Update

Yujun Zhou

February 25, 2019

Next analysis steps will be to:

(1) Include data from years 2009-14  
  
  
We have a lot of missing in the previous year’s data. The observations with all the variables of interest is 1142 during 2011-2015 and 741 of them are from 2014 and 2015.  
  
  
  
(2) Account for bloom period, when we expect effects of sunflower pollen to occur, in two ways:  
  
 (a) restrict analysis to data from August & September only, when sunflower is in bloom, comparing same location with blooming vs no blooming

Restricting month to August and September leaves an observation of 443 and we are seeing significant results on all the mites regressions. But not for Nosema regressions.

A sample regression result:

Call:

lm(formula = mites ~ log(sunflower + 1) + average\_mintemp + average\_precip +

as.factor(month), data = bee.bloom.month)

Residuals:

Min 1Q Median 3Q Max

-57.22 -30.04 -12.67 10.63 442.28

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 56.20923 12.34402 4.554 6.99e-06 \*\*\*

log(sunflower + 1) -2.40462 0.74737 -3.217 0.0014 \*\*

average\_mintemp -1.04389 0.76885 -1.358 0.1753

average\_precip -0.07842 0.04878 -1.608 0.1087

as.factor(month)9 9.44134 5.25354 1.797 0.0731 .

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 49.66 on 404 degrees of freedom

(34 observations deleted due to missingness)

Multiple R-squared: 0.04942, Adjusted R-squared: 0.04001

F-statistic: 5.251 on 4 and 404 DF, p-value: 0.0003932

(b) include all months and an interaction term between acres of sunflower and month, to account for effects of location.   
   
an interaction of bloom month dummies and sunflower acres does have some impact the number of mites in the apiaries and surprisingly on the nosema as well.

Sample Mites regression:

Call:

lm(formula = mites ~ log(sunflower + 1) \* bloom\_dummy + average\_mintemp +

average\_precip + as.factor(month), data = bee1115)

Residuals:

Min 1Q Median 3Q Max

-51.56 -25.72 -9.40 8.15 442.82

Coefficients: (1 not defined because of singularities)

Estimate Std. Error t value Pr(>|t|)

(Intercept) 7.91526 9.98253 0.793 0.428016

log(sunflower + 1) -0.72444 0.67596 -1.072 0.284105

bloom\_dummy 45.37759 11.07884 4.096 4.54e-05 \*\*\*

average\_mintemp -0.70322 0.36351 -1.935 0.053331 .

average\_precip 0.01940 0.02505 0.774 0.438961

as.factor(month)2 -0.74416 24.87760 -0.030 0.976143

as.factor(month)3 1.54912 17.09140 0.091 0.927799

as.factor(month)4 5.94885 12.95345 0.459 0.646155

as.factor(month)5 5.85713 12.88800 0.454 0.649593

as.factor(month)6 11.02440 11.50300 0.958 0.338095

as.factor(month)7 25.50274 11.68713 2.182 0.029331 \*

as.factor(month)8 -11.33904 4.65182 -2.438 0.014959 \*

as.factor(month)10 40.42337 10.58498 3.819 0.000142 \*\*\*

as.factor(month)11 13.38118 12.75641 1.049 0.294440

as.factor(month)12 0.12409 17.64973 0.007 0.994392

log(sunflower + 1):bloom\_dummy -1.48082 0.93924 -1.577 0.115200

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 45.42 on 1007 degrees of freedom

(119 observations deleted due to missingness)

Multiple R-squared: 0.0967, Adjusted R-squared: 0.08324

F-statistic: 7.186 on 15 and 1007 DF, p-value: 3.418e-15

Sample Nosema regression:

Call:

lm(formula = nosema ~ log(sunflower + 1) \* bloom\_dummy + average\_mintemp +

average\_precip + as.factor(month), data = bee1115)

Residuals:

Min 1Q Median 3Q Max

-2.0123 -0.1503 -0.1073 -0.0399 11.7061

Coefficients: (1 not defined because of singularities)

Estimate Std. Error t value Pr(>|t|)

(Intercept) 1.8099545 0.1760278 10.282 < 2e-16 \*\*\*

log(sunflower + 1) 0.0342507 0.0119184 2.874 0.00414 \*\*

bloom\_dummy -1.6283576 0.1954049 -8.333 2.55e-16 \*\*\*

average\_mintemp -0.0045213 0.0064101 -0.705 0.48076

average\_precip 0.0001428 0.0004416 0.323 0.74651

as.factor(month)2 -1.1971148 0.4386800 -2.729 0.00647 \*\*

as.factor(month)3 -0.2657184 0.3013827 -0.882 0.37817

as.factor(month)4 -0.9496691 0.2284155 -4.158 3.49e-05 \*\*\*

as.factor(month)5 -1.3538878 0.2272603 -5.957 3.54e-09 \*\*\*

as.factor(month)6 -1.4310374 0.2028400 -7.055 3.21e-12 \*\*\*

as.factor(month)7 -1.6007287 0.2060009 -7.770 1.92e-14 \*\*\*

as.factor(month)8 -0.0115387 0.0820903 -0.141 0.88824

as.factor(month)10 -1.6987608 0.1866510 -9.101 < 2e-16 \*\*\*

as.factor(month)11 -1.5542420 0.2249413 -6.910 8.61e-12 \*\*\*

as.factor(month)12 -1.3730776 0.3112273 -4.412 1.14e-05 \*\*\*

log(sunflower + 1):bloom\_dummy -0.0339912 0.0165626 -2.052 0.04040 \*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.801 on 1007 degrees of freedom

(119 observations deleted due to missingness)

Multiple R-squared: 0.1389, Adjusted R-squared: 0.1261

F-statistic: 10.83 on 15 and 1007 DF, p-value: < 2.2e-16

(3) Determine if there are enough virus data points to include viruses as responses virus variable more significant (testing is expensive)

Not enough virus data and only in recent years data.

(4) Kathy also suggested doing the analysis restricted to only states where sunflower grows, to account for some location effects  
  
This would limit the number of observations to about 130, but it does show significant impacts on having more sunflower acreage. Interacting the blooming months dummies with the sunflower acreage is still not significant on this subset of data.   
  
  
(5) Hannah says that apiary size has a big effect on management style and pathogen/pest pressure. It may be a good idea to include number of colonies in the apiary as a covariate.

we only have the colonies in the apiary variable in the 2014-2015 data. Regressing on the subset data of 2014-2015, it proves to be significant and helps to reduce standard errors in coefficient estimates.

(2) I forgot to discuss 'number of alive bees', which your document listed as a response. We are finding it very hard to use measures of colony growth/size in our apiary study since beekeepers often move frames. It would make sense to use to drop this response and just focus on pathogens/pests (Nosema, mites, viruses).

Right, we don’t have enough observations anyway.