ISSN 2349-0837
Volume 7 Number 2
JOURNAL OF ADVANCES IN AGRICULTURE

Evolutionary Dynamics of Bee Colony Collapse Disorder: First Steps toward a Mathematical Model of the Contagion Hypothesis

Jonathan David Farley [1]

[1] Department of Mathematics, California Institute of Technology, Pasadena, California 91125, United States of America. Institut für Algebra, Johannes Kepler Universität Linz, A-4040 Linz, Austria. Department of Computer Science, University of Maine, Orono, Maine 04469, United States of America.

[1] Present address: Department of Mathematics, Morgan State University, 1700 E. Cold Spring Lane, Baltimore, MD 21251, United States of America (lattice.theory@gmail.com).

Abstract.

The disappearance of honey bees from many managed colonies in the United States and Europe in 2006 and 2007 is modeled under the assumption that the cause is some contagion. Based on the limited data available, we use a simple model to suggest that colony collapse disorder will not destroy all colonies in the United States. To predict the evolution of future outbreaks, however, and perhaps trace their origins, it is recommended that graph-theoretic data be collected, and that census data be collected on a more frequent basis, concerning bee populations.

Keywords. Honey bee; logistic curve; mathematical epidemiology; commercial pollination.

Introduction

In the closing months of 2006, a dramatic decline in the number of managed honey bee colonies in the United States was observed. [1] [2] Colonies have been found almost completely devoid of bees, a phenomenon now known as "colony collapse disorder" (CCD). In a recent study, vanEngelsdorp *et al.* looked at the evolution of roughly 150,000 managed bee colonies in 15 American states from September 2006 to March 2007. [3] They ascertain that roughly one-third of all colonies surveyed were lost in this period (not all of these to CCD, it must be noted).

With the value of commercial bee pollination in the United States estimated at $14.6 billion annually, and with 3 million bee colonies, CCD is an issue of great concern to policy makers, environmentalists, and the general public. [4] [5] The cause of CCD is as yet unknown, but it is clearly important to try to predict the course of CCD with whatever tools we have available. Mathematical epidemiology has yielded useful insights regarding how to manage other diseases. [6]

1. Results

Under the assumption that the cause of CCD is some sort of contagion, we apply standard techniques to deduce that, if our assumptions hold, CCD will not wipe out all managed colonies, and that, if the number of infections has only been falling since the start of the time period of concern, then the number of infected but uncollapsed colonies at the beginning of that time period must have been more than 5% of all colonies. (If future observations greatly contradict the conclusions of the model, then we may regard this as evidence that the contagion hypothesis is incorrect.)

The more significant conclusion from our work, however, is that we must start collecting more data, and different sorts of data, so that we will be better able to predict the course of, or trace the origins of, another outbreak of this or some other disorder in the American bee population.

2. Review of a Basic Model for the Spread of Infectious Disease [7]

Let n be the total number of colonies. Let x be the number of healthy colonies still susceptible to CCD (with x_0 being the number of healthy colonies at the beginning of the time period in question), let y be the number of uncollapsed colonies afflicted with the (hypothesized) contagion that leads to CCD, and let z be the number of "dead" (collapsed) colonies.

If an infected colony is somehow brought into contact with a healthy colony, the probability of infection per unit time interval will be denoted β. The probability an infected colony collapses per unit time interval will be denoted γ. Thus, the dynamics of the disorder are governed by the following system of differential equations ([7], equations 4.5):

$$dx/dt = -\beta xy$$

There is no conflict of interest.

[2] There have been similar unexplained losses in the past. [2]

July 2017
www.cirworld.com
DOI : 10.24297/jaa.v7i2.6205

$$dy/dt = \beta xy - \gamma y$$
$$dz/dt = \gamma y.$$

We assume that at time $t = 0$ there are no dead colonies ($z_0 = 0$). As is stated in Bailey, the epidemic cannot build unless

$$\rho < x_0,$$

where $\rho = \gamma/\beta$. Moreover, we get

$$dz/dt = \gamma(n - z - x_0\, e^{-z/\rho}).$$

The expected total number of dead colonies is

$$z_\infty = (\rho^2 / x_0)\, [(x_0 / \rho) - 1 + \alpha]$$

where

$$\alpha = \{[(x_0 / \rho) - 1]^2 + (2x_0\, y_0 / \rho^2)\}^{1/2}$$

([7], equations 4.9, 4.10, and 4.12).[3]

3. Under the Assumption of Contagion, How Many Colonies Were Infected at the Beginning of the Time Period of Interest?

Suppose we now observe more collapsed colonies evolving. What can we conclude about the number of infected colonies at the beginning of the period of concern (assuming there is some contagion)?

Our assumption says that dz/dt is positive, which means

$$0 \le n - z - x_0\, e^{-z/\rho}.$$

If

$$n = 150000$$

is the total number of colonies and

$$z = 50000$$

is the current number of dead colonies, then we deduce that

$$50000 / \ln (x_0 / 100000) \ge \rho,$$

if $x_0 > 100000$.

As stated above, a necessary and sufficient condition for the number of infected cases to continue to rise is that $\rho < x$, so if this number (y) had reached its peak at the beginning of the time period in question, the inequality

$$x_0 \le 50000 / \ln (x_0 / 100000),$$

would be satisfied. Figure 1 shows the values of the right side of the inequality for various values of x_0.

The function on the right-hand side of the inequality is monotonically decreasing in x_0, and $x_0 = 142153$ is a fixed point. The conclusion we can draw is that, if the number of infected but not yet collapsed colonies has been on the decline since the beginning of the time period in question, then x_0, the initial value of susceptible (but uninfected) colonies, must have been *below* about 140,000. (Given the fact that we are rounding our figures, it is certainly meaningless to go beyond the two significant digits.) In other words, there would have been at least 10,000 infected (but not yet collapsed) colonies at the start of the time period of concern.

4. What Will the Long-Term Damage Be?

Beekeepers have expressed concern that this will be the end of the industry.[4] Since $x_0 + y_0 = n$, we can calculate the total number of expected "dead" colonies just given x_0 and ρ. Unfortunately, the available data do not give us these values, but for every value of x_0 (which, of course, must be less than n), we can select a value of ρ (which, as stated above, can be at most x_0) and then calculate z_∞.

We selected all values of x_0 from 1000 to $n = 150000$, going in increments of 1000, and for each such value, we let ρ range from 10 to x_0, going in increments of 10. (The BASIC computer code generating the table is given in Figure 3.) It appears as if the values of z_∞ calculated are *unimodal*, that is, z_∞ increases as ρ increases, until a maximum for z_∞ is reached, and then z_∞ decreases as ρ increases. For $x_0 \le 100000$, the value of ρ for which the maximum z_∞ is achieved is always $\rho = x_0$. The highest z_∞ value reached, however, is never above 110,000, suggesting that *a significant proportion of colonies will survive the outbreak*, if our assumptions hold, of course (Figure 2).

(Note: It would be interesting to prove analytically the statements above concerning unimodality and the maximum values of z_∞.)

5. Conclusion

In our application of the basic mathematical model, we only used round figures. The crude "one-third" figure for the number of dead colonies includes colonies that were not affected by CCD. Indeed, some beekeepers experienced

[3] Readers might be more familiar with the "basic reproductive ratio" R_0. At some point it drops below 1 and the epidemic ceases.

[4] "It's a matter of whether we are going to stay in the bee business," says David Hackenberg, the largest beekeeper in Pennsylvania. [8]

losses of up to 90% and some beekeepers experienced normal losses. It is necessary not only to use the most accurate numbers available, but also, for the future, to keep track of the network of interactions between various colonies. Perhaps the colonies not experiencing CCD did not have any interaction with those that did, so our analysis is tainted by the inclusion of these data. In other words, we must begin to use the tools of social network analysis or graph theory, and not just statistics.[5] Our analysis is also hampered by the fact that the census of colonies seems not to have been done monthly or even yearly. (It may be possible, with a more careful analysis, to use data regarding honey production[6] to help trace the disorder.) Our second recommendation is that a sample of colonies should be selected for inspection on a more frequent basis. While this might mean sacrificing the inspected colonies, more data might enable us to work out the rate of transmission or probability that an infected colony dies (essentially, β and γ). This kind of "continuous stream" of data might even enable us to work out information about the beginning of the outbreak, as well as the estimated number of deaths. It would also be helpful for insurers. (Moreover, while it makes sense to treat colonies as a unit, it might also be helpful to have a census of the number of bees, given that the number of bees in a colony can vary by an order of magnitude. [11])

It must be remembered that a model is only as accurate as its input, and, moreover, the assumptions may still be wildly wrong. For example, the predicted number of individuals that will eventually become infected in an epidemic can be far off the mark if one incorrectly assumes that the population is homogeneous.[7] Nevertheless, even imprecise mathematical models have been proven to be useful in other emergencies [14],[8] and, in this case, perhaps models can help sort out the various hypotheses regarding the cause.[9] [16] The urgency of the CCD outbreak justifies, in our view, these first tentative steps.

Acknowledgements.

The author would like to thank Professor Dennis vanEngelsdorp of the Pennsylvania Department of Agriculture for providing him with useful references. The author thanks Professor Glenn Webb of Vanderbilt University, Dr. Dominic van der Zypen of Allianz Zürich, Florida Apiary Inspection Assistant Chief Jerry Hayes, Dr. Dean Wilkening of Stanford University's Center for International Security and Cooperation, and Professor Lord Robert May of Oxford for their comments.

References

[1] "Buzz Off: Investigating Colony Collapse Disorder," *The Economist* (April 27, 2007).

[2] Robyn M. Underwood and Dennis vanEngelsdorp, "Colony Collapse Disorder: Have We Seen This Before?" *Bee Culture* (2007).

[3] Dennis vanEngelsdorp, Robyn Underwood, Dewey Caron, and Jerry Hayes, Jr., "An Estimate of Managed Colony Losses in the Winter of 2006-2007: A Report Commissioned by the Apiary Inspectors of America," *American Bee Journal* (July 2007).

[4] Roger A. Morse and Nicholas W. Calderone, "The Value of Honey Bees As Pollinators of U.S. Crops in 2000," *Bee Culture* **128** (2000).

[5] Renée Johnson, "Recent Honey Bee Colony Declines," Congressional Research Service Report for Congress (March 31, 2007).

[6] Marc Lipsitch *et al.*, "Transmission Dynamics and Control of Severe Acute Respiratory Syndrome," *Science* **300** (2003), 1966-1970.

[7] Norman T. J. Bailey, *The Mathematical Theory of Epidemics* (Charles Griffin and Company Limited, London, 1957).

[8] Rick Wills, "Apple Growers in Pa. Enter Crucial Pollination Period," *Pittsburgh Tribune-Review*, May 4, 2007.

[9] E. Lieberman, C. Hauert, and M. Nowak, "Evolutionary Dynamics on Graphs," *Nature* **433** (2005), 312-316.

[10] National Agricultural Statistics Service, "Honey" (February 28, 2007, February 28, 2006, February 28, 2005, February 27, 2004).

[11] Dennis vanEngelsdorp, personal communication (2007).

[12] Roy M. Anderson and Robert M. May, *Infectious Diseases of Humans: Dynamics and Control* (Oxford University Press, Oxford, 1991).

[13] Lord Robert May, personal communication (2007).

---

[5] See, for instance, [9].

[6] U.S. honey production was down 11% in 2006 from the year before and down 5% in 2005; but it was *up* 1% in 2004 and up 5% in 2003, suggesting that the outbreak may have begun earlier than 2006. [10]

[7] See [12], Figure 11.22, p. 272. Professor Lord May adds that the length of an epidemic can only be properly ascertained with a stochastic analysis. [13]

[8] For another "back of the envelope" calculation, see Box 1 of [15].

[9] Professor Webb also points out that some of the hypothesized causes of CCD *may* extinguish the populations in question.

[14] Carlos Castillo-Chavez, Carlos W. Castillo-Garsow, and Abdul-Aziz Yakubu, "Mathematical Models of Isolation and Quarantine," *Journal of the American Medical Association* **290** (2003), 2876-2877.

[15] M. J. Keeling, M. E. J. Woolhouse, R. M. May, G. Davies, and B. T. Grenfell, "Modelling Vaccination Strategies against Foot-and-Mouth Disease," *Nature* **421** (2003), 136-142.

[16] Glenn Webb, personal communication (2007).

Figure 1. A necessary and sufficient condition for the number of infected cases to continuerise.

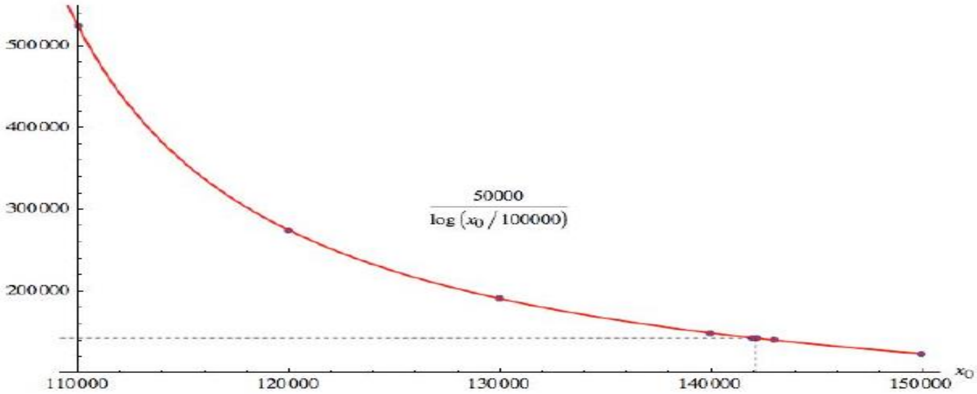

Figure 1

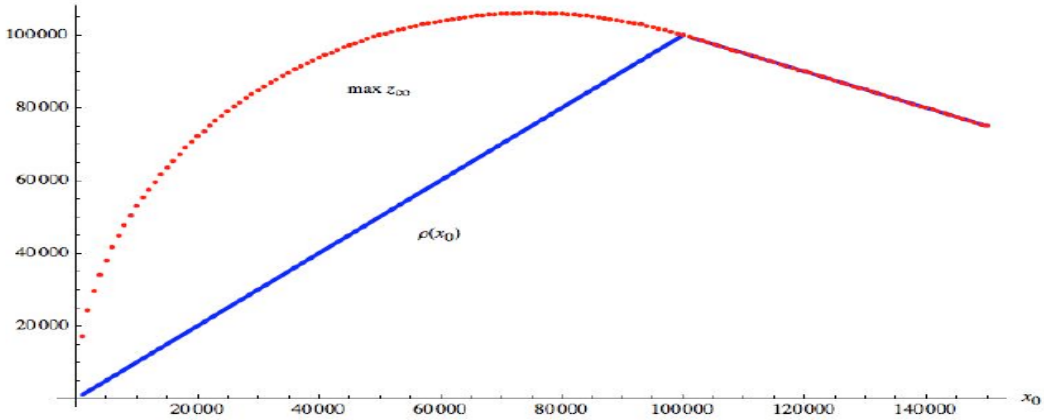

Figure 2. Estimating the total number of dead colonies.

Figure 2. Estimating the total number of dead colonies. x_0

50000 / ln (x_0 / 100000)

110000	524603
120000	274241
130000	190575
140000	148601
142000	142590
142150	142162
142152	142156
142153	**142153**
143000	139792
150000	123315

Table 1.

x_0	max z_∞	corresponding ρ
1000	17262.6765	1000
2000	24331.0501	2000
3000	29698.4848	3000
4000	34176.015	4000
5000	38078.8655	5000
6000	41569.2194	6000
7000	44743.7146	7000
8000	47665.5012	8000
9000	50378.5669	9000
10000	52915.0262	10000
11000	55299.1863	11000
12000	57549.9783	12000
13000	59682.4933	13000
14000	61708.9945	14000
15000	63639.6103	15000
16000	65482.8222	16000
17000	67245.8177	17000
18000	68934.7518	18000
19000	70554.9431	19000
20000	72111.0255	20000
21000	73607.0649	21000
22000	75046.6522	22000
23000	76432.9772	23000
24000	77768.8884	24000
25000	79056.9415	25000
26000	80299.4396	26000
27000	81498.4663	27000
28000	82655.9133	28000
29000	83773.5042	29000
30000	84852.8138	30000
31000	85895.2851	31000
32000	86902.2439	32000
33000	87874.9111	33000
34000	88814.4133	34000
35000	89721.7922	35000
36000	90598.0132	36000
37000	91443.9719	37000
38000	92260.5008	38000
39000	93048.3745	39000
40000	93808.3152	40000
41000	94540.9964	41000
42000	95247.0472	42000
43000	95927.0556	43000
44000	96581.5718	44000
45000	97211.1105	45000
46000	97816.1541	46000
47000	98397.1545	47000
48000	98954.535	48000
49000	99488.6928	49000
50000	100000	50000
51000	100488.805	51000
52000	100955.436	52000
53000	101400.197	53000
54000	101823.377	54000
55000	102225.242	55000
56000	102606.043	56000
57000	102966.014	57000
58000	103305.373	58000
59000	103624.322	59000
60000	103923.049	60000
61000	104201.727	61000

62000	104460.519	62000
63000	104699.57	63000
64000	104919.016	64000
65000	105118.98	65000
66000	105299.573	66000
67000	105460.893	67000
68000	105603.03	68000
69000	105726.061	69000
70000	105830.052	70000
71000	105915.06	71000
72000	105981.13	72000
73000	106028.298	73000
74000	106056.589	74000
75000	106066.017	75000
76000	106056.589	76000
77000	106028.298	77000
78000	105981.13	78000
79000	105915.06	79000
80000	105830.053	80000
81000	105726.061	81000
82000	105603.03	82000
83000	105460.893	83000
84000	105299.573	84000
85000	105118.98	85000
86000	104919.016	86000
87000	104699.57	87000
88000	104460.519	88000
89000	104201.728	89000
90000	103923.049	90000
91000	103624.322	91000
92000	103305.373	92000
93000	102966.014	93000
94000	102606.043	94000
95000	102225.242	95000
96000	101823.377	96000
97000	101400.197	97000
98000	100955.436	98000
99000	100488.805	99000
100000	100000	100000
101000	99500	99500
102000	99000	99000
103000	98500	98500
104000	98000	98000
105000	97500	97500
106000	97000	97000
107000	96500	96500
108000	96000	96000
109000	95500	95500
110000	95000	95000
111000	94500	94500
112000	94000	94000
113000	93500	93500
114000	93000	93000
115000	92500	92500
116000	92000	92000
117000	91500	91500
118000	91000	91000
119000	90500	90500
120000	90000	90000
121000	89500	89500
122000	89000	89000
123000	88500	88500
124000	88000	88000
125000	87500	87500
126000	87000	87000
127000	86500	86500
128000	86000	86000

129000	85500	85500
130000	85000	85000
131000	84500	84500
132000	84000	84000
133000	83500	83500
134000	83000	83000
135000	82500	82500
136000	82000	82000
137000	81500	81500
138000	81000	81000
139000	80500	80500
140000	80000	80000
141000	79500	79500
142000	79000	79000
143000	78500	78500
144000	78000	78000
145000	77500	77500
146000	77000	77000
147000	76500	76500
148000	76000	76000
149000	75500	75500
150000	75000	75000

Table 2. Estimating the total number of dead colonies.

```
20 F=OPENOUT "BEEDATA"

30 LET N=150000
40 REM IR=INCREMENT FOR R IN LOOP
50 IR=10
60 REM IX=INCREMENT FOR X IN LOOP
70 IX=1000
80 REM O=AMOUNT WE WILL ALLOW RHO TO GO ABOVE X0
90 O=0
100 INPUT "START X0 / FINISH X0 ",X1,X2
110 FOR X0=X1 TO X2 STEP IX
120   Z9=0
130   R9=0
140   REM LET R8=1 ONCE THE SEQUENCE OF Z'S STARTS DECREASING
150   R8=0
160   REM LET R7=1 IF THE SEQUENCE IS NOT UNIMODAL
170   R7=0
180   FOR R=IR TO X0+O STEP IR
190     Y0=N-X0
200     LET A1=((X0/R)-1)^2
210     LET A2=2*X0*Y0/(R^2)
220     LET A=SQR(A1+A2)
230     LET Z=((R^2)/X0)*((X0/R)-1+A)
240     IF Z<=Z9 THEN LET R8=1
250     IF Z>Z9 THEN LET R9=R
260     IF Z>Z9 AND R8=1 THEN LET R7=1
270     IF Z>Z9 THEN LET Z9=Z
290   NEXT R
300   PRINT#F,"X0=",STR$(X0)," ZMAX=",STR$(Z9)," RHO=",STR$(R9)
310   IF R7=1 THEN PRINT#F,"Z'S NOT UNIMODAL"
320   PRINT#F,CHR$(10)
330   NEXT X0
340   CLOSE#F
```

Figure 3. The BASIC code generating the data of Figure 2.

DOI : 10.24297/jaa.v7i2.6205