form=~0) # fixed effect only  
summary(m1)

## Linear mixed model fit by REML ['lmerMod']  
## Formula: kdr\_percent ~ location/days + (1 | plot)  
## Data: d  
## Weights: kdr\_insects.tested  
##   
## REML criterion at convergence: 1258.5  
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.6772 -0.4446 0.0764 0.5468 2.5684   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## plot (Intercept) 39.91 6.317   
## Residual 4180.33 64.655   
## Number of obs: 151, groups: plot, 24  
##   
## Fixed effects:  
## Estimate Std. Error t value  
## (Intercept) 86.112627 2.796065 30.798  
## locationTintinara -42.928332 4.115048 -10.432  
## locationArthur River:days -0.008797 0.007736 -1.137  
## locationTintinara:days -0.013672 0.008595 -1.591  
##   
## Correlation of Fixed Effects:  
## (Intr) lctnTn lctAR:  
## locatnTntnr -0.679   
## lctnArtRvr: -0.565 0.384   
## lctnTntnr:d 0.000 -0.460 0.000

Calculate the confidence intervals for the rate of decrease (% resistance per day) for each location.

confint(m1, parm=5:6)

## Computing profile confidence intervals ...

## Warning in profile.merMod(object, which = parm, signames = oldNames, ...):  
## non-monotonic profile for locationArthur River:days

## Warning in profile.merMod(object, which = parm, signames = oldNames, ...):  
## non-monotonic profile for locationTintinara:days

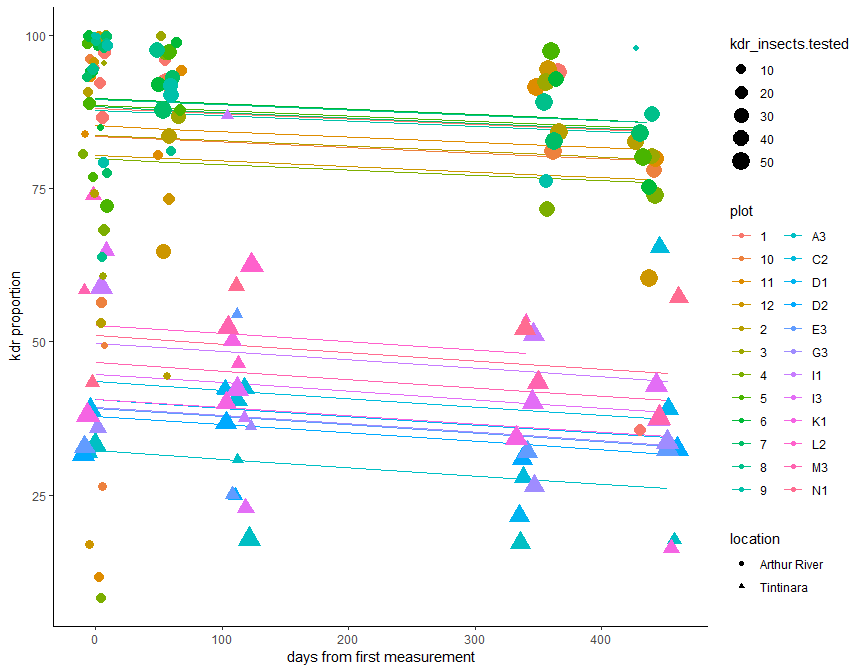
## Warning in confint.thpr(pp, level = level, zeta = zeta): bad spline fit for  
## locationArthur River:days: falling back to linear interpolation

## Warning in confint.thpr(pp, level = level, zeta = zeta): bad spline fit for  
## locationTintinara:days: falling back to linear interpolation

## 2.5 % 97.5 %  
## locationArthur River:days -0.008910305 -0.008646071  
## locationTintinara:days -0.013624738 -0.013339278

Plot kdr frequency for each replicate. Letters show predicted values while lines show observed data.

p1 = ggplot(d,   
 aes(x=days, colour=plot)) +  
 # geom\_text(aes(y=kdr\_percent.mean, label = plot)) +  
 geom\_line(aes(y=pred\_m1)) +  
 # geom\_line(aes(y=pred\_fixed\_m1, linetype = location), size =2, colour ='black') +  
 geom\_point(aes(y=kdr\_percent, group = plot, size = kdr\_insects.tested, shape = location),   
 position = position\_jitter(width = 10)) +  
 xlab('days from first measurement') +   
 ylab('kdr proportion') +  
 theme\_classic()  
ggsave(plot=p1, 'plots/full\_model.png', height=7, width=9)  
p1



Interestingly, if we ignore individual effects of plots we reduce the statistical power of the analysis.

m2 = lm(kdr\_percent ~ location/days, weights = kdr\_insects.tested,   
 data=d)  
summary(m2)

##   
## Call:  
## lm(formula = kdr\_percent ~ location/days, data = d, weights = kdr\_insects.tested)  
##   
## Weighted Residuals:  
## Min 1Q Median 3Q Max   
## -260.232 -32.446 8.913 39.650 169.418   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 86.632837 2.315481 37.415 <2e-16 \*\*\*  
## locationTintinara -44.329522 3.486068 -12.716 <2e-16 \*\*\*  
## locationArthur River:days -0.010745 0.008444 -1.273 0.205   
## locationTintinara:days -0.012279 0.009297 -1.321 0.189   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 71.48 on 147 degrees of freedom  
## Multiple R-squared: 0.7326, Adjusted R-squared: 0.7272   
## F-statistic: 134.3 on 3 and 147 DF, p-value: < 2.2e-16

Summarise results across all replicates and plot means and standard error.

dsum = d %>%  
 group\_by(year\_month, location, trtcode) %>%   
 summarise(  
 days = mean(days),  
 days\_s = mean(days\_s),  
 kdr\_se = sd(kdr\_percent, na.rm=T)/sqrt(n()),   
 kdr\_percent = mean(kdr\_percent, na.rm=T),   
 n = n()  
 )  
dsum$pred\_fixed\_m1 = predict(m1, re.form=~0, newdata = dsum)  
  
p2 = ggplot(dsum, aes(x=year\_month, y=kdr\_percent, colour=location)) +  
 geom\_point(aes()) +  
 geom\_line(aes(y=pred\_fixed\_m1, group = location)) +   
 geom\_errorbar(aes(ymin=kdr\_percent-kdr\_se, ymax=kdr\_percent+kdr\_se),  
 width = 0.1) +   
 xlab('date') +   
 ylab('kdr proportion') +  
 theme\_classic()  
ggsave(plot=p2, 'plots/kdr\_summary.png')

## Saving 5 x 4 in image

p2

